Modelos no-lineales para la teoría de muestreo

Anastasio, Magalí

2011

UNIVERSIDAD DE BUENOS AIRES

Facultad de Ciencias Exactas y Naturales

Departamento de Matemática

Modelos no-lineales para la teoría de muestreo

Tesis presentada para optar al título de Doctora de la Universidad de Buenos Aires en el área Ciencias Matemáticas

Magalí Anastasio

Director de tesis: Carlos Cabrelli

Consejero de estudios: Carlos Cabrelli

Buenos Aires, 2011.

Modelos no-lineales para la teoría de muestreo

(Resumen)

Un nuevo paradigma en la teoría de muestreo fue desarrollado recientemente. El clásico modelo lineal es reemplazado por un modelo no-lineal pero estructurado, que consiste en una unión de subespacios. Este es el enfoque natural para la nueva teoría de muestreo comprimido, señales con representaciones ralas y con tasa finita de innovación.

En esta tesis estudiamos algunos problemas relacionados con el proceso de muestreo en uniones de subespacios. Primero centramos nuestra atención en el problema de hallar una unión de subespacios que mejor aproxime a un conjunto finito de vectores. Utilizamos técnicas de reducción dimensional para disminuir los costos de algoritmos diseñados para hallar uniones de subespacios óptimos.

Luego estudiamos el problema de muestreo para señales que pertenecen a una unión de espacios invariantes por traslaciones enteras. Mostramos que las condiciones para la inyectividad y estabilidad del operador de muestreo son válidas en el caso general de espacios invariantes por traslaciones enteras generados por marcos de traslaciones en lugar de bases ortonormales.

A raíz del estudio de los problemas mencionados anteriormente, surgen dos cuestiones que están relacionadas con la estructura de los espacios invariantes por traslaciones enteras. La primera es si la suma de dos de estos espacios es un subespacio cerrado. Usando el ángulo de Friedrichs entre subespacios, obtenemos condiciones necesarias y suficientes para que la suma de dos espacios invariantes por traslaciones enteras sea cerrada.

En segundo lugar se estudian propiedades de invariancia de espacios invariantes por traslaciones enteras en varias variables. Presentamos condiciones necesarias y suficientes como para que un espacio invariante por traslaciones enteras sea invariante por un subgrupo cerrado de $\mathbb{R}^d$. Además probamos la existencia de espacios invariantes por traslaciones enteras que son exactamente invariantes para un subgrupo cerrado dado. Como aplicación, relacionamos la extra invariancia con el tamaño de los soportes de la transformada de Fourier de los generadores de los espacios.

Palabras Claves: muestreo; espacios invariantes por traslaciones enteras; marcos; bases de Riesz; operador Gramiano; fibras; reducción dimensional; desigualdades de concentración; ángulos entre subespacios.

Non-linear models in sampling theory

(Abstract)

A new paradigm in sampling theory has been developed recently. The classical linear model is replaced by a non-linear, but structured model consisting of a union of subspaces. This is the natural approach for the new theory of compressed sensing, representation of sparse signals and signals with finite rate of innovation.

In this thesis we study some problems concerning the sampling process in a union of subspaces. We first focus our attention in the problem of finding a union of subspaces that best explains a finite data of vectors. We use techniques of dimension reduction to avoid the expensiveness of algorithms which were developed to find optimal union of subspaces.

We then study the sampling problem for signals which belong to a union of shift-invariant spaces. We show that, the one to one and stability conditions for the sampling operator, are valid for the general case in which the subspaces are describe in terms of frame generators instead of orthonormal bases.

As a result of the study of the problems mentioned above, two questions concerning the structure of shift-invariant spaces arise. The first one is if the sum of two shift-invariant spaces is a closed subspace. Using the Friedrichs angle between subspaces, we obtain necessary and sufficient conditions for the closedness of the sum of two shift-invariant spaces.

The second problem involves the study of invariance properties of shift-invariant spaces in higher dimensions. We state and prove several necessary and sufficient conditions for a shift-invariant space to be invariant under a given closed subgroup of $\mathbb{R}^d$, and prove the existence of shift-invariant spaces that are exactly invariant for each given subgroup. As an application we relate the extra invariance to the size of support of the Fourier transform of the generators of the shift-invariant space.

Key words: sampling; shift-invariant spaces; frames; Riesz bases; Gramian operator; fibers; dimensionality reduction; concentration inequalities; angle between subspaces.

Agradecimientos

A Carlos por su guía, por haberme alentado y apoyado en todas mis decisiones, por su calidez, su confianza y por hacer tan ameno el trabajo junto a él.

A mis viejos por su apoyo incondicional y por el amor que me brindan día a día. A mis hermanas y cuñados por su cariño y por alegrarme los domingos. A mis sobris: Santi, Benji, Gian y Joaco por ser hermosos y porque los adoro con toda mi alma.

A las chicas: Ani, Caro y Vicky. Por ser unas amigas de fierro, por estar siempre, por escucharme y aconsejarme en todos los cafecitos. Me bancaron en todo, no sé qué hubiera hecho sin su apoyo y compañía.

A los chicos de la ofi: Turco, Roman, Mazzi, Dani, Martin S. y Chris, por su buena onda, por prestarnos yerba y por soportar nuestro cotorreo.

A quienes me acompañaron en toda esta etapa: Sole, Eli, Martin, Sigrid, Marian, Lean Delpe, Juli G., Isa, Lean L., San, Pablis, Colo, Dani R., Fer, Caro, Leo Z., Guille, Muro, Santi L., Rela, Dano, Vendra, Mer, Vero, Coti y Manu.

Al grupo de Análisis Armónico y Geometría Fractal, por los seminarios, congresos y charlas compartidas. A José Luis por haberme sugerido el estudio del tema que dio origen al capítulo 4.

A Akram, Úrsula y Vicky, porque fue un placer trabajar con ellos.

Contents

Introduction

A classical assumption in sampling theory is that the signals to be sampled belong to a single space of functions, for example the Paley-Wiener space of band-limited functions. In this case, the Kotelnikov-Shannon-Whittaker (KSW) theorem states that any function $f \in L^2(\mathbb{R})$ whose Fourier transform is supported within $[-\frac{1}{2}, \frac{1}{2}]$ can be completely reconstructed from its samples $\{f(k)\}_{k\in\mathbb{Z}}$. More specifically, if PW denotes the Paley Wiener space

$$PW = \left\{ f \in L^2(\mathbb{R}) \ : \ \operatorname{supp}(\widehat{f}) \subseteq \left[-\frac{1}{2}, \frac{1}{2}\right] \right\},$$

then $\{\operatorname{sinc}(\cdot - k)\}_{k\in\mathbb{N}}$ forms an orthonormal basis for PW, where $\operatorname{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$. Moreover, for all $f \in PW$ we have that $f(k) = \langle f, \operatorname{sinc}(\cdot - k)\rangle$ and

$$f(t) = \sum_{k\in\mathbb{Z}} f(k)\operatorname{sinc}(t-k), \tag{0.1}$$

with the series on the right converging uniformly on $\mathbb{R}$, as well as in $L^2(\mathbb{R})$.

The KSW theorem is fundamental in digital signal processing since it provides a method to convert an analog signal f to a digital signal $\{f(k)\}_{k\in\mathbb{Z}}$ and it also gives a reconstruction formula.

The Paley-Wiener space is invariant under integer translations, i.e. if $f \in PW$ then $f(\cdot - k) \in PW$ for any $k \in \mathbb{Z}$. The closed subspaces of $L^2(\mathbb{R}^d)$ which are invariant under integer translates are called shift-invariant spaces (SISs).

A shift-invariant space V is said to be generated by a set of functions $\{\varphi_j\}_{j\in J} \subseteq L^2(\mathbb{R}^d)$ if every function in V is a limit of linear combinations of integer shifts of the functions φ_j. That is,

$$V = \overline{\operatorname{span}}\{\varphi_j(\cdot - k) \ : \ j \in J, k \in \mathbb{Z}^d\},$$

where the closure is taken in the L^2-norm. We will say that the SIS is finitely generated if there exists a finite set of generators for the space. The Paley-Wiener space is an example of a shift-invariant space which is generated by $\varphi(t) = \operatorname{sinc}(t)$.

The function $\operatorname{sinc}(t)$ is well-localized in frequency but is poorly localized in time. This makes the formula (0.1) unstable in the presence of noise. To avoid this disadvantage other spaces of functions were considered as signal models. Mainly, shift-invariant spaces (SISs) generated by functions with better joint time-frequency localization or with compact support. One of the goals of the sampling problem in SISs is studying conditions on the generators of a SIS V in order that every function of V can be reconstructed from its values in a discrete sequence of samples as in the band-limited case. The sampling problem for SISs was thoroughly treated in [AG01, Sun05, Wal92, ZS99].

Assume now that we want to sample the signals in a finite set $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq L^2(\mathbb{R}^d)$ and that they do not belong to a computationally tractable SIS. For example, if the cardinality of the data set m is large, the SIS generated by $\mathcal{F}$ contains all the data, but it is too large to be an appropriate model for use in applications. So, a space with less generators would be more suitable. In order to model the set $\mathcal{F}$ by a manageable SIS we consider the following problem: given $k << m$, the goal is to find a SIS with k generators that best models the data set $\mathcal{F}$. That is, we would like to find a SIS V_0 generated by at most k functions that is closest to the set $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq L^2(\mathbb{R}^d)$ in the sense that

$$V_0 = \text{argmin}_{V \in \mathcal{L}_k} \sum_{i=1}^{m} \|f_i - P_V f_i\|^2_{L^2}, \tag{0.2}$$

where $\mathcal{L}_k$ is the set of all the SISs generated by at most k functions, and P_V is the orthogonal projection from $L^2(\mathbb{R}^d)$ onto V.

In [ACHM07] the authors proved the existence of an optimal space that satisfies (0.2), they gave a way to construct the generators of such space and estimated the error between the optimal space and the data set $\mathcal{F}$. To obtain their results they reduced the problem to the finite dimensional problem of finding a subspace of dimension at most k that best approximates a finite data set of vectors in the Hilbert space $\ell^2(\mathbb{Z}^d)$. This last problem can be solved by an extension of the Eckart-Young's Theorem. We will review some of these results in Chapter 2.

Recently, a new approach for the sampling theory has been developed. The classical linear model is replaced by a non-linear, but structured model consisting of a union of subspaces. More specifically, Lu and Do [LD08] extended the sampling problem assuming that the signals to be sampled belong to a union of subspaces instead of a single subspace. To understand the importance of this new approach let us introduce some examples.

i) *Compressed sensing:* In the compressed sensing setting ([CRT06], [CT06], [Don06]) the signal $x \in \mathbb{R}^N$ is assumed to be sparse in an orthonormal basis of $\mathbb{R}^N$. That is, given $\Phi = \{\phi_j\}_{j=1}^N$ an orthonormal basis for $\mathbb{R}^N$, x has at most k non-zero coefficients in Φ, where $k << N$. In other words, if $\theta_j(x) = \langle x, \phi_j \rangle$, then

$$\|\theta(x)\|_0 := \#\{\, j \in \{1, \dots, N\} \ : \ \theta_j(x) \neq 0\} \leq k.$$

The sparse signals live in the union of k-dimensional subspaces, given by

$$\bigcup_{1 \leq j_1 < \dots < j_k \leq N} V_{j_1, \dots, j_k}, \tag{0.3}$$

with $V_{j_1, \dots, j_k} = \text{span}\{\phi_{j_1}, \dots, \phi_{j_k}\}$.

ii) *Blind Spectral Support:* Let $[\omega_0, \omega_0 + N] \subseteq \mathbb{R}$ be an interval which is partitioned into N equal intervals $C_j = [\omega_0 + j, \omega_0 + j + 1]$ for $0 \leq j \leq N - 1$. Assume we have a function $f \in L^2(\mathbb{R})$ whose Fourier transform is supported in at most k intervals $C_{j_1}, \dots, C_{j_k}$ (with $k << N$), but we do not know the indices $j_1, \dots, j_k$.

If we define

$$V_{j_1,\dots,j_k} := \{g \in L^2(\mathbb{R}) \ : \ \text{supp}(\widehat{g}) \subseteq C_{j_1} \cup \dots \cup C_{j_k}\},$$

then the function f belongs to the union of subspaces

$$\bigcup_{1 \leq j_1 < \dots < j_k \leq N} V_{j_1,\dots,j_k}.$$

This class of signals are called multiband signals with unknown spectral support (see [FB96]).

iii) *Stream of Diracs:* Given $k \in \mathbb{N}$ consider the stream of k Diracs

$$x(t) = \sum_{j=1}^{k} c_j \delta(t - t_j),$$

where $\{t_j\}_{j=1}^k$ are unknown locations and $\{c_j\}_{j=1}^k$ are unknown weights.

If the k locations are fixed, then the signals live in a k-dimensional subspace. Thus, they live in an infinite union of k-dimensional subspaces.

These signals have $2k$ degrees of freedom: k for the weights and k for the locations of the Diracs. Sampling theorems for this class of signals have been studied in the framework of signals with *finite rate of innovation*. They receive this name since they have a finite number of degrees of freedom per unit of time. In [VMB02] it was proved that only $2k$ samples are sufficient to reconstruct these signals .

Note that if we considered the signals from a union of subspaces as elements of the subspace generated by the union of these spaces, we would be able to apply the linear sampling techniques for signals lying in only one subspace. But the problem is that we would not be having into account an additional information about the signals. For example, in the case of k-sparse signals (see Example i) from above) we only need $2k$ samples to reconstruct a signal $x \in \mathbb{R}^N$, k for the support of the coefficients $\theta(x)$ and k for the value of the non-zero coefficients. On the other side, if we considered the signal x as an element of the subspace generated by the union (0.3) (i.e. $\mathbb{R}^N$) we would need N samples to reconstruct it.

The model proposed by Lu y Do [LD08] in which the signals live in a union of subspaces instead of a single vector space represented a new paradigm for signal sampling and reconstruction. Since for each class of signals the starting point of this new theory is the knowledge of the signal space, the first step for implementing the theory is to find an appropriate signal model from a set of observed data. In [ACM08] the authors studied the problem of finding a union of subspaces $\cup_i V_i \subseteq \mathcal{H}$ that best explains the data $\mathcal{F} = \{f_1, \ldots, f_m\}$ in a Hilbert space $\mathcal{H}$ (finite or infinite dimensional). They proved that if the subspaces V_i belong to a family of closed subspaces C which satisfies the so called Minimum Subspace Approximation Property (MSAP), an optimal solution to the non-linear subspace modeling problem that best fit the data exists, and algorithms to find these subspaces were developed.

In some applications the model is a finite union of subspaces and $\mathcal{H}$ is finite dimensional. Once the model is found, the given data points can be clustered and classified according to their distances from the subspaces, giving rise to the so called *subspace clustering problem* (see e.g., [CL09] and the references therein). Thus a dual problem is to first find a "best partition" of the data. Once this partition is obtained, the associated optimal subspaces can be easily found. In any case, the search for an optimal partition or optimal subspaces usually involves heavy computations that dramatically increases with the dimensionality of $\mathcal{H}$. Thus, one important feature is to map the data into a lower dimensional space, and solve the transformed problem in this lower dimensional space. If the mapping is chosen appropriately, the original problem can be solved exactly or approximately using the solution of the transformed data.

In Chapter 2, we concentrate on the non-linear subspace modeling problem when the model is a finite union of subspaces of $\mathbb{R}^N$ of dimension $k << N$. Our goal is to find transformations from a high dimensional space to lower dimensional spaces with the aim of solving the subspace modeling problem using the low dimensional transformed data. We find the optimal data partition for the transformed data and use this partition for the original data to obtain the subspace model associated to this partition. We then estimate the error between the model thus found and the optimal subspaces model for the original data.

Once the union of subspaces that best explains a data set is found, it is interesting to study the sampling process for signals which belong to this kind of models. The sampling results which are applied for signals lying in a single subspace are not longer valid for signals in a union of subspaces since the linear structure is lost. The approach of Lu and Do [LD08] had a great impact in many applications in signal processing, in particular in the emerging theory of compressed sensing [CT06], [CRT06], [Don06] and signals with finite rate of innovations [VMB02].

To understand the problem, let us now describe the process of sampling in a union of subspaces. Assume that $\mathcal{X}$ is a union of subspaces from some Hilbert space $\mathcal{H}$ and a signal s is extracted from $\mathcal{X}$. We take some measurements of that signal. These measurements can be thought of as the result of the application of a series of functionals $\{\varphi_\alpha\}_\alpha$ to our signal s. The problem is then to reconstruct the signal using only the measurements $\{\varphi_\alpha(s)\}_\alpha$ and some description of the subspaces in $\mathcal{X}$. The series of functionals define an operator, *the sampling operator*, acting on the ambient space $\mathcal{H}$ and taking values in a suitable sequence space. Under some hypothesis on the structure of the subspaces, Lu and Do [LD08] found necessary and sufficient conditions on these functionals in order for the sampling operator to be stable and one-to-one when restricted to the union of the subspaces. These conditions were obtained in two settings. In the euclidean space and in $L^2(\mathbb{R}^d)$. In this latter case the subspaces considered were finitely generated shift-invariant spaces.

Blumensath and Davies [BD09] studied the problem of sampling in union of subspaces in the finite dimensional case, extending some of the results in Lu and Do [LD08]. They applied their results to compressed sensing models and sparse signals. In [EM09], Eldar developed a general framework for robust and efficient recovery of a signal from a given set of samples. The signal is a finite length vector that is sparse in some given basis and is assumed to lie in a union of subspaces.

There are two technical aspects in the approach of Lu and Do that restrict the applicability of their results in the shift-invariant space case. The first one is due to the fact that the conditions are obtained in terms of Riesz bases of translates of the SISs involved, and it is well known that not every SIS has a Riesz basis of translates (see Example 1.5.11). The second one is that the approach is based upon the sum of every two of the SISs in the union. The conditions on the sampling operator are then obtained using fiberization techniques on that sum. This requires that the sum of each of two subspaces is a closed subspace, which is not true in general.

In Chapter 3 we obtain the conditions for the sampling operator to be one-to-one and stable in terms of frames of translates of the SISs instead of orthonormal bases. This extends the previous results to arbitrary SISs and in particular removes the first restriction mentioned above. It is very important to have conditions based on frames, specially for applications, since frames are more flexible and simpler to construct. Frames of translates for shift-invariant spaces with generators that are smooth and with good decay can be easily obtained.

In Chapter 3, we give necessary and sufficient conditions for the stability of the sampling operator in a union of arbitrary SISs. We also show that, without the assumption of the closedness of the sum of every two of the SISs in the union, we can only obtain sufficient conditions for the injectivity of the sampling operator.

Using known results from the theory of SISs, in Chapter 4 we obtain necessary and sufficient conditions for the closedness of the sum of two shift-invariant spaces. As a consequence, we determine families of subspaces on which the conditions for the injectivity of the sampling operator are necessary and sufficient.

An important and interesting question in the study of SISs is whether they have the property to be invariant under translations other than integers. A limit case is when the space is invariant under translations by all real numbers. In this case the space is called *translation invariant*. However there exist shift-invariant spaces with some *extra* invariance that are not necessarily translation invariant. That is, there are some intermediate cases between shift-invariance and translation invariance. The question is then, how can we identify them?

Recently, Hogan and Lakey defined the *discrepancy* of a shift-invariant space as a way to quantify the *non-translation invariance* of the subspace, (see [HL05]). The discrepancy measures how far a unitary norm function of the subspace, can move away from it, when translated by non integers. A translation invariant space has discrepancy zero.

In another direction, Aldroubi et al, (see [ACHKM10]) studied shift-invariant spaces of $L^2(\mathbb{R})$ that have some extra invariance. They show that if V is a shift-invariant space, then its *invariance set*, is a closed additive subgroup of $\mathbb{R}$ containing $\mathbb{Z}$. The invariance set associated to a shift-invariant space is the set M of real numbers satisfying that for each $p \in M$ the translations by p of every function in V, belongs to V. As a consequence, since every additive subgroup of $\mathbb{R}$ is either discrete or dense, there are only two possibilities left for the extra invariance. That is, either V is invariant under translations by the group $(1/n)\mathbb{Z}$, for some positive integer n (and not invariant under any bigger subgroup) or it is translation invariant. They found different characterizations, in terms of the Fourier transform, of when a shift invariant space is $(1/n)\mathbb{Z}$-invariant.

A natural question arises in this context. Are the characterizations of extra invariance that hold on the line, still valid in several variables?

The invariance set $M \subseteq \mathbb{R}^d$ associated to a shift-invariant space V, that is, the set of vectors that leave V invariant when translated by its elements, is again, as in the 1-dimensional case, a closed subgroup of $\mathbb{R}^d$ (see Proposition 5.2.1). The problem of the extra invariance can then be reformulated as finding necessary and sufficient conditions for a shift-invariant space to be invariant under a closed additive subgroup $M \subseteq \mathbb{R}^d$ containing $\mathbb{Z}^d$.

The main difference here with the one dimensional case, is that the structure of the subgroups of $\mathbb{R}^d$ when d is bigger than one, is not as simple.

The results obtained for the 1-dimensional case translate very well in the case in which the invariance set M is a lattice, (i.e. a discrete group) or when M is dense, that is $M = \mathbb{R}^d$. However, there are subgroups of $\mathbb{R}^d$ that are neither discrete nor dense. So, can there exist shift-invariant spaces which are M-invariant for such a subgroup M and are not translation invariant?

In Chapter 5 we study the extra invariance of shift-invariant spaces in higher dimensions. We obtain several characterizations paralleling the 1-dimensional results. In addition our results show the existence of shift-invariant spaces that are *exactly* M-invariant for every closed subgroup $M \subseteq \mathbb{R}^d$ containing $\mathbb{Z}^d$. By 'exactly M-invariant' we mean that they are not invariant under any other subgroup containing M. We apply our results to obtain estimates on the size of the support of the Fourier transform of the generators of the space.

At the end of Chapter 5 we also give a brief description of the generalization of the extra invariance results to the context of locally compact abelian (LCA) groups.

Thesis outline

Chapter 1 contains the notation and some preliminary tools used throughout this thesis. We present basic definitions and results regarding frames and Riesz bases in Hilbert spaces. We give some characterizations and properties of shift-invariant spaces. We also define the range function and the notion of fibers for shift-invariant spaces.

In Chapter 2 we study the problem of finding models which best explain a finite data set of signals. We first review some results about finding a subspace that is closest to a given finite data set. We then study the general case of unions of subspaces which best approximate a set of signals. The results are proved in a general setting and then applied to the case of low dimensional subspaces of $\mathbb{R}^N$ and to infinite dimensional shift-invariant spaces of $L^2(\mathbb{R}^d)$.

For the euclidean case $\mathbb{R}^N$, the problem of optimal union of subspaces increases dramatically with the dimension N. In Chapter 2, we study a class of transformations that map the problem into another one in lower dimension. We use the best model in the low dimensional space to approximate the best solution in the original high dimensional space. We then estimate the error produced between this solution and the optimal solution in the high dimensional space.

The purpose of Chapter 3 is the extension of the results of [LD08] for sampling in a union of subspaces for the case that the subspaces in the union are arbitrary shift-invariant spaces. We describe the subspaces by means of frame generators instead of orthonormal bases. We give necessary and sufficient conditions for the stability of the sampling operator in a union of arbitrary SISs. We also show that, without the assumption of the closedness of the sum of every two of the SISs in the union, we can only obtain sufficient conditions for the injectivity of the sampling operator.

In Chapter 4 we obtain necessary and sufficient conditions for the closedness of the sum of two shift-invariant spaces in terms of the Friedrichs angle between subspaces. As a consequence of this, we determine families of subspaces on which the conditions for injectivity of the sampling operator are necessary and sufficient.

Finally, in Chapter 5 we study invariance properties of shift-invariant spaces in higher dimensions. We state and prove several necessary and sufficient conditions for a shift-invariant space to be invariant under a given closed subgroup of $\mathbb{R}^d$, and prove the existence of shift-invariant spaces that are exactly invariant for each given subgroup. As an application we relate the extra invariance to the size of support of the Fourier transform of the generators of the shift-invariant space. We extend recent results obtained for the case of one variable to several variables. We also give in this chapter a brief description of the extra invariance results obtained in [ACP10a] for the general case of locally compact abelian (LCA) groups.

Publications from this thesis

The new results in Chapter 2, 3, 4, and 5 have originated the following publications:

- A. Aldroubi, M. Anastasio, C. Cabrelli and U. M. Molter, *A dimension reduction scheme for the computation of optimal unions of subspaces*, Sampl. Theory Signal Image Process., **10**(1-2), 2011, 135–150. (Chapter 2)

- M. Anastasio and C. Cabrelli, *Sampling in a union of frame generated subspaces*, Sampl. Theory Signal Image Process, **8**(3), 2009, 261–286. (Chapter 3 and Chapter 4)

- M. Anastasio, C. Cabrelli and V. Paternostro, *Invariance of a Shift-Invariant Space in Several Variables*, Complex Analysis and Operator Theory, **5**(4), 2011, 1031–1050. (Chapter 5)

- M. Anastasio, C. Cabrelli and V. Paternostro, *Extra invariance of shift-invariant spaces on LCA groups*, J. Math. Anal. Appl., **370**(2), 2010, 530–537. (Chapter 5)

Theorem 1.2.1. *Every separable Hilbert space $\mathcal{H}$ has an orthonormal basis.*

Example 1.2.2. Let $\{e_j\}_{j\in J}$ be the sequence in $\ell^2(J)$ defined by $(e_j)_i = \delta_{i,j}$ for every $i, j \in J$. Then $\{e_j\}_{j\in J}$ is an orthonormal basis for $\ell^2(J)$ and it is called the *canonical basis*.

We will now introduce the definition of Riesz bases. We will see later that they can be considered as a generalization of orthonormal bases.

Definition 1.2.3. A sequence $\{x_j\}_{j\in J}$ in $\mathcal{H}$ is a *Riesz basis* for $\mathcal{H}$ if it is complete in $\mathcal{H}$ and there exist constants $0 < \alpha \leq \beta < +\infty$ such that

$$\alpha \sum_{j\in J} |c_j|^2 \leq \Big\| \sum_{j\in J} c_j x_j \Big\|^2 \leq \beta \sum_{j\in J} |c_j|^2 \quad \forall \{c_j\}_{j\in J} \in \ell^2(J). \tag{1.2}$$

The following proposition states a relationship between Riesz bases, bases and orthonormal bases.

Proposition 1.2.4. *Let $\{x_j\}_{j\in J}$ be a sequence in $\mathcal{H}$. The following statements are equivalent.*

i) $\{x_j\}_{j\in J}$ *is a Riesz basis for* $\mathcal{H}$.

ii) $\{x_j\}_{j\in J}$ *is a basis for* $\mathcal{H}$*, and*

$$\sum_{j\in J} c_j x_j \quad \text{converges if and only if} \quad \{c_j\}_{j\in J} \in \ell^2(J).$$

iii) There exist a bounded linear operator $T : \mathcal{H} \to \mathcal{H}$ and an orthonormal basis $\{e_j\}_{j\in J}$ for $\mathcal{H}$ such that $T(e_j) = x_j$ for all $j \in J$.

Taking T as the identity operator in item iii), we have that all orthonormal bases are Riesz bases. The following proposition states that the converse is true when the constants of the inequality (1.2) are equal to one.

Proposition 1.2.5. *Let $\{x_j\}_{j\in J}$ be a sequence in $\mathcal{H}$. Then, $\{x_j\}_{j\in J}$ is a an orthonormal basis if and only if it is a Riesz basis with constants $\alpha = \beta = 1$.*

We will now introduce the concept of frames which can be seen as a generalization of Riesz bases.

Definition 1.2.6. A sequence $\{x_j\}_{j\in J}$ in $\mathcal{H}$ is a *frame* for $\mathcal{H}$ if there exist constants $0 < \alpha \leq \beta < +\infty$ such that

$$\alpha \|h\|^2 \leq \sum_{j\in J} |\langle h, x_j\rangle|^2 \leq \beta \|h\|^2 \quad \forall h \in \mathcal{H}. \tag{1.3}$$

The constants α, β are called *frame bounds*. If $\{x_j\}_{j\in J}$ satisfies the right inequality from (1.3) we will call it a *Bessel sequence*.

The frame is *tight* if $\alpha = \beta$. A *Parseval frame* is a tight frame with constants $\alpha = \beta = 1$.

The frame is *exact* if it ceases to be a frame whenever any single element is deleted from the sequence.

We will say that $\{x_j\}_{j\in J}$ is a *frame sequence* if it is a frame for the subspace $\overline{\text{span}}\{x_j\}_{j\in J}$.

Remark 1.2.7. Although a Parseval frame satisfy the Parseval's identity (1.1), it might not be an orthogonal system. In fact, it is orthogonal if and only if every element of the set has unitary norm. A simple example is the family $X = \{\frac{1}{\sqrt{2}}e_1, \frac{1}{\sqrt{2}}e_1, e_n\}_{n\geq 2}$ where $\{e_n\}_{n\in\mathbb{N}}$ is an orthonormal basis for an infinite dimensional Hilbert space $\mathcal{H}$. X is a Parseval frame that is not orthogonal and it is not even a basis.

As we have mentioned above, frames can be considered as a generalization of Riesz bases. The next proposition gives necessary and sufficient conditions in order for a frame to be a Riesz basis.

Proposition 1.2.8. *Let $\{x_j\}_{j\in J}$ be a sequence in $\mathcal{H}$. Then $\{x_j\}_{j\in J}$ is a Riesz basis if and only if it is an exact frame for $\mathcal{H}$.*

We will now introduce some operators which play a crucial role in the theory of sampling.

Definition 1.2.9. If $X = \{x_j\}_{j\in J}$ is a Bessel sequence in $\mathcal{H}$, we define the *analysis operator* as

$$B_X : \mathcal{H} \to \ell^2(J), \quad B_X h = \{\langle h, x_j\rangle\}_{j\in J}.$$

The adjoint of B is the *synthesis operator*, given by

$$B_X^* : \ell^2(J) \to \mathcal{H}, \quad B_X^* c = \sum_{j\in J} c_j x_j.$$

The Bessel condition guarantees the boundedness of B_X and as a consequence, that of B_X^*.

By composing B_X^* and B_X, we obtain the *frame operator*

$$S : \mathcal{H} \to \mathcal{H}, \quad Sh := B_X^* B_X h = \sum_{j\in J} \langle h, x_j\rangle x_j.$$

Frame sequences can be characterized through its synthesis operators as it is stated in the following proposition.

Proposition 1.2.10. *A sequence $X = \{x_j\}_{j\in J}$ in $\mathcal{H}$ is a frame sequence if and only if the synthesis operator B_X^* is well-defined on $\ell^2(J)$ and has closed range.*

As a consequence of the previous proposition, if $\{x_j\}_{j\in J}$ is a frame for the subspace $V := \overline{\text{span}}\{x_j\}_{j\in J}$, then

$$V = \left\{ \sum_{j\in J} c_j x_j \; : \; \{c_j\} \in \ell^2(J) \right\}. \tag{1.4}$$

Using this, it is possible to construct for any infinite dimensional separable Hilbert space a Bessel sequence which is complete in $\mathcal{H}$ and it is not a frame sequence.

Example 1.2.11. Let $\{e_n\}_{n\in\mathbb{N}}$ be an orthonormal basis for an infinite dimensional separable Hilbert space $\mathcal{H}$ and define $f_n = e_n + e_{n+1}$ for $n \in \mathbb{N}$. This is a Bessel sequence since, for $h \in \mathcal{H}$,

$$\begin{aligned}\sum_{k\in\mathbb{N}} |\langle h, e_n + e_{n+1}\rangle|^2 &= \sum_{n\in\mathbb{N}} |\langle h, e_n\rangle + \langle h, e_{n+1}\rangle|^2 \\ &\leq 2\sum_{n\in\mathbb{N}} |\langle h, e_n\rangle|^2 + 2\sum_{n\in\mathbb{N}} |\langle h, e_{n+1}\rangle|^2 \\ &\leq 4\|h\|^2.\end{aligned}$$

We also have that $\{f_n\}_{n\in\mathbb{N}}$ is complete in $\mathcal{H}$ because if there exists $h \in \mathcal{H}$ such that $\langle h, e_n + e_{n+1}\rangle = 0$ for all $n \in \mathbb{N}$, then $\langle h, e_n\rangle = -\langle h, e_{n+1}\rangle$ for all n. Thus $|\langle h, e_n\rangle|$ is constant. Using the Parseval's identity (1.1), we conclude that $\langle h, e_n\rangle = 0$ for all n, so $h = 0$. Therefore, $\{f_n\}_{n\in\mathbb{N}}$ is complete in $\mathcal{H}$.

Observe that for $h = e_1 \in \mathcal{H}$ there exists no $\{c_n\} \in \ell^2(\mathbb{N})$ such that $h = \sum_{n\in\mathbb{N}} c_n f_n$. By (1.4) this proves that $\{f_n\}_{n\in\mathbb{N}}$ is a Bessel sequence which is not a frame sequence.

Remark 1.2.12. Note that from Proposition 1.2.10, any finite sequence $\{x_1, \ldots, x_m\}$ in a Hilbert space $\mathcal{H}$ is a frame for the closed subspace $V = \text{span}\{x_1, \ldots, x_m\}$.

The next proposition announces important properties about the frame operator. It also states one of the most important results about frames which is that every element in $\mathcal{H}$ has a representation as an infinite linear combination of the elements of the frame.

Recall that a series $\sum_{j\in J} x_j$ is unconditionally convergent if $\sum_{j\in J} x_{\sigma(j)}$ converges for every permutation σ of J.

Proposition 1.2.13. *If $X = \{x_j\}_{j\in J}$ is a frame for $\mathcal{H}$ with frame bounds α,β, then the following statements hold.*

i) The frame operator S is bounded, invertible, self-adjoint, positive, and satisfies

$$\alpha\|h\|^2 \leq \langle Sh, h\rangle \leq \beta\|h\|^2 \quad \forall h \in \mathcal{H}.$$

ii) $\{S^{-1}x_j\}_{j\in J}$ is a frame for $\mathcal{H}$, with frame bounds $0 < \beta^{-1} \leq \alpha^{-1}$.

iii) The following series converge unconditionally for each $h \in \mathcal{H}$

$$h = \sum_{j\in J}\langle h, S^{-1}x_j\rangle x_j = \sum_{j\in J}\langle h, x_j\rangle S^{-1}x_j.$$

iv) If the frame is tight, then $S = \alpha I$ and $S^{-1} = \alpha^{-1}I$.

Let $\{x_j\}_{j\in J}$ be a frame for $\mathcal{H}$, a Bessel sequence $\{y_j\}_{j\in J}$ is said to be a *dual frame* of $\{x_j\}_{j\in J}$ if

$$h = \sum_{j\in J}\langle h, y_j\rangle x_j = \sum_{j\in J}\langle h, x_j\rangle y_j \quad \forall\, h \in \mathcal{H}.$$

By Proposition 1.2.13, we have that $\{S^{-1}x_j\}_{j\in J}$ is a dual frame of $\{x_j\}_{j\in J}$ which is called the *canonical dual*. When $\{x_j\}_{j\in J}$ is a Riesz basis, the unique dual is the canonical dual.

A frame which is not a Riesz basis is said to be *overcomplete*. When the frame is overcomplete there exist dual frames which are different from the canonical dual.

As a consequence of item iii) of Proposition 1.2.13 every element $h \in \mathcal{H}$ has a representation of the form $h = \sum_{j\in J} c_j x_j$ with coefficients $c_j = \langle h, S^{-1}x_j\rangle$. If $\{x_j\}_{j\in J}$ is an overcomplete frame, the representation given before is not unique, that is, there are other coefficients $\{c'_j\}_{j\in J} \in \ell^2(J)$ for which $h = \sum_{j\in J} c'_j x_j$.

Note that by Theorem 1.2.1, every closed subspace of a separable Hilbert space has an orthonormal basis. A question that arises then is why studying frames if in every closed subspace there exists an orthonormal basis. One of the advantages of frames is their redundancy. If the frame is overcomplete there are several choices for the coefficients c_j in the representation of an element $h \in \mathcal{H}$ as $h = \sum_{j\in J} c_j x_j$. Thus, due to this redundancy, if some of the coefficients are missing or unknown it is still possible to recover the signal from the incomplete data.

Another application which shows the importance of working with frames will be shown in future sections. We will study in this chapter the structure of closed subspaces of $L^2(\mathbb{R}^d)$ which are invariant under integer translations (shift-invariant spaces). We will show that every shift-invariant subspace has a frame formed by integer translates of functions. We will also prove that there exist shift-invariant subspaces that do no have Riesz bases of translates. Thus, for these spaces is essential to work with frames instead of bases.

1.3 The Gramian operator

In this section we will introduce the Gramian operator associated to a Bessel sequence. We will see that there exists a relationship between the spectrum of this operator and the fact that the sequence is a frame.

Definition 1.3.1. Suppose $X = \{x_j\}_{j\in J}$ is a Bessel sequence in $\mathcal{H}$ and B_X is the analysis operator. The Gramian of the system X is defined by

$$G_X : \ell^2(J) \to \ell^2(J), \quad G_X := B_X B_X^*.$$

We identify G_X with its matrix representation.

$$(G_X)_{j,k} = \langle x_k, x_j\rangle \quad \forall\, j,k \in J.$$

Given a Hilbert space $\mathcal{K}$ and a bounded linear operator $T : \mathcal{K} \to \mathcal{K}$, we will denote by $\sigma(T)$ the spectrum of T, that is

$$\sigma(T) = \{\lambda \in \mathbb{C} \,:\, \lambda I - T \text{ is not invertible }\},$$

where I denotes the identity operator of $\mathcal{K}$.

The following lemmas will be useful to prove a property which relates a frame sequence with the spectrum of its Gramian.

Lemma 1.3.2. *Let $T : \mathcal{K} \to \mathcal{K}$ be a positive semi-definite self-adjoint operator and $\widetilde{T} : \ker(T)^{\perp} \to \ker(T)^{\perp}$ the restriction of T to $\ker(T)^{\perp}$. Assume $0 < \alpha \leq \beta < +\infty$. The following conditions are equivalent:*

i) $\sigma(T) \subseteq \{0\} \cup [\alpha, \beta]$

ii) $\sigma(\widetilde{T}) \subseteq [\alpha, \beta]$.

iii) $\alpha \|x\|^2 \leq \langle Tx, x\rangle \leq \beta \|x\|^2 \quad \forall x \in \ker(T)^{\perp}$

Proof. We will first prove that i) implies ii). Assume that $\sigma(T) \subseteq \{0\} \cup [\alpha, \beta]$. Given $\lambda \in \sigma(\widetilde{T})$, since $\sigma(\widetilde{T}) \subseteq \sigma(T)$, it follows that $\lambda \in \{0\} \cup [\alpha, \beta]$. If $\lambda = 0$, then λ is an isolated point of $\sigma(\widetilde{T})$. Using that T is self-adjoint, we have that $\widetilde{T}$ is self-adjoint. Thus, λ must be an eigenvalue of $\widetilde{T}$ (see [Con90]). Hence, $\ker(\widetilde{T}) \neq 0$, which is a contradiction.

The deduction of i) from ii) is left to the reader.

As $\widetilde{T} : \ker(T)^{\perp} \to \ker(T)^{\perp}$ is self-adjoint, we have that $\sigma(\widetilde{T}) \subseteq [\alpha, \beta]$ if and only if

$$\alpha \|x\|^2 \leq \langle \widetilde{T}x, x\rangle \leq \beta \|x\|^2 \quad \forall x \in \ker(T)^{\perp}.$$

Thus, the equivalence between ii) and iii) is straightforward.

□

The next lemma is proved in [Chr03, Lemma 5.5.4].

Lemma 1.3.3. *Let $X := \{x_j\}_{j \in J} \subseteq \mathcal{H}$ be a Bessel sequence, then X is a frame sequence with constants α and β if and only if the synthesis operator B_X^* satisfies*

$$\alpha \|c\|^2 \leq \|B_X^* c\|^2 \leq \beta \|c\|^2 \quad \forall c \in \ker(B_X^*)^{\perp}.$$

The following is a well known property, its proof can be deduced from Lemma 1.3.2 and Lemma 1.3.3.

Theorem 1.3.4. *Let $X := \{x_j\}_{j \in J} \subseteq \mathcal{H}$ be a Bessel sequence, then X is a frame sequence with constants α and β if and only if*

$$\sigma(G_X) \subseteq \{0\} \cup [\alpha, \beta].$$

We also have the property from below which relates the dimension of the subspace spanned by a finite set of vectors with the rank of the Gramian matrix.

Proposition 1.3.5. *Let $X = \{x_1, \ldots, x_m\}$ be a finite set of vectors in $\mathcal{H}$. Then*

$$\operatorname{rank}[G_X] = \dim(\operatorname{span}\{x_1, \ldots, x_m\}).$$

Proof. Since $G_X = B_X B_X^* \in \mathbb{C}^{m \times m}$, we have that

$$\operatorname{rank}[G_X] = \dim(\operatorname{range}(B_X^*)) = \dim(\operatorname{span}\{x_1, \ldots, x_m\}).$$

□

1.4 Shift-invariant spaces

In this section we introduce some definitions and basic properties of shift-invariant spaces. For a detailed treatment of the subject see [dBDR94, dBDVR94, Bow00, Hel64, RS95] and the references therein.

Definition 1.4.1. A closed subspace $V \subseteq L^2(\mathbb{R}^d)$ is a *shift-invariant space* (SIS) if $f \in V$ implies $t_k f \in V$ for any $k \in \mathbb{Z}^d$, where t_k is the translation by k.

Given Φ a set of functions in $L^2(\mathbb{R}^d)$, we denote by $E(\Phi)$ the set,

$$E(\Phi) := \{t_k\varphi : k \in \mathbb{Z}^d, \varphi \in \Phi\}.$$

When $\Phi = \{\varphi\}$, we will write $E(\varphi)$.

The SIS generated by Φ is

$$V(\Phi) := \overline{\operatorname{span}}(E(\Phi)) = \overline{\operatorname{span}}\{t_k\varphi : \varphi \in \Phi,\ k \in \mathbb{Z}^d\}.$$

We call Φ a *set of generators* for $V(\Phi)$. When $\Phi = \{\varphi\}$, we simply write $V(\varphi)$.

The *length* of a shift-invariant space V is the cardinality of a smallest generating set for V, that is

$$\operatorname{len}(V) := \min\{\#\Phi \ : \ V = V(\Phi)\}.$$

A SIS of length one is called a *principal shift-invariant space* (PSIS). A SIS of finite length is a *finitely generated shift-invariant space* (FSIS).

Remark 1.4.2. If $\varphi \in L^2(\mathbb{R}^d)$ and $\varphi \neq 0$ then the functions $\{t_k\varphi \ : \ k \in \mathbb{Z}^d\}$ are linearly independent (see [Chr03, Proposition 7.4.2] or [HSWW10a] for more details). So, every non trivial SIS is an infinite dimensional linear space.

As a consequence of the integer invariance of the SISs we have the following lemma.

Lemma 1.4.3. *Let $V \subseteq L^2(\mathbb{R}^d)$ be a SIS and P_V the orthogonal projection onto V. Then*

$$t_k P_V = P_V t_k \quad \forall\, k \in \mathbb{Z}^d.$$

Let us remark here that if $\Phi \subseteq L^2(\mathbb{R}^d)$ is a set of generators for a shift-invariant space V, that is $V = V(\Phi)$, then the set $E(\Phi)$ does not need to be a frame for V, even for finitely generated SISs (see Example 1.5.15). However it is always true that there exists a set of generators for V such that its integer translates form a frame for V. This is the result of the next theorem.

Theorem 1.4.4. *Given V a SIS of $L^2(\mathbb{R}^d)$, there exists a subset $\Phi = \{\varphi_j\}_{j\in J} \subseteq V$ such that $E(\Phi)$ is a Parseval frame for V. If V is finitely generated, the cardinal of J can be chosen to be the length of V.*

We would like to note here that although a SIS always has a frame of translates, there are SISs which have no Riesz bases of translates (see Example 1.5.11). This fact shows the importance of considering frames instead of Riesz bases when we are studying the structure of SISs.

1.4.1 Sampling in shift-invariant spaces

Our aim in this section is to give a brief description of sampling in shift-invariant spaces, for more details we refer the reader to [AG01, Sun05, Wal92, ZS99].

We will begin by studying the structure of the canonical dual of a frame of translates. We will show that the canonical dual is formed by translates of functions.

Proposition 1.4.5. *Let $\Phi = \{\varphi_j\}_{j\in J}$ be a set of functions of $L^2(\mathbb{R}^d)$. Assume $E(\Phi)$ is a frame for a closed space $V \subseteq L^2(\mathbb{R}^d)$. Then, the dual frame of $E(\Phi)$ is the set of translates $E(\widetilde{\Phi}) = \{t_k\widetilde{\varphi_j}\}_{j\in J, k\in\mathbb{Z}^d}$, where $\widetilde{\varphi_j} = S^{-1}\varphi_j$ and S is the frame operator associated to $E(\Phi)$ given by*

$$S : V \to V \quad Sf = \sum_{k\in\mathbb{Z}^d}\sum_{j\in J}\langle f, t_k\varphi_j\rangle t_k\varphi_j.$$

Proof. Recall from Proposition 1.2.13 that the canonical dual of $E(\Phi)$ is given by $\{S^{-1}(t_k\varphi_j) : k \in \mathbb{Z}^d, j \in J\}$. It is easily seen that the operator S commutes with integer translates. So, its inverse also commutes with integer translates. Thus, the canonical dual is given by $\{t_k(S^{-1}\varphi_j) : k \in \mathbb{Z}^d, j \in J\}$.

□

As we have mentioned in the Introduction, the Kotelnikov-Shannon-Whittaker (KSW) theorem states that a band-limited function f can be reconstructed from its values in the integers using the formula

$$f(t) = \sum_{k\in\mathbb{Z}} f(k)\mathrm{sinc}(t-k),$$

with the series on the right converging uniformly on $\mathbb{R}$, as well as in $L^2(\mathbb{R})$ (see (0.1)).

The space of band-limited functions $PW = \{f \in L^2(\mathbb{R}) : \mathrm{supp}(\widehat{f}) \subseteq [-\frac{1}{2}, \frac{1}{2}]\}$ is a principal shift-invariant space generated by the function $\varphi = \mathrm{sinc}$. That is, $PW = V(\mathrm{sinc})$.

As a generalization of the KSW theorem, the sampling problem in SISs consists in studying conditions on the generators of a SIS V in order that every function of V can be reconstructed from its values in a discrete sequence of samples.

In this section we will focus our attention in the problem of sampling in principal shift-invariant spaces. We will describe some of the conditions which a generator φ for a PSIS must satisfy in order to have a reconstruction formula in $V(\varphi)$ similar to the one given in the KSW theorem.

A closed subspace $V \subseteq L^2(\mathbb{R}^d)$ of continuous functions will be called a *reproducing kernel Hilbert space* (RKHS) if for each $x \in \mathbb{R}^d$ the evaluation function

$$f \mapsto f(x)$$

is a continuous linear functional on V. If this condition is verified, by the Riesz's representation theorem (see for instance [Con90]), for every $x \in \mathbb{R}^d$ there exists a unique function $N_x \in V$ such that

$$f(x) = \langle f, N_x\rangle \quad \forall f \in V.$$

The set of functions $\{N_x\}_{x\in\mathbb{R}^d}$ is called the *reproducing kernel*.

Assume now that for a given $\varphi \in L^2(\mathbb{R}^d)$, the set $E(\varphi)$ is a frame for $V = V(\varphi)$ and that V is a RKHS. Then, for every $f \in V$ and $k \in \mathbb{Z}^d$,

$$\langle f, t_k N_0\rangle = \langle t_{-k} f, N_0\rangle = t_{-k} f(0) = f(k) = \langle f, N_k\rangle.$$

That is, $t_k N_0 = N_k$. If, in addition, $E(N_0)$ is a frame for V with dual frame $E(\widetilde{N_0})$ (see Proposition 1.4.5), by Proposition 1.2.13 we have that

$$f(x) = \sum_{k\in\mathbb{Z}^d} \langle f, t_k N_0\rangle t_k \widetilde{N_0} = \sum_{k\in\mathbb{Z}^d} f(k) t_k \widetilde{N_0}. \tag{1.5}$$

If the convergence of the previous series is uniform, then every function $f \in V$ can be reconstructed from its values in the integers. In this way, we obtain in $V(\varphi)$ a result similar to the one in the KSW theorem.

As a consequence of the previous analysis, we obtain that the sampling problem for principal shift-invariant spaces is based on studying conditions on the generator φ so that every function of $V = V(\varphi)$ is continuous, the space V is a RKHS, the set $E(N_0)$ is a frame for V, and the convergence in (1.5) is uniform. All of these conditions were studied in [ZS99, Sun05], for a fuller treatment of this problem we refer the reader to these papers.

1.5 Range function and fibers for shift-invariant spaces

A useful tool in the theory of shift-invariant spaces is based on early work of Helson [Hel64]. An $L^2(\mathbb{R}^d)$ function is decomposed into "fibers". This produces a characterization of SISs in terms of closed subspaces of $\ell^2(\mathbb{Z}^d)$ (the fiber spaces). The advantage of this approach is that, although the FSISs are infinite-dimensional subspaces (see Remark 1.4.2), most of their properties can be translated into properties on the fibers of the spanning sets. That allows to work with finite-dimensional subspaces of $\ell^2(\mathbb{Z}^d)$.

In the sequel, we will give the definition and some properties of the fibers. For a detailed description of this approach, see [Bow00] and the references therein.

The Hilbert space of square integrable vector functions $L^2([0,1)^d, \ell^2(\mathbb{Z}^d))$, consists of all vector valued measurable functions $F : [0,1)^d \to \ell^2(\mathbb{Z}^d)$ such that

$$\|F\| := \left(\int_{[0,1)^d} \|F(x)\|^2_{\ell^2}\, dx\right)^{\frac{1}{2}},$$

is finite.

Proposition 1.5.1. *The function $\tau : L^2(\mathbb{R}^d) \to L^2([0,1)^d, \ell^2(\mathbb{Z}^d))$ defined for $f \in L^2(\mathbb{R}^d)$ by*

$$\tau f(\omega) := \{\widehat{f}(\omega + k)\}_{k\in\mathbb{Z}^d},$$

is an isometric isomorphism between $L^2(\mathbb{R}^d)$ and $L^2([0,1)^d, \ell^2(\mathbb{Z}^d))$.

The sequence $\{\widehat{f}(\omega + k)\}_{k\in\mathbb{Z}^d}$ is called the fiber of f at ω.

Definition 1.5.2. A *range function* is a mapping

$$J : [0,1)^d \to \{\text{closed subspaces of } \ell^2(\mathbb{Z}^d)\}.$$

J is measurable if the operator valued function of the orthogonal projections $\omega \mapsto P_{J(\omega)}$ is weakly measurable. In a separable Hilbert space measurability is equivalent to weak measurability. Therefore, the measurability of J is equivalent to $\omega \mapsto P_{J(\omega)}(a)$ being vector measurable for each $a \in \ell^2(\mathbb{Z}^d)$, or $\omega \mapsto P_{J(\omega)}(F(\omega))$ being vector measurable for each fixed vector measurable function $F : [0,1)^d \to \ell^2(\mathbb{Z}^d)$.

Shift-invariant spaces can be characterized through range functions.

Proposition 1.5.3. *A closed subspace $V \subseteq L^2(\mathbb{R}^d)$ is shift-invariant if and only if*

$$V = \{f \in L^2(\mathbb{R}^d) \,:\, \tau f(\omega) \in J_V(\omega) \text{ for a.e. } \omega \in [0,1)^d\},$$

where J_V is a measurable range function. The correspondence between V and J_V is one-to-one.

Moreover, if $V = V(\Phi)$ for some countable set $\Phi \subseteq L^2(\mathbb{R}^d)$, then

$$J_V(\omega) = \overline{\operatorname{span}}\{\tau\varphi(\omega) \,:\, \varphi \in \Phi\} \quad \text{for a.e. } \omega \in [0,1)^d.$$

The subspace $J_V(\omega)$ is called the fiber space of V at ω.

Note that if $V \subseteq L^2(\mathbb{R}^d)$ is an FSIS generated by the set of functions $\Phi = \{\varphi_1, \dots, \varphi_m\}$, then

$$J_V(\omega) = \operatorname{span}\{\tau\varphi_1(\omega), \dots, \tau\varphi_m(\omega)\}.$$

So, even though V is an infinite dimensional subspace of $L^2(\mathbb{R}^d)$, the fiber spaces $J_V(\omega)$ are all finite dimensional subspaces of $\ell^2(\mathbb{Z}^d)$.

We have the following property concerning fibers of SISs.

Proposition 1.5.4. *Let V be a SIS of $L^2(\mathbb{R}^d)$ and $f \in L^2(\mathbb{R}^d)$, then*

$$\tau(P_V f)(\omega) = P_{J_V(\omega)}(\tau f(\omega)) \quad \text{for a.e. } \omega \in [0,1)^d.$$

As a consequence of the previous proposition, we obtain the following.

Proposition 1.5.5. *Let V_1 and V_2 be SISs. If $V = V_1 \dot{\oplus} V_2$, then*

$$J_V(\omega) \;=\; J_{V_1}(\omega) \dot{\oplus} J_{V_2}(\omega), \quad \text{a.e. } \omega \in [0,1)^d.$$

The converse of this proposition is also true, but will not be needed for the subjects developed in this thesis.

Let us now introduce the concept of dimension function for SISs.

Definition 1.5.6. Given V a SIS of $L^2(\mathbb{R}^d)$, the *dimension function* associated to V is defined by

$$\dim_V : [0,1)^d \to \mathbb{N}_0 \cup \{\infty\}, \quad \dim_V(\omega) = \dim(J_V(\omega)).$$

Here $\mathbb{N}_0$ denotes the set of non-negative integers.

We have the following property which relates the essential supremum of the dimension function to the length of an FSIS.

Proposition 1.5.7 ([dBDVR94])**.** *Let $V \subseteq L^2(\mathbb{R}^d)$ be an FSIS. Then*

$$\operatorname{len}(V) = \operatorname{ess-sup}\{\dim_V(\omega) : \omega \in [0,1)^d\}.$$

1.5.1 Riesz bases and frames for shift-invariant spaces

The next two theorems characterize Bessel sequences, frames and Riesz bases of translates in terms of fibers. The main idea is that every property of the set $E(\Phi)$ (being a Bessel sequence, a frame or a Riesz basis) is equivalent to its fibers satisfying an analogous property in a uniform way.

Theorem 1.5.8. *Let Φ be a countable subset of $L^2(\mathbb{R}^d)$. The following are equivalent.*

i) *$E(\Phi)$ is a Bessel sequence in $L^2(\mathbb{R}^d)$ with constant β.*

ii) *$\tau\Phi(\omega) := \{\tau\varphi(\omega) : \varphi \in \Phi\}$ is a Bessel sequence in $\ell^2(\mathbb{Z}^d)$ with constant β for a.e. $\omega \in [0,1)^d$.*

Theorem 1.5.9. *Let $V = V(\Phi)$, where Φ is a countable subset of $L^2(\mathbb{R}^d)$. Then the following holds:*

i) *$E(\Phi)$ is a frame for V with constants α and β if and only if $\tau\Phi(\omega)$ is a frame for $J_V(\omega)$ with constants α and β for a.e. $\omega \in [0,1)^d$.*

ii) *$E(\Phi)$ is a Riesz basis for V with constants α and β if and only if $\tau\Phi(\omega)$ is a Riesz basis for $J_V(\omega)$ with constants α and β for a.e. $\omega \in [0,1)^d$.*

Furthermore, if Φ is finite, V has a Riesz basis of translates if and only if the dimension function associated to V is constant a.e. $\omega \in [0,1)^d$.

Remark 1.5.10. If V is an FSIS generated by $\Phi = \{\varphi_1, \dots, \varphi_m\} \subseteq L^2(\mathbb{R}^d)$, then $J_V(\omega) = \operatorname{span}\{\tau\varphi_1(\omega), \dots, \tau\varphi_m(\omega)\}$ a.e. $\omega \in [0,1)^d$. So, by Remark 1.2.12, $\tau\Phi(\omega)$ is a frame for $J_V(\omega)$ for a.e. ω. But, as we will see in Example 1.5.15, $E(\Phi)$ might not be a frame for $V(\Phi)$ in general. This is due to the fact that we need a pair of uniform positive frame bounds α and β for the frame $\tau\Phi(\omega)$ which are independent of ω in order for $E(\Phi)$ to be a frame for $V(\Phi)$.

As we have mentioned in Theorem 1.4.4 every SIS has a frame of translates. Using fiberization techniques, we will give below an example of a SIS which do not have a Riesz basis of translates.

Example 1.5.11. Consider the shift-invariant space V generated by $\varphi \in L^2(\mathbb{R})$, where $\widehat{\varphi}(\omega) = \chi_{[0,\frac{1}{2})}(\omega)$. Since $\dim_V(\omega) = 1$ for a.e. $\omega \in [0, \frac{1}{2})$ and $\dim_V(\omega) = 0$ for a.e. $\omega \in [\frac{1}{2}, 1)$, it follows by Theorem 1.5.9 that V has no Riesz bases of translates.

1.5.2 The Gramian operator for shift-invariant spaces

Definition 1.5.12. Let $\Phi = \{\varphi_j\}_{j\in J}$ be a countable set of functions in $L^2(\mathbb{R}^d)$ such that $E(\Phi)$ is a Bessel sequence. The Gramian of Φ at $\omega \in [0, 1)^d$ is $\mathcal{G}_\Phi(\omega) : \ell^2(J) \to \ell^2(J)$,

$$(\mathcal{G}_\Phi(\omega))_{i,j} = \langle \tau\varphi_j(\omega), \tau\varphi_i(\omega)\rangle_{\ell^2(\mathbb{Z}^d)} = \sum_{k\in\mathbb{Z}^d} \widehat{\varphi}_i(\omega + k)\overline{\widehat{\varphi}_j(\omega + k)} \quad \forall\, i, j \in J. \tag{1.6}$$

In the notation of Definition 1.3.1, $\mathcal{G}_\Phi(\omega)$ is the Gramian operator associated to the Bessel sequence $\tau\Phi(\omega) = \{\tau\varphi_j(\omega)\}_{j\in J}$ in $\ell^2(\mathbb{Z}^d)$, that is $\mathcal{G}_\Phi(\omega) = G_{\tau\Phi(\omega)}$.

When $\Phi = \{\varphi\}$, the Gramian will be denoted by $\mathcal{G}_\varphi$ and its expression is

$$\mathcal{G}_\varphi(\omega) = \langle \tau\varphi(\omega), \tau\varphi(\omega)\rangle_{\ell^2(\mathbb{Z}^d)} = \sum_{k\in\mathbb{Z}^d} |\widehat{\varphi}(\omega + k)|^2.$$

From Theorem 1.5.8 and Theorem 1.5.9 we obtain the following result (see [Bow00] for more details).

Theorem 1.5.13. *Let $\Phi = \{\varphi_j\}_{j\in J} \subseteq L^2(\mathbb{R}^d)$. Then,*

i) $E(\Phi)$ is a Bessel sequence with constant β if and only if

$$\text{ess-sup}_{\omega\in[0,1)^d} \|\mathcal{G}_\Phi(\omega)\|_{\text{op}} \le \beta.$$

ii) $E(\Phi)$ is a frame for $V(\Phi)$ with constants α and β if and only if for almost all $\omega \in [0, 1)^d$,

$$\alpha\langle \mathcal{G}_\Phi(\omega)c, c\rangle \le \langle \mathcal{G}_\Phi^2(\omega)c, c\rangle \le \beta\langle \mathcal{G}_\Phi(\omega)c, c\rangle \quad \forall\, c \in \ell^2(J).$$

iii) $E(\Phi)$ is a Riesz basis for $V(\Phi)$ with constants α and β if and only if for almost all $\omega \in [0, 1)^d$,

$$\alpha\|c\|^2 \le \langle \mathcal{G}_\Phi(\omega)c, c\rangle \le \beta\|c\|^2 \quad \forall\, c \in \ell^2(J).$$

Remark 1.5.14. As a consequence of Theorem 1.5.13, for a PSIS $V(\varphi)$ we have

i) $E(\varphi)$ is a frame for $V(\varphi)$ with constants α and β if and only if

$$\alpha \le \mathcal{G}_\varphi(\omega) \le \beta \text{ for almost all } \omega \in N_\varphi,$$

where $N_\varphi := \{\omega \in [0, 1)^d \;:\; \mathcal{G}_\varphi(\omega) \neq 0\}$.

ii) $E(\varphi)$ is a Riesz basis for $V(\varphi)$ with constants α and β if and only if

$$\alpha \le \mathcal{G}_\varphi(\omega) \le \beta \text{ for almost all } \omega \in [0, 1)^d.$$

We refer the reader to [HSWW10a] to see how other properties of $E(\varphi)$ (such as being a Schauder basis for $V(\varphi)$) correspond to those of $\mathcal{G}_\varphi$.

We will present now an example from [Chr03] which shows a function φ whose translates are not a frame for the SIS generated by φ.

Example 1.5.15. Let $\varphi = \chi_{[-1,2)}$. It can be shown (see [Chr03]) that $\mathcal{G}_\varphi(\omega) = 3 + 4\cos(2\pi\omega) + 2\cos(4\pi\omega)$. Note that $\mathcal{G}_\varphi$ is continuous and has two isolated zeros $\mathcal{G}_\varphi(\frac{1}{3}) = \mathcal{G}_\varphi(\frac{2}{3}) = 0$. So, by Remark 1.5.14, $E(\varphi)$ is not a frame for $V(\varphi)$.

2

Optimal signal models and dimensionality reduction for data clustering

2.1 Introduction

In this chapter we are going to study the problem of finding models which best explain a finite data set of signals. We will first review some results about finding a subspace that is closest to a given finite data set. More precisely, if $\mathcal{F} = \{f_1, \ldots, f_m\}$ is a set of vectors of a Hilbert space $\mathcal{H}$, we will study the problem of finding an optimal subspace $V_0 \subseteq \mathcal{H}$ that minimizes the expression

$$\mathcal{E}(\mathcal{F}, V) := \sum_{i=1}^{m} d^2(f_i, V) = \sum_{i=1}^{m} \|f_i - P_V f_i\|^2$$

over all possible choices of subspaces V belonging to an appropriate class C of subspaces of $\mathcal{H}$.

We will focus our attention in finding optimal subspaces for two cases: when $\mathcal{H} = \mathbb{R}^N$ and C is the set of subspaces of dimension at most k with $k << N$, and when $\mathcal{H} = L^2(\mathbb{R}^d)$ with C being the family of FSISs of length at most k.

Following the new paradigm for signal sampling and reconstruction developed recently by Lu y Do [LD08] which assumes that the signals live in a union of subspaces instead of a single vector space, we will study the problem of finding an appropriate signal model $\mathcal{X} = \cup_i V_i$ from a set of observed data $\mathcal{F} = \{f_1, \ldots, f_m\}$.

We will review the results from [ACM08], which find subspaces $V_1, \ldots, V_l$, of some Hilbert space $\mathcal{H}$ that minimize the expression

$$e(\mathcal{F}, \{V_1, \ldots, V_l\}) = \sum_{i=1}^{m} \min_{1\leq j\leq l} d^2(f_i, V_j),$$

over all possible choices of l subspaces belonging to an appropriate class of subspaces of $\mathcal{H}$.

If the subspaces V_i belong to a family of closed subspaces C which satisfies the so called Minimum Subspace Approximation Property (MSAP), an optimal solution to the non-linear subspace modeling problem that best fit the data exists, and algorithms to find these subspaces were developed in [ACM08].

The results from [ACM08] are proved in a general setting and then applied to the case of low dimensional subspaces of $\mathbb{R}^N$ and to infinite dimensional shift-invariant spaces of $L^2(\mathbb{R}^d)$.

For the euclidean case $\mathbb{R}^N$, the problem of finding a union of subspaces of dimension $k << N$ that best explains a data set $\mathcal{F} = \{f_1, \ldots, f_m\} \subseteq \mathbb{R}^N$ increases dramatically with the dimension N. In the present chapter we have focused on the computational complexity of finding optimal union of subspaces in $\mathbb{R}^N$. More precisely, we study techniques of dimension reduction for the algorithm proposed in [ACM08]. These techniques can also be used in a wide variety of situations and are not limited to this particular application.

We use random linear transformations to map the data to a lower dimensional space. The "projected" signals are then processed in that space, (i.e. finding the optimal union of subspaces) in order to produce an optimal partition. Then we apply this partition to the original data to obtain the associated model for that partition and obtain a bound for the error.

We analyze two situations. First we study the case when the data belongs to a union of subspaces (ideal case with no noise). In that case we obtain the optimal model using almost any transformation (see Proposition 2.4.3).

In the presence of noise, the data usually doesn't belong to a union of low dimensional subspaces. Thus, the distances from the data to an optimal model add up to a positive error. In this case, we need to restrict the admissible transformations. We apply recent results on distributions of matrices satisfying concentration inequalities, which also proved to be very useful in the theory of compressed sensing.

We are able to prove that the model obtained by our approach is quasi optimal with a high probability. That is, if we map the data using a random matrix from one of the distributions satisfying the concentration law, then with high probability, the distance of the data to the model is bounded by the optimal distance plus a constant. This constant depends on the parameter of the concentration law, and the parameters of the model (number and dimension of the subspaces allowed in the model).

Let us remark here that the problem of finding the optimal union of subspaces that fit a given data set is also known as "Projective clustering". Several algorithms have been proposed in the literature to solve this problem. Particularly relevant is [DRVW06] (see also references therein) where the authors used results from volume and adaptive sampling to obtain a polynomial-time approximation scheme. See [AM04] for a related algorithm.

The rest of the chapter is organized as follows: in Section 2.2 we present the Eckart-Young's Theorem, which solves the problem of finding a subspace of dimension less than or equal to k that best approximates a finite set of vectors of $\mathbb{R}^N$. We also review the results from [ACHM07] to find an FSIS which best fits a finite data set of functions of $L^2(\mathbb{R}^d)$.

In Section 2.3 we state the results from [ACM08] which find, for a given set of vectors in a Hilbert space, a union of subspaces minimizing the sum of the square of the distances between each vector and its closest subspace in the collection. We also review the iterative algorithm proposed in [ACM08] for finding the solution subspaces.

In Section 2.4 we concentrate on the non-linear subspace modeling problem when the model is a finite union of subspaces of $\mathbb{R}^N$ of dimension $k << N$. We study a class of transformations that map the problem into another one in lower dimension. We use the best model in the low dimensional space to approximate the best solution in the original high dimensional space. We then estimate the error produced between this solution and the optimal solution in the high dimensional space.

In Section 2.5 we give the proofs of the results from Subsection 2.4.2.

2.2 Optimal subspaces as signal models

Given a set of vectors $\mathcal{F} = \{f_1, \dots, f_m\}$ in a separable Hilbert space $\mathcal{H}$ and a family of closed subspaces C of $\mathcal{H}$, the problem of finding a subspace $V \in C$ that best models the data $\mathcal{F}$ has many applications to mathematics and engineering.

Since one of our goals is to model a set of data by a closed subspace, we first provide a measure of how well a given data set can be modeled by a subspace.

Definition 2.2.1. Given a set of vectors $\mathcal{F} = \{f_1, \dots, f_m\}$ in a separable Hilbert space, the distance from a closed subspace $V \subseteq \mathcal{H}$ to $\mathcal{F}$ will be denoted by

$$\mathcal{E}(\mathcal{F}, V) := \sum_{i=1}^{m} d^2(f_i, V) = \sum_{i=1}^{m} \|f_i - P_V f_i\|^2.$$

We will say that a family of subspaces C has the Minimum Subspace Approximation Property (MSAP) if for any finite set $\mathcal{F}$ of vectors in $\mathcal{H}$ there exists a subspace $V_0 \in C$ such that

$$\mathcal{E}(\mathcal{F}, V_0) = \inf\{\mathcal{E}(\mathcal{F}, V) : V \in C\} \leq \mathcal{E}(\mathcal{F}, V), \quad \forall\, V \in C. \tag{2.1}$$

Any subspace $V_0 \in C$ satisfying (2.1) will be called an *optimal subspace* for $\mathcal{F}$.

Necessary and sufficient conditions for C to satisfy the MSAP are obtained in [AT10].

Let us denote by $\mathcal{E}_0(\mathcal{F}, C)$ the minimal error defined by

$$\mathcal{E}_0(\mathcal{F}, C) := \inf\{\mathcal{E}(\mathcal{F}, V) : V \in C\}. \tag{2.2}$$

In this section we will study the problem of finding optimal subspaces for two cases: when $\mathcal{H} = \mathbb{R}^N$ and C is the set of subspaces of dimension at most k (with $k << N$), and when $\mathcal{H} = L^2(\mathbb{R}^d)$ with C being the family of FSISs of length at most k.

We will begin by studying the euclidean case $\mathcal{H} = \mathbb{R}^N$. Assume we have a finite data set $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq \mathbb{R}^N$. Our goal is to find a subspace V_0 such that $\dim(V_0) \leq k$ and

$$\mathcal{E}(\mathcal{F}, V_0) = \mathcal{E}_0(\mathcal{F}, C_k) = \inf\{\mathcal{E}(\mathcal{F}, V) : V \in C_k\},$$

where $\mathcal{C}_k$ is the family of subspaces of $\mathbb{R}^N$ with dimension at most k.

This well-known problem is solved by the Eckart-Young's Theorem (see [Sch07]) which uses the Singular Value Decomposition (SVD) of a matrix. Before stating the theorem, we will briefly recall the SVD of a matrix (for a detailed treatment see for example [Bha97]).

Let $M = [f_1, \dots, f_m] \in \mathbb{R}^{N\times m}$ and $d := \operatorname{rank}(M)$. Consider the matrix $M^*M \in \mathbb{R}^{m\times m}$. Since M^*M is self-adjoint and positive semi-definite, it has eigenvalues $\lambda_1 \geq \cdots \geq \lambda_d > 0 = \lambda_{d+1} = \cdots = \lambda_m$. The associated eigenvectors $y_1, \dots, y_m$ can be chosen to form an orthonormal basis of $\mathbb{R}^m$. The left singular vectors $u_1, \dots, u_d$ can then be obtained from

$$u_i = \lambda_i^{-1/2} M y_i = \lambda_i^{-1/2} \sum_{j=1}^{m} y_{ij} f_j \quad \forall\, 1 \leq i \leq d.$$

The remaining left singular vectors $u_{d+1}, \dots, u_m$ can be chosen to be any orthonormal collection of $m - d$ vectors in $\mathbb{R}^N$ that are perpendicular to the subspace spanned by the columns of M. One obtain the following SVD of M

$$M = U\Lambda^{1/2}Y^*,$$

where $U \in \mathbb{R}^{N\times m}$ is the matrix with columns $\{u_1, \dots, u_m\}$, $\Lambda^{1/2} = \operatorname{diag}(\lambda_1^{1/2}, \dots, \lambda_m^{1/2})$, and $Y = \{y_1, \dots, y_m\} \in \mathbb{R}^{m\times m}$ with $U^*U = I_m = Y^*Y = YY^*$.

We are now able to state the Eckart-Young's Theorem.

Theorem 2.2.2. *Let $\mathcal{F} = \{f_1, \dots, f_m\}$ be a set of vectors in $\mathbb{R}^N$ and let $M = [f_1, \dots, f_m] \in \mathbb{R}^{N\times m}$ be the matrix with columns f_i. Suppose that M has a SVD $M = U\Lambda^{1/2}Y^*$ and that $0 < k \leq d$, with $d := \operatorname{rank}(M)$. If $V_0 = \operatorname{span}\{u_1, \dots, u_k\}$, then*

$$\mathcal{E}(\mathcal{F}, V_0) = \mathcal{E}_0(\mathcal{F}, \mathcal{C}_k) = \inf\{\mathcal{E}(\mathcal{F}, V) \,:\, V \in \mathcal{C}_k\}.$$

Furthermore,

$$\mathcal{E}_0(\mathcal{F}, \mathcal{C}_k) = \sum_{j=k+1}^{d} \lambda_j,$$

where $\lambda_1 \geq \cdots \geq \lambda_d > 0$ are the positive eigenvalues of M^*M.

The previous theorem proves that in $\mathcal{H} = \mathbb{R}^N$, the class $\mathcal{C}_k$ of subspaces of dimension at most k has the MSAP. Therefore, for any finite set $\mathcal{F} = \{f_1, \dots, f_m\}$ of vectors in $\mathbb{R}^N$ there exists an optimal subspace $V_0 \in \mathcal{C}_k$ which best approximates the data set $\mathcal{F}$. Moreover, Theorem 2.2.2 gives a way to construct the generators of an optimal subspace and estimates the minimal error $\mathcal{E}_0(\mathcal{F}, \mathcal{C}_k)$.

Let us now study the problem of finding an FSIS of $L^2(\mathbb{R}^d)$ that best approximates a finite data set of functions of $L^2(\mathbb{R}^d)$. More specifically, given a set of functions $\mathcal{F} = \{f_1, \dots, f_m\}$ in $L^2(\mathbb{R}^d)$, our goal is to find an FSIS V_0 of length at most k (with k much smaller than m) that is closest to $\mathcal{F}$ in the sense that

$$\mathcal{E}(\mathcal{F}, V_0) = \mathcal{E}_0(\mathcal{F}, \mathcal{L}_k) = \inf\{\mathcal{E}(\mathcal{F}, V) \,:\, V \in \mathcal{L}_k\}, \tag{2.3}$$

where $\mathcal{L}_k$ is the set of all the SISs of length less than or equal to k.

To solve this problem, in [ACHM07] the authors used fiberization techniques to reduce it to the finite dimensional problem of finding a subspace of dimension at most k that best approximates a finite data set of vectors in the Hilbert space $\ell^2(\mathbb{Z}^d)$. This last problem can be solved by an extension of the Eckart-Young's Theorem (for more details see [ACHM07]).

The following theorem states the existence of an optimal subspace which solves problem (2.3). Recall from Definition 1.6 that for a given set $\mathcal{F} = \{f_1, \dots, f_m\}$ of functions in $L^2(\mathbb{R}^d)$, the Gramian matrix $\mathcal{G}_{\mathcal{F}}(\omega) \in \mathbb{C}^{m\times m}$ is defined by $(\mathcal{G}_{\mathcal{F}}(\omega))_{i,j} = \langle \tau f_i(\omega), \tau f_j(\omega)\rangle$ for every $1 \le i, j \le m$, where $\tau f(\omega) = \{\widehat{f}(\omega + k)\}_{k\in\mathbb{Z}^d}$.

Theorem 2.2.3. *Assume $\mathcal{F} = \{f_1, \dots, f_m\}$ is a set of functions in $L^2(\mathbb{R}^d)$, let $\lambda_1(\omega) \ge \lambda_2(\omega) \ge \dots \ge \lambda_m(\omega)$ be the eigenvalues of the Gramian $\mathcal{G}_{\mathcal{F}}(\omega)$. Then*

i) The eigenvalues $\lambda_i(\omega)$, $1 \le i \le m$ are $\mathbb{Z}^d$-periodic, measurable functions in $L^2([0,1)^d)$ and

$$\mathcal{E}_0(\mathcal{F}, \mathcal{L}_k) = \sum_{i=k+1}^{m} \int_{[0,1)^d} \lambda_i(\omega) d\omega,$$

where $\mathcal{L}_k$ is the set of all the FSISs of length less than or equal to k.

ii) Let $N_i := \{\omega : \lambda_i(\omega) \neq 0\}$, and define $\tilde{\sigma}_i(\omega) = \lambda_i^{-1/2}(\omega)$ on N_i and $\tilde{\sigma}_i(\omega) = 0$ on N_i^c. Then, there exists a choice of measurable left eigenvectors $y_1(\omega), \dots, y_k(\omega)$ associated with the first k largest eigenvalues of $\mathcal{G}_{\mathcal{F}}(\omega)$ such that the functions defined by

$$\widehat{\varphi_i}(\omega) = \tilde{\sigma}_i(\omega) \sum_{j=1}^{m} y_{ij}(\omega) \widehat{f_j}(\omega), \quad i = 1, \dots, k, \ \omega \in \mathbb{R}^d$$

are in $L^2(\mathbb{R}^d)$. Furthermore, the corresponding set of functions $\Phi = \{\varphi_1, \dots, \varphi_k\}$ is a generator for an optimal space V_0 and the set $E(\Phi)$ is a Parseval frame for V_0.

As a consequence of the previous theorem we obtain that the class $\mathcal{L}_k$ of FSISs of $L^2(\mathbb{R}^d)$ of length at most k satisfies the MSAP. So, problem (2.3) always has a solution. Moreover, Theorem 2.2.3 gives a way to construct the generators of an optimal subspace and estimates the minimal error $\mathcal{E}_0(\mathcal{F}, \mathcal{L}_k)$.

2.3 Optimal union of subspaces as signal models

In this section we will study the problem of finding a union of subspaces that best approximates a finite data set in a Hilbert space $\mathcal{H}$.

Let C be a family of closed subspaces of $\mathcal{H}$ containing the zero subspace. Given $l \in \mathbb{N}$, denote by $\mathcal{B}$ the collection of *bundles* of subspaces in C,

$$\mathcal{B} = \{B = \{V_1, \dots, V_l\} : V_i \in C, \ i = 1, ..., l\}.$$

For a set of vectors $\mathcal{F} = \{f_1, \ldots, f_m\}$ in $\mathcal{H}$, the error between a bundle $B = \{V_1, \ldots, V_l\} \in \mathcal{B}$ and $\mathcal{F}$ will be defined by

$$e(\mathcal{F}, B) = \sum_{i=1}^{m} \min_{1 \le j \le l} d^2(f_i, V_j),$$

where d stands for the distance in $\mathcal{H}$ (see Figure 2.1 for an example).

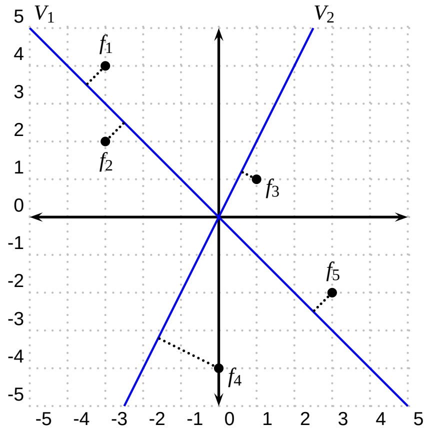

Figure 2.1: An example of a data set $\mathcal{F} = \{f_1, \ldots, f_5\}$ in $\mathbb{R}^2$ and a bundle $B = \{V_1, V_2\}$ of two lines. In this case $e(\mathcal{F}, B) = d^2(f_1, V_1) + d^2(f_2, V_1) + d^2(f_3, V_2) + d^2(f_4, V_2) + d^2(f_5, V_1)$. The partition generated by the bundle B is $S_1 = \{1, 2, 5\}$ and $S_2 = \{3, 4\}$.

Observe that for the case $l = 1$ the error e coincides with the error $\mathcal{E}$ defined in the previous section. That is,

$$e(\mathcal{F}, \{V\}) = \mathcal{E}(\mathcal{F}, V) = \sum_{i=1}^{m} d^2(f_i, V).$$

Recall from the previous section that a family of subspaces C has the Minimum Subspace Approximation Property (MSAP) if for any finite set $\mathcal{F}$ of vectors in $\mathcal{H}$ there exists a subspace $V_0 \in C$ such that

$$\mathcal{E}(\mathcal{F}, V_0) = \inf\{\mathcal{E}(\mathcal{F}, V) \,:\, V \in C\} \le \mathcal{E}(\mathcal{F}, V), \quad \forall\, V \in C.$$

The following theorem states that the problem of finding an optimal union of subspaces has solution for every finite data set $\mathcal{F} \subseteq \mathcal{H}$ and every $l \ge 1$ if and only if C has the MSAP.

Theorem 2.3.1 ([ACM08]). *Let $\mathcal{F} = \{f_1, \ldots, f_m\}$ be vectors in $\mathcal{H}$, and let l be given ($l < m$). If C satisfies the MSAP, then there exists a bundle $B_0 = \{V_1^0, \ldots, V_l^0\} \in \mathcal{B}$ such that*

$$e(\mathcal{F}, B_0) = e_0(\mathcal{F}) := \inf\{e(\mathcal{F}, B) \,:\, B \in \mathcal{B}\}. \tag{2.4}$$

Any bundle $B_0 \in \mathcal{B}$ satisfying (2.4) will be called an optimal bundle *for $\mathcal{F}$.*

Remark 2.3.2. In the context of the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^d)$, Teorem 2.2.3 proves that the family $\mathcal{L}_k$ of shift-invariant spaces with length less than or equal to k has the MSAP. Thus, by Theorem 2.3.1 there exists a solution for the problem of optimal union of FSISs.

In the case of $\mathcal{H} = \mathbb{R}^N$, Theorem 2.2.2 states that the family C_k of subspaces of dimension at most k has the MSAP. So, also in this case there exists a union of k-dimensional subspaces which is closest to a given data set.

2.3.1 Bundles associated to a partition and partitions associated to a bundle

The following relations between partitions of the indices $\{1, \ldots, m\}$ and bundles will be relevant for understanding the solution to the problem of optimal models. From now on we will assume that the class C has the MSAP.

We will denote by $\mathbf{\Pi}_l(\{1, \ldots, m\})$ the set of all l-sequences $\mathbf{S} = \{S_1, \ldots, S_l\}$ of subsets of $\{1, \ldots, m\}$ satisfying the property that for all $1 \leq i, j \leq l$,

$$\bigcup_{r=1}^{l} S_r = \{1, \ldots, m\} \quad \text{and} \quad S_i \cap S_j = \emptyset \text{ for } i \neq j.$$

We want to emphasize that this definition does not exclude the case when some of the S_i are the empty set. By abuse of notation, we will still call the elements of $\mathbf{\Pi}_l(\{1, \ldots, m\})$ *partitions* of $\{1, \ldots, m\}$.

Definition 2.3.3. Given a bundle $B = \{V_1, \ldots, V_l\} \in \mathcal{B}$, we can split the set $\{1, \ldots, m\}$ into a partition $\mathbf{S} = \{S_1, \ldots, S_l\} \in \mathbf{\Pi}_l(\{1, \ldots, m\})$ with respect to that bundle, by grouping together into S_i the indices of the vectors in $\mathcal{F}$ that are closer to a given subspace V_i than to any other subspace V_j, $j \neq i$. Thus, the partitions generated by B are defined by $\mathbf{S} = \{S_1, \ldots, S_l\} \in \mathbf{\Pi}_l(\{1, \ldots, m\})$, where

$$j \in S_i \quad \text{if and only if} \quad d(f_j, V_i) \leq d(f_j, V_h), \quad \forall\, h = 1, \ldots, l.$$

We can also associate to a given partition $\mathbf{S} \in \mathbf{\Pi}_l$ the bundles in $\mathcal{B}$ as follows:

Definition 2.3.4. Given a partition $\mathbf{S} = \{S_1, \ldots, S_l\} \in \mathbf{\Pi}_l$, we will denote by $\mathcal{F}_i$ the set $\mathcal{F}_i = \{f_j\}_{j \in S_i}$. A bundle $B = \{V_1, \ldots, V_l\} \in \mathcal{B}$ is generated by $\mathbf{S}$ if and only if for every $i = 1, \ldots, l$,

$$\mathcal{E}(\mathcal{F}_i, V_i) = \mathcal{E}_0(\mathcal{F}_i, C) = \inf\{\mathcal{E}(\mathcal{F}_i, V) \;:\; V \in C\}.$$

In this way, for a given data set $\mathcal{F}$, every bundle has a set of associated partitions (those that are generated by the bundle) and every partition has a set of associated bundles (those that are generated by the partition). Note however, that the fact that $\mathbf{S}$ is generated by B does not imply that B is generated by $\mathbf{S}$, and vice versa (an example is given in Figure 2.2). However, if B_0 is an optimal bundle that solves the problem for the data $\mathcal{F}$ as in

Theorem 2.3.1, then in this case, the partition $\mathbf{S}_0$ generated by B_0 also generates B_0. On the other hand not every pair $(B, \mathbf{S})$ with this property produces the minimal error $e_0(\mathcal{F})$.

Here and subsequently, the partition $\mathbf{S}_0$ generated by the optimal bundle B_0 will be called an optimal partition for $\mathcal{F}$.

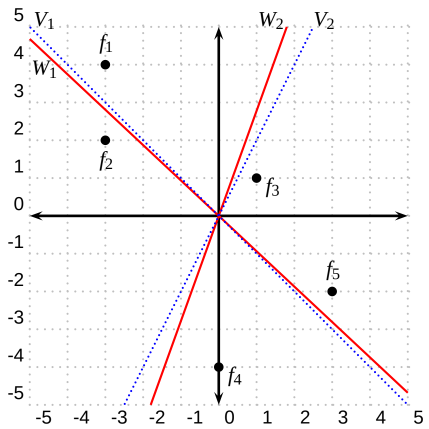

Figure 2.2: The data set $\mathcal{F} = \{f_1, \ldots, f_5\}$ is the same as in Figure 2.1. The bundle $B = \{V_1, V_2\}$ generates the partition $\mathbf{S} = \{\{1, 2, 5\}, \{3, 4\}\}$. This partition generates the bundle $B' = \{W_1, W_2\}$.

An algorithm to solve the problem of finding an optimal union of subspaces was proposed in [ACM08]. It consists in picking any partition $\mathbf{S}_1 \in \mathbf{\Pi}_l$ and finding a bundle B_1 generated by $\mathbf{S}_1$. Then find a partition $\mathbf{S}_2$ generated by the bundle B_1 and calculate the bundle B_2 associated to $\mathbf{S}_2$. Iterate this procedure until obtaining the optimal bundle (see [ACM08] for more details).

2.3.2 The euclidean case: sparsity and dictionaries

In this section we will focus our attention in the problem of optimal union of subspaces for the euclidean case. The study of optimal union of subspaces models for the case $\mathcal{H} = \mathbb{R}^N$ has applications to mathematics and engineering [CL09, EM09, EV09, Kan01, KM02, LD08, AC09, VMS05]. In the previous section we have shown that the problem of finding a union of l subspaces of dimension less than or equal to k that best approximates a data set in $\mathbb{R}^N$ has a solution (see Remark 2.3.2).

In this section, we will relate the existence of optimal union of subspaces in $\mathbb{R}^N$ with the problem of finding a dictionary in which the data set has a certain sparsity. We will also analyze the applicability of the algorithm given in the previous section for the euclidean case.

Definition 2.3.5. Given a set of vectors $\mathcal{F} = \{f_1, \ldots, f_m\}$ in $\mathbb{R}^N$, a real number $\rho \geq 0$ and positive integers $l, k < N$ we will say that the data $\mathcal{F}$ is (l, k, ρ)-sparse if there exist subspaces $V_1, \ldots, V_l$ of $\mathbb{R}^N$ with dimension at most k, such that

$$e(\mathcal{F}, \{V_1, \ldots, V_l\}) = \sum_{i=1}^{m} \min_{1 \le j \le l} d^2(f_i, V_j) \le \rho,$$

where d stands for the euclidean distance in $\mathbb{R}^N$.

When $\mathcal{F}$ is $(l, k, 0)$-sparse, we will simply say that $\mathcal{F}$ is (l, k)-sparse.

Note that if $\mathcal{F}$ is (l, k)-sparse, there exist l subspaces $V_1, \ldots, V_l$ of dimension at most k, such that

$$\mathcal{F} \subseteq \cup_{i=1}^{l} V_i.$$

For the general case $\rho > 0$, the (l, k, ρ)-sparsity of the data implies that $\mathcal{F}$ can be partitioned into a small number of subsets, in such a way that each subset belongs to or is at no more than ρ-distance from a low dimensional subspace. The collection of these subspaces provides an optimal non-linear sparse model for the data.

Observe that if the data $\mathcal{F}$ is (l, k, ρ)-sparse, a model which verifies Definition 2.3.5 provides a dictionary of length not bigger than lk (and in most cases much smaller) in which our data can be represented using at most k atoms with an error smaller than ρ.

More precisely, let $\{V_1, \ldots, V_l\}$ be a collection of subspaces which satisfies Definition 2.3.5 and D a set of vectors from $\bigcup_j V_j$ that is minimal with the property that its span contains $\bigcup_j V_j$. Then for each $f \in \mathcal{F}$ there exists $\Lambda \subseteq D$ with $\#\Lambda \le k$ such that

$$\|f - \sum_{g \in \Lambda} \alpha_g g\|_2^2 \le \rho, \qquad \text{for some scalars } \alpha_g.$$

In [MT82] Megiddo and Tamir showed that it is NP-complete to decide whether a set $\mathcal{F}$ of m points in $\mathbb{R}^2$ can be covered by l lines. This implies that the problem of finding a union a subspaces that best explains a data set is NP-Complete even in the planar case.

The algorithm from [ACM08] described in the previous section involves the calculation of optimal bundles, which depends on finding an optimal subspace for a data set. Recall that the solution to the case $l = 1$ is given by the SVD of a matrix (see Eckart-Young's Theorem). The running time of the SVD method for a matrix $M \in \mathbb{R}^{N \times m}$ is $O(\min\{mN^2, Nm^2\})$ (for further details see [TB97]). Thus the implementation of the algorithm can be very expensive if N is very large.

In the following section we study techniques of dimension reduction to avoid the expensiveness of the algorithm described above. These techniques can also be used in a wide variety of situations and are not limited to this particular application.

2.4 Dimensionality reduction

The problem of finding the optimal union of subspaces that best models a given set of data $\mathcal{F}$ when the dimension of the ambient space N is large is computationally expensive.

When the dimension k of the subspaces is considerably smaller than N, it is natural to map the data onto a lower-dimensional subspace, solve an associated problem in the lower dimensional space and map the solution back into the original space. Specifically, given the data set $\mathcal{F} = \{f_1, \ldots, f_m\} \subseteq \mathbb{R}^N$ which is (l, k, ρ)-sparse and a matrix $A \in \mathbb{R}^{r\times N}$, with $r << N$, find the optimal partition of the projected data $\mathcal{F}' := A(\mathcal{F}) = \{Af_1, \ldots, Af_m\} \subseteq \mathbb{R}^r$, and use this partition to find an approximate solution to the optimal model for $\mathcal{F}$.

2.4.1 The ideal case $\rho = 0$

In this section we will assume that the data $\mathcal{F} = \{f_1, \ldots, f_m\} \subseteq \mathbb{R}^N$ is (l, k)-sparse, i.e., there exist l subspaces of dimension at most k such that $\mathcal{F}$ lies in the union of these subspaces. For this ideal case, we will show that we can always recover the optimal solution to the original problem from the optimal solution to the problem in the low dimensional space as long as the low dimensional space has dimension $r > k$.

We will begin with the proof that for any matrix $A \in \mathbb{R}^{r\times N}$, the projected data $\mathcal{F}' = A(\mathcal{F})$ is (l, k)-sparse in $\mathbb{R}^r$.

Lemma 2.4.1. *Assume the data $\mathcal{F} = \{f_1, \ldots, f_m\} \subseteq \mathbb{R}^N$ is (l, k)-sparse and let $A \in \mathbb{R}^{r\times N}$. Then $\mathcal{F}' := A(\mathcal{F}) = \{Af_1, \ldots, Af_m\} \subseteq \mathbb{R}^r$ is (l, k)-sparse.*

Proof. Let $V_1^0, \ldots, V_l^0$ be optimal spaces for $\mathcal{F}$. Since

$$\dim(A(V_i^0)) \leq \dim(V_i^0) \leq k \quad \forall\, 1 \leq i \leq l,$$

and

$$\mathcal{F}' \subseteq \bigcup_{i=1}^{l} A(V_i^0),$$

it follows that $B := \{A(V_1^0), \ldots, A(V_l^0)\}$ is an optimal bundle for $\mathcal{F}'$ and $e(\mathcal{F}', B) = 0$.

□

Let $\mathcal{F} = \{f_1, \ldots, f_m\} \subseteq \mathbb{R}^N$ be (l, k)-sparse and $A \in \mathbb{R}^{r\times N}$. By Lemma 2.4.1, $\mathcal{F}'$ is (l, k)-sparse. Thus, there exists an optimal partition $\mathbf{S} = \{S_1, \ldots, S_l\}$ for $\mathcal{F}'$ in $\mathbf{\Pi}_l(\{1, \ldots, m\})$, such that

$$\mathcal{F}' \subseteq \bigcup_{i=1}^{l} W_i,$$

where $W_i := \operatorname{span}\{Af_j\}_{j\in S_i}$ and $\dim(W_i) \leq k$. Note that $\{W_1, \ldots, W_l\}$ is an optimal bundle for $\mathcal{F}'$.

We can define the bundle $B_{\mathbf{S}} = \{V_1, \ldots, V_l\}$ by

$$V_i := \operatorname{span}\{f_j\}_{j\in S_i}, \quad \forall\, 1 \leq i \leq l. \tag{2.5}$$

Since $\mathbf{S} \in \mathbf{\Pi}_l(\{1, \ldots, m\})$, we have that

$$\mathcal{F} \subseteq \bigcup_{i=1}^{l} V_i.$$

Thus, the bundle $B_{\mathbf{S}}$ will be optimal for $\mathcal{F}$ if $\dim(V_i) \leq k$, $\forall\, 1 \leq i \leq l$. The above discussion suggests the following definition:

Definition 2.4.2. Let $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq \mathbb{R}^N$ be (l,k)-sparse. We will call a matrix $A \in \mathbb{R}^{r\times N}$ *admissible* for $\mathcal{F}$ if for every optimal partition $\mathbf{S}$ for $\mathcal{F}'$, the bundle $B_{\mathbf{S}}$ defined by (2.5) is optimal for $\mathcal{F}$.

The next proposition states that almost all $A \in \mathbb{R}^{r\times N}$ are admissible for $\mathcal{F}$.

The Lebesgue measure of a set $E \subseteq \mathbb{R}^q$ will be denoted by $|E|$.

Proposition 2.4.3. *Assume the data $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq \mathbb{R}^N$ is (l,k)-sparse and let $r > k$. Then, almost all $A \in \mathbb{R}^{r\times N}$ are admissible for $\mathcal{F}$.*

Proof. If a matrix $A \in \mathbb{R}^{r\times N}$ is not admissible, there exists an optimal partition $\mathbf{S} \in \mathbf{\Pi}_l$ for $\mathcal{F}'$ such that the bundle $B_{\mathbf{S}} = \{V_1, \dots, V_l\}$ is not optimal for $\mathcal{F}$.

Let $\mathcal{D}_k$ be the set of all the subspaces V in $\mathbb{R}^N$ of dimension bigger than k, such that $V = \operatorname{span}\{f_j\}_{j\in S}$ with $S \subseteq \{1, \dots, m\}$.

Thus, we have that the set of all the matrices of $\mathbb{R}^{r\times N}$ which are not admissible for $\mathcal{F}$ is contained in the set

$$\bigcup_{V\in\mathcal{D}_k} \{A \in \mathbb{R}^{r\times N} \;:\; \dim(A(V)) \leq k\}.$$

Note that the set $\mathcal{D}_k$ is finite, since there are finitely many subsets of $\{1, \dots, m\}$. Therefore, the proof of the proposition is complete by showing that for a fixed subspace $V \subseteq \mathbb{R}^N$, such that $\dim(V) > k$, it is true that

$$|\{A \in \mathbb{R}^{r\times N} \;:\; \dim(A(V)) \leq k\}| = 0. \tag{2.6}$$

Let then V be a subspace such that $\dim(V) = t > k$. Given $\{v_1, \dots, v_t\}$ a basis for V, by abuse of notation, we continue to write V for the matrix in $\mathbb{R}^{N\times t}$ with vectors v_i as columns. Thus, proving (2.6) is equivalent to proving that

$$|\{A \in \mathbb{R}^{r\times N} \;:\; \operatorname{rank}(AV) \leq k\}| = 0. \tag{2.7}$$

As $\min\{r,t\} > k$, the set $\{A \in \mathbb{R}^{r\times N} \;:\; \operatorname{rank}(AV) \leq k\}$ is included in

$$\{A \in \mathbb{R}^{r\times N} \;:\; \det(V^*A^*AV) = 0\}. \tag{2.8}$$

Since $\det(V^*A^*AV)$ is a non-trivial polynomial in the $r \times N$ coefficients of A, the set (2.8) has Lebesgue measure zero. Hence, (2.7) follows.

$\square$

2.4.2 The non-ideal case $\rho > 0$

Even if a set of data is drawn from a union of subspaces, in practice it is often corrupted by noise. Thus, in general $\rho > 0$, and our goal is to estimate the error produced when we solve the associated problem in the lower dimensional space and map the solution back into the original space.

Intuitively, if $A \in \mathbb{R}^{r\times N}$ is an arbitrary matrix, the set $\mathcal{F}' = A\mathcal{F}$ will preserve the original sparsity only if the matrix A does not change the geometry of the data in an essential way. One can think that in the *ideal* case, since the data is sparse, it actually lies in an union of low dimensional subspaces (which is a very thin set in the ambient space).

However, when the data is not 0-sparse, but only ρ-sparse with $\rho > 0$, the optimal subspaces plus the data do not lie in a thin set. This is the main obstacle in order to obtain an analogous result as in the ideal case.

Far from having the result that for *almost any* matrix A the geometry of the data will be preserved, we have the Johnson-Lindenstrauss (JL) lemma [JL84], that guaranties - for a given data set - the existence of *one* Lipschitz mapping which approximately preserves the relative distances between the data points.

Several proofs of the JL lemma have been made in the past years. It is of our interest the proof of an improved version of the JL lemma given in [Ach03] that uses random matrices which verify a concentration inequality. In what follows we will announce this concentration inequality. The aim of this chapter is to use these random matrices to obtain positive results for the problem of optimal union of subspaces in the $\rho > 0$ case.

Let $(\Omega, \Pr)$ be a probability measure space. Given $r, N \in \mathbb{N}$, a random matrix $A_\omega \in \mathbb{R}^{r\times N}$ is a matrix with entries $(A_\omega)_{i,j} = a_{i,j}(\omega)$, where $\{a_{i,j}\}$ are independent and identically distributed random variables for every $1 \le i \le r$ and $1 \le j \le N$.

Given $x \in \mathbb{R}^d$, we write $\|x\|$ for the ℓ^2 norm of x in $\mathbb{R}^d$.

Definition 2.4.4. We say that a random matrix $A_\omega \in \mathbb{R}^{r\times N}$ satisfies the concentration inequality if for every $0 < \varepsilon < 1$, there exists $c_0 = c_0(\varepsilon) > 0$ (independent of r, N) such that for any $x \in \mathbb{R}^N$,

$$\Pr\Big((1-\varepsilon)\|x\|^2 \le \|A_\omega x\|^2 \le (1+\varepsilon)\|x\|^2\Big) \ge 1 - 2e^{-rc_0} \tag{2.9}$$

Such matrices are easy to come by as the next proposition shows [Ach03]. We will denote by $\mathcal{N}(0, \frac{1}{r})$ the Normal distribution with mean 0 and variance $\frac{1}{r}$.

Proposition 2.4.5. *Let $A_\omega \in \mathbb{R}^{r\times N}$ be a random matrix whose entries are chosen independently from either $\mathcal{N}(0, \frac{1}{r})$ or $\{\frac{-1}{\sqrt{r}}, \frac{1}{\sqrt{r}}\}$ Bernoulli. Then A_ω satisfies (2.9) with $c_0(\varepsilon) = \frac{\varepsilon^2}{4} - \frac{\varepsilon^3}{6}$.*

To prove the proposition from above in [Ach03], the author showed that for any $x \in \mathbb{R}^N$, the expectation of the random variable $\|A_\omega x\|^2$ is $\|x\|^2$. Then, it was proved that for any $x \in \mathbb{R}^N$ the random variable $\|A_\omega x\|^2$ is strongly concentrated about its expected value.

In what follows we will state and prove a simpler version of the JL lemma included in [Ach03]. This lemma states that any set of points from a high-dimensional Euclidean space can be embedded into a lower dimensional space without suffering great distortion.

Here and subsequently, the *union bound* will refer to the property which states that for any finite or countable set of events, the probability that at least one of the events happens is no greater than the sum of the probabilities of the individual events. That is, for a countable set of events $\{B_i\}_{i\in I}$, it holds that

$$\Pr\left(\bigcup_i B_i\right) \leq \sum_i \Pr(B_i).$$

This property follows from the fact that a probability measure is σ-sub-additive.

Lemma 2.4.6. *Let $\mathcal{F} = \{f_1, \ldots, f_m\}$ be a set of points in $\mathbb{R}^N$ and let $0 < \varepsilon < 1$. If $r > 24\varepsilon^{-2}\ln(m)$, there exists a matrix $A \in \mathbb{R}^{r\times N}$ such that*

$$(1-\varepsilon)\|f_i - f_j\|^2 \leq \|Af_i - Af_j\|^2 \leq (1+\varepsilon)\|f_i - f_j\|^2 \quad \forall\, 1 \leq i, j \leq m. \tag{2.10}$$

Proof. Let $A_\omega \in \mathbb{R}^{r\times N}$ be a random matrix with entries having any one of the two distributions from Proposition 2.4.5.

Using the union bound property, we have that

$$\begin{aligned}
& \Pr\Big((1-\varepsilon)\|f_i - f_j\|^2 \leq \|A_\omega(f_i - f_j)\|^2 \leq (1+\varepsilon)\|f_i - f_j\|^2 \quad \forall\, 1 \leq i, j \leq m\Big) \\
= \; & 1 - \Pr\Big(\big|\|A_\omega(f_i - f_j)\|^2 - \|f_i - f_j\|^2\big| \geq \varepsilon\|f_i - f_j\|^2 \quad \text{for some } 1 \leq i, j \leq m\Big) \\
\geq \; & 1 - \sum_{1\leq i,j\leq m} \Pr\Big(\big|\|A_\omega(f_i - f_j)\|^2 - \|f_i - f_j\|^2\big| \geq \varepsilon\|f_i - f_j\|^2\Big) \\
\geq \; & 1 - \sum_{1\leq i,j\leq m} 2e^{-rc_0} \\
= \; & 1 - m(m-1)e^{-rc_0}.
\end{aligned}$$

Recall from Proposition 2.4.5 that $c_0 = \frac{\varepsilon^2}{4} - \frac{\varepsilon^3}{6} \geq \frac{\varepsilon^2}{12}$. If $r > 24\varepsilon^{-2}\ln(m)$, it follows that $1 - m(m-1)e^{-rc_0} > 0$, thus (2.10) is verified with positive probability.

□

In this section we will use random matrices A_ω satisfying (2.9) to produce the lower dimensional data set $\mathcal{F}' = A_\omega\mathcal{F}$, with the aim of recovering with high probability an optimal partition for $\mathcal{F}$ using the optimal partition of $\mathcal{F}'$.

Below we will state the main results of Subsection 2.4.2 and we will give their proofs in Section 2.5.

Note that by Lemma 2.4.1, if $\mathcal{F} = \{f_1, \ldots, f_m\} \subseteq \mathbb{R}^N$ is $(l, k, 0)$-sparse, then $A_\omega\mathcal{F}$ is $(l, k, 0)$-sparse for all $\omega \in \Omega$. The following proposition is a generalization of Lemma 2.4.1 to the case where $\mathcal{F}$ is (l, k, ρ)-sparse with $\rho > 0$.

Proposition 2.4.7. *Assume the data $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq \mathbb{R}^N$ is (l, k, ρ)-sparse with $\rho > 0$. If $A_\omega \in \mathbb{R}^{r\times N}$ is a random matrix which satisfies (2.9), then $A_\omega\mathcal{F}$ is $(l, k, (1+\varepsilon)\rho)$-sparse with probability at least $1 - 2me^{-rc_0}$.*

Hence if the data is mapped with a random matrix which satisfies the concentration inequality, then with high probability, the sparsity of the transformed data is close to the sparsity of the original data. Further, as the following theorem shows, we obtain an estimation for the error between $\mathcal{F}$ and the bundle generated by the optimal partition for $\mathcal{F}' = A_\omega\mathcal{F}$.

Note that, given a constant $\alpha > 0$, the scaled data $\alpha\mathcal{F} = \{\alpha f_1, \dots, \alpha f_m\}$ satisfies that $e(\alpha\mathcal{F}, B) = \alpha^2 e(\mathcal{F}, B)$ for any bundle B. So, an optimal bundle for $\mathcal{F}$ is optimal for $\alpha\mathcal{F}$, and vice versa. Therefore, we can assume that the data $\mathcal{F} = \{f_1, \dots, f_m\}$ is *normalized*, that is, the matrix $M \in \mathbb{R}^{N\times m}$ which has the vectors $\{f_1, \dots, f_m\}$ as columns has unitary Frobenius norm. Recall that the Frobenius norm of a matrix $M \in \mathbb{R}^{N\times m}$ is defined by

$$\|M\|^2 := \sum_{i=1}^{N}\sum_{j=1}^{m} M_{i,j}^2, \tag{2.11}$$

where $M_{i,j}$ are the coefficients of M.

Theorem 2.4.8. *Let $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq \mathbb{R}^N$ be a normalized data set and $0 < \varepsilon < 1$. Assume that $A_\omega \in \mathbb{R}^{r\times N}$ is a random matrix satisfying (2.9) and $\mathbf{S}_\omega$ is an optimal partition for $\mathcal{F}' = A_\omega\mathcal{F}$ in $\mathbb{R}^r$. If B_ω is a bundle generated by the partition $\mathbf{S}_\omega$ and the data $\mathcal{F}$ in $\mathbb{R}^N$ as in Definition 2.3.3, then with probability exceeding $1 - (2m^2 + 4m)e^{-rc_0}$, we have*

$$e(\mathcal{F}, B_\omega) \le (1+\varepsilon)e_0(\mathcal{F}) + \varepsilon c_1, \tag{2.12}$$

where $c_1 = (l(d-k))^{1/2}$ and $d = \operatorname{rank}(\mathcal{F})$.

Finally, we can use this theorem to show that the set of matrices which are *η-admissible* (see definition below) is large.

The following definition generalizes Definition 2.4.2 to the ρ-sparse setting, with $\rho > 0$.

Definition 2.4.9. Assume $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq \mathbb{R}^N$ is (l, k, ρ)-sparse and let $0 < \eta < 1$. We will say that a matrix $A \in \mathbb{R}^{r\times N}$ is *η-admissible* for $\mathcal{F}$ if for any optimal partition $\mathbf{S}$ for $\mathcal{F}' = A\mathcal{F}$ in $\mathbb{R}^r$, the bundle $B_{\mathbf{S}}$ generated by $\mathbf{S}$ and $\mathcal{F}$ in $\mathbb{R}^N$, satisfies

$$e(\mathcal{F}, B_{\mathbf{S}}) \le \rho + \eta.$$

We have the following generalization of Proposition 2.4.3, which provides an estimate on the size of the set of η-admissible matrices.

Corollary 2.4.10. *Let $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq \mathbb{R}^N$ be a normalized data set and $0 < \eta < 1$. Assume that $A_\omega \in \mathbb{R}^{r\times N}$ is a random matrix which satisfies property (2.9) for $\varepsilon = \eta\,(1 + \sqrt{l(d-k)})^{-1}$. Then A_ω is η-admissible for $\mathcal{F}$ with probability at least $1-(2m^2+4m)e^{-rc_0(\varepsilon)}$.*

Proof. Using the fact that $e_0(\mathcal{F}) \leq \mathcal{E}(\mathcal{F}, \{0\}) = \|\mathcal{F}\|^2 = 1$, we conclude from Theorem 2.4.8 that

$$\Pr\Big(e(\mathcal{F}, B_\omega) \leq e_0(\mathcal{F}) + \varepsilon(1 + c_1)\Big) \geq 1 - c_2 e^{-rc_0(\varepsilon)}, \tag{2.13}$$

where $c_1 = (l(d-k))^{1/2}$, $d = \mathrm{rank}(\mathcal{F})$, and $c_2 = 2m^2 + 4m$. That is,

$$\Pr\Big(e(\mathcal{F}, B_\omega) \leq e_0(\mathcal{F}) + \eta\Big) \geq 1 - (2m^2 + 4m)e^{-rc_0(\varepsilon)}.$$

□

As a consequence of the previous corollary, we have a bound on the dimension of the lower dimensional space to obtain a bundle which produces an error at η-distance of the minimal error with high probability.

Now, using that $c_0(\varepsilon) \geq \frac{\varepsilon^2}{12}$ for random matrices with gaussian or Bernoulli entries (see Proposition 2.4.5), from Theorem 2.4.8 we obtain the following corollary.

Corollary 2.4.11. *Let $\eta, \delta \in (0,1)$, be given. Assume that $A_\omega \in \mathbb{R}^{r\times N}$ is a random matrix whose entries are as in Proposition 2.4.5.*

Then for every r satisfying,

$$r \geq \frac{12(1 + \sqrt{l(d-k)})^2}{\eta^2} \ln\left(\frac{2m^2 + 4m}{\delta}\right)$$

with probability at least $1 - \delta$ we have that

$$e(\mathcal{F}, B_\omega) \leq e_0(\mathcal{F}) + \eta.$$

We want to remark here that the results of Subsection 2.4.2 are valid for any probability distribution that satisfies the concentration inequality (2.9). The bound on the error is still valid for $\rho = 0$. However in that case we were able to obtain sharp results in Subsection 2.4.1.

2.5 Proofs

In this section we give the proofs for Subsection 2.4.2.

2.5.1 Background and supporting results

Before proving the results of the previous section we need several known theorems, lemmas, and propositions below.

Given $M \in \mathbb{R}^{m\times m}$ a symmetric matrix, let $\lambda_1(M) \geq \lambda_2(M) \geq \cdots \geq \lambda_m(M)$ be its eigenvalues and $s_1(M) \geq s_2(M) \geq \cdots \geq s_m(M) \geq 0$ be its singular values.

Recall that the Frobenius norm defined in (2.11) satisfies that

$$\|M\|^2 = \sum_{1\le i,j\le m} M_{i,j}^2 = \sum_{i=1}^{m} s_i^2(M),$$

where $M_{i,j}$ are the coefficients of M.

Theorem 2.5.1. *[Bha97, Theorem III.4.1]*

Let $A, B \in \mathbb{R}^{m\times m}$ be symmetric matrices. Then for any choice of indices $1 \le i_1 < i_2 < \cdots < i_k \le m$,

$$\sum_{j=1}^{k} (\lambda_{i_j}(A) - \lambda_{i_j}(B)) \le \sum_{j=1}^{k} \lambda_j(A - B).$$

Corollary 2.5.2. *Let $A, B \in \mathbb{R}^{m\times m}$ be symmetric matrices. Assume k and d are two integers which satisfy $0 \le k \le d \le m$, then*

$$\Big|\sum_{j=k+1}^{d} (\lambda_j(A) - \lambda_j(B))\Big| \le (d-k)^{1/2}\|A - B\|.$$

Proof. Since $A-B$ is symmetric, it follows that for each $1 \le j \le m$ there exists $1 \le i_j \le m$ such that

$$|\lambda_j(A - B)| = s_{i_j}(A - B).$$

From this and Theorem 2.5.1 we have

$$\begin{aligned}
\sum_{j=k+1}^{d} (\lambda_j(A) - \lambda_j(B)) &\le \sum_{j=1}^{d-k} \lambda_j(A - B) \le \sum_{j=1}^{d-k} s_{i_j}(A - B) \\
&\le \sum_{j=1}^{d-k} s_j(A - B) \le (d-k)^{1/2}\Big(\sum_{j=1}^{d-k} s_j^2(A - B)\Big)^{1/2} \\
&\le (d-k)^{1/2}\|A - B\|.
\end{aligned}$$

□

Remark 2.5.3. Note that the bound of the previous corollary is sharp. Indeed, let $A \in \mathbb{R}^{m\times m}$ be the diagonal matrix with coefficients $a_{ii} = 2$ for $1 \le i \le d$, and $a_{ii} = 0$ otherwise. Let $B \in \mathbb{R}^{m\times m}$ be the diagonal matrix with coefficients $b_{ii} = 2$ for $1 \le i \le k$, $b_{ii} = 1$ for $k+1 \le i \le d$, and $b_{ii} = 0$ otherwise. Thus,

$$\Big|\sum_{j=k+1}^{d} (\lambda_j(A) - \lambda_j(B))\Big| = \Big|\sum_{j=k+1}^{d} (2-1)\Big| = d - k.$$

Further $\|A - B\| = (d-k)^{1/2}$, and therefore

$$\Big|\sum_{j=k+1}^{d} (\lambda_j(A) - \lambda_j(B))\Big| = (d-k)^{1/2}\|A - B\|.$$

The next lemma was stated in [AV06], but we will give its proof since it shows an important property satisfied by matrices verifying the concentration inequality.

Lemma 2.5.4. *[AV06] Suppose that $A_\omega \in \mathbb{R}^{r\times N}$ is a random matrix which satisfies (2.9) and $u, v \in \mathbb{R}^N$, then*

$$|\langle u, v\rangle - \langle A_\omega u, A_\omega v\rangle| \leq \varepsilon \|u\|\|v\|,$$

with probability at least $1 - 4e^{-rc_0}$.

Proof. It suffices to show that for $u, v \in \mathbb{R}^N$ such that $\|u\| = \|v\| = 1$ we have that

$$|\langle u, v\rangle - \langle A_\omega u, A_\omega v\rangle| \leq \varepsilon,$$

with probability at least $1 - 4e^{-rc_0}$.

Applying (2.9) to the vectors $u+v$ and $u-v$ we obtain with probability at least $1-4e^{-rc_0}$ that

$$(1-\varepsilon)\|u-v\|^2 \leq \|A_\omega(u-v)\|^2 \leq (1+\varepsilon)\|u-v\|^2$$

and

$$(1-\varepsilon)\|u+v\|^2 \leq \|A_\omega(u+v)\|^2 \leq (1+\varepsilon)\|u+v\|^2.$$

Thus,

$$\begin{aligned} 4\langle A_\omega u, A_\omega v\rangle &= \|A_\omega(u+v)\|^2 - \|A_\omega(u-v)\|^2 \\ &\geq (1-\varepsilon)\|u+v\|^2 - (1-\varepsilon)\|u-v\|^2 \\ &= 4\langle u, v\rangle - 2\varepsilon(\|u\|^2 + \|v\|^2) \\ &= 4\langle u, v\rangle - 4\varepsilon. \end{aligned}$$

The other inequality follows similarly.

□

The following proposition was proved in [Sar08], but we include its proof for the sake of completeness.

Proposition 2.5.5. *Let $A_\omega \in \mathbb{R}^{r\times N}$ be a random matrix which satisfies (2.9) and let $M \in \mathbb{R}^{N\times m}$ be a matrix. Then, we have*

$$\|M^*M - M^*A_\omega^*A_\omega M\| \leq \varepsilon\|M\|^2,$$

with probability at least $1 - 2(m^2 + m)e^{-rc_0}$.

Proof. Set $Y_{i,j}(\omega) = (M^*M - M^*A_\omega^*A_\omega M)_{i,j} = \langle f_i, f_j\rangle - \langle A_\omega f_i, A_\omega f_j\rangle$. By Lemma 2.5.4 with probability at least $1 - 4e^{-rc_0}$ we have that

$$|Y_{i,j}(\omega)| \leq \varepsilon\|f_i\|\|f_j\| \tag{2.14}$$

Note that if (2.14) holds for all $1 \leq i \leq j \leq m$, then

$$\begin{aligned}||M^*M - M^*A_\omega^*A_\omega M||^2 &= \sum_{1\le i,j\le m} Y_{i,j}(\omega)^2 \\ &\le \varepsilon^2 \sum_{1\le i,j\le m} ||f_i||^2||f_j||^2 = \varepsilon^2||M||^4.\end{aligned}$$

Thus, by the union bound, we obtain

$$\begin{aligned}&\Pr\Big(||M^*M - M^*A_\omega^*A_\omega M|| \le \varepsilon||M||^2\Big) \\ &\ge \Pr\Big(|Y_{i,j}(\omega)| \le \varepsilon||f_i||||f_j|| \quad \forall\, 1 \le i \le j \le m\Big) \\ &\ge 1 - \textstyle\sum_{1\le i\le j\le m} 4e^{-rc_0} = 1 - 2(m^2+m)e^{-rc_0}.\end{aligned}$$

□

2.5.2 New results and proof of Theorem 2.4.8

Given a set of vectors $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq \mathbb{R}^N$ let $\mathcal{E}_0(\mathcal{F}, C_k)$ be as in (2.2), that is

$$\mathcal{E}_0(\mathcal{F}, C_k) = \inf\{\mathcal{E}(\mathcal{F}, V) \ :\ V \text{ is a subspace with } \dim(V) \le k\},$$

where $\mathcal{E}(\mathcal{F}, V) = \sum_{i=1}^m d^2(f_i, V)$. For simplicity of notation, we will write $\mathcal{E}_0(\mathcal{F}, C_k)$ as $\mathcal{E}_k(\mathcal{F})$.

Assume $M \in \mathbb{R}^{N\times m}$ is the matrix with columns $\mathcal{F} = \{f_1, \dots, f_m\}$. If $d := \operatorname{rank}(M)$, recall that Theorem 2.2.2 (Eckart-Young) states that

$$\mathcal{E}_k(\mathcal{F}) = \sum_{j=k+1}^{d} \lambda_j(M^*M), \tag{2.15}$$

where $\lambda_1(M^*M) \ge \cdots \ge \lambda_d(M^*M) > 0$ are the positive eigenvalues of M^*M.

Lemma 2.5.6. *Assume that $M \in \mathbb{R}^{N\times m}$ and $A \in \mathbb{R}^{r\times N}$ are arbitrary matrices. Let $S \in \mathbb{R}^{N\times s}$ be a submatrix of M. If $d := \operatorname{rank}(M)$ is such that $0 \le k \le d$, then*

$$|\mathcal{E}_k(\mathcal{S}) - \mathcal{E}_k(A\mathcal{S})| \le (d-k)^{1/2}\, ||S^*S - S^*A^*AS||,$$

where $\mathcal{S} \subseteq \mathbb{R}^N$ is the set formed by the columns of S.

Proof. Let $d_s := \operatorname{rank}(S)$. We have $\operatorname{rank}(AS) \le d_s$. If $d_s \le k$, the result is trivial. Otherwise by (2.15) and Corollary 2.5.2, we obtain

$$\begin{aligned}|\mathcal{E}_k(\mathcal{S}) - \mathcal{E}_k(A\mathcal{S})| &= \Big|\sum_{j=k+1}^{d_s} (\lambda_j(S^*S) - \lambda_j(S^*A^*AS))\Big| \\ &\le (d_s - k)^{1/2}||S^*S - S^*A^*AS||.\end{aligned}$$

As S is a submatrix of M, we have that $(d_s - k)^{1/2} \leq (d-k)^{1/2}$, which proves the lemma. ☐

Recall that $e_0(\mathcal{F})$ is the optimal value for the data $\mathcal{F}$, and $e_0(A_\omega \mathcal{F})$ is the optimal value for the data $\mathcal{F}' = A_\omega \mathcal{F}$ (see (2.4)). A relation between these two values is given by the following lemma.

Lemma 2.5.7. *Let $\mathcal{F} = \{f_1, \ldots, f_m\} \subseteq \mathbb{R}^N$ and $0 < \varepsilon < 1$. If $A_\omega \in \mathbb{R}^{r\times N}$ is a random matrix which satisfies (2.9), then with probability exceeding $1 - 2me^{-rc_0}$, we have*

$$e_0(A_\omega \mathcal{F}) \leq (1+\varepsilon) e_0(\mathcal{F}).$$

Proof. Let $V \subseteq \mathbb{R}^N$ be a subspace. Using (2.9) and the union bound, with probability at least $1 - 2me^{-rc_0}$ we have that

$$\begin{aligned}\mathcal{E}(A_\omega \mathcal{F}, A_\omega V) &= \sum_{i=1}^m d^2(A_\omega f_i, A_\omega V) \leq \sum_{i=1}^m \|A_\omega f_i - A_\omega(P_V f_i)\|^2 \\ &\leq (1+\varepsilon)\sum_{i=1}^m \|f_i - P_V f_i\|^2 = (1+\varepsilon)\mathcal{E}(\mathcal{F}, V),\end{aligned}$$

where P_V is the orthogonal projection onto V.

Assume that $\mathbf{S} = \{S_1, \ldots, S_l\}$ is an optimal partition for $\mathcal{F}$ and $\{V_1, \ldots, V_l\}$ is an optimal bundle for $\mathcal{F}$. Let $\mathcal{F}_i = \{f_j\}_{j \in S_i}$. From what has been proved above and the union bound, with probability exceeding $1 - \sum_{i=1}^l 2m_i e^{-rc_0} = 1 - 2me^{-rc_0}$, it holds

$$e_0(A_\omega \mathcal{F}) \leq \sum_{i=1}^l \mathcal{E}(A_\omega \mathcal{F}_i, A_\omega V_i) \leq (1+\varepsilon)\sum_{i=1}^l \mathcal{E}(\mathcal{F}_i, V_i) = (1+\varepsilon)e_0(\mathcal{F}).$$

☐

Proof of Proposition 2.4.7. This is a direct consequence of Lemma 2.5.7. ☐

Proof of Theorem 2.4.8. Let $\mathbf{S}_\omega = \{S_\omega^1, \ldots, S_\omega^l\}$ and $\mathcal{F}_\omega^i = \{f_j\}_{j \in S_\omega^i}$. Since $B_\omega = \{V_\omega^1, \ldots, V_\omega^l\}$ is generated by $\mathbf{S}_\omega$ and $\mathcal{F}$, it follows that $\mathcal{E}(\mathcal{F}_\omega^i, V_\omega^i) = \mathcal{E}_k(\mathcal{F}_\omega^i)$. And as $\mathbf{S}_\omega$ is an optimal partition for $A_\omega \mathcal{F}$ in $\mathbb{R}^r$, we have that $\sum_{i=1}^l \mathcal{E}_k(A_\omega \mathcal{F}_\omega^i) = e_0(A_\omega \mathcal{F})$.

Let $m_\omega^i = \#(S_\omega^i)$ and $M_\omega^i \in \mathbb{R}^{N\times m_\omega^i}$ be the matrices which have $\{f_j\}_{j\in S_\omega^i}$ as columns. Using Lemma 2.5.6, Lemma 2.5.7, and Proposition 2.5.5, with high probability it holds that

$$\begin{aligned}
e(\mathcal{F}, B_\omega) &\leq \sum_{i=1}^{l} \mathcal{E}(\mathcal{F}_\omega^i, V_\omega^i) = \sum_{i=1}^{l} \mathcal{E}_k(\mathcal{F}_\omega^i) \\
&\leq \sum_{i=1}^{l} \mathcal{E}_k(A_\omega \mathcal{F}_\omega^i) + (d-k)^{1/2} \sum_{i=1}^{l} \|M_\omega^{i*} M_\omega^i - M_\omega^{i*} A_\omega^* A_\omega M_\omega^i\| \\
&\leq e_0(A_\omega \mathcal{F}) + (l(d-k))^{1/2} \Big(\sum_{i=1}^{l} \|M_\omega^{i*} M_\omega^i - M_\omega^{i*} A_\omega^* A_\omega M_\omega^i\|^2 \Big)^{1/2} \\
&\leq (1+\varepsilon) e_0(\mathcal{F}) + (l(d-k))^{1/2} \|M^* M - M^* A_\omega^* A_\omega M\| \\
&\leq (1+\varepsilon) e_0(\mathcal{F}) + \varepsilon (l(d-k))^{1/2},
\end{aligned}$$

where $M \in \mathbb{R}^{N\times m}$ is the unitary Frobenius norm matrix which has the vectors $\{f_1, \ldots, f_m\}$ as columns.

The right side of (2.12) follows from Proposition 2.5.5, Lemma 2.5.7, and the fact that

$$\begin{aligned}
&\Pr\Big(e(\mathcal{F}, B_\omega) \leq (1+\varepsilon) e_0(\mathcal{F}) + \varepsilon (l(d-k))^{1/2}\Big) \\
&\quad \geq \Pr\Big(\|M^* M - M^* A_\omega^* A_\omega M\| \leq \varepsilon \text{ and } e_0(A_\omega \mathcal{F}) \leq (1+\varepsilon) e_0(\mathcal{F})\Big) \\
&\quad \geq 1 - (2(m^2+m)e^{-rc_0} + 2me^{-rc_0}) = 1 - (2m^2 + 4m)e^{-rc_0}.
\end{aligned}$$

□

3

Sampling in a union of frame generated subspaces

3.1 Introduction

In the previous chapter we have studied the problem of finding a union of subspaces that best explains a data set. Our goal in this chapter is to study the sampling process for signals which lie in this kind of models.

We will begin by describing the problem of sampling in a union of subspaces. Assume $\mathcal{H}$ is a separable Hilbert space and $\{V_\gamma\}_{\gamma\in\Gamma}$ are closed subspaces in $\mathcal{H}$, with Γ an arbitrary index set. Let $\mathcal{X}$ be the union of subspaces defined as

$$\mathcal{X} := \bigcup_{\gamma\in\Gamma} V_\gamma.$$

Suppose now that a signal x is extracted from $\mathcal{X}$ and we take some measurements of that signal. These measurements can be thought of as the result of the application of a series of functionals $\{\varphi_i\}_{i\in I}$ to our signal x. The problem is then to reconstruct the signal using only the measurements $\{\varphi_i(x)\}_{i\in I}$ and some description of the subspaces in $\mathcal{X}$.

Assume the series of functionals define an operator, *the sampling operator*,

$$A : \mathcal{H} \to \ell^2(I), \quad Ax := \{\varphi_i(x)\}_{i\in I}.$$

From the Riesz's representation theorem [Con90], there exists a unique set of vectors $\Psi := \{\psi_i\}_{i\in I}$, such that

$$Ax = \{\langle x, \psi_i\rangle\}_{i\in I}.$$

The sampling problem consists of reconstructing a signal $x \in \mathcal{X}$ using the data $\{\langle x, \psi_i\rangle\}_{i\in I}$. The first thing required is that the signals are uniquely determined by the data. That is, the sampling operator A should be one-to-one on $\mathcal{X}$. Another important property that is usually required for a sampling operator, is stability. That is, the existence of two constants $0 < \alpha \le \beta < +\infty$ such that

$$\alpha\|x_1 - x_2\|^2_{\mathcal{H}} \le \|Ax_1 - Ax_2\|^2_{\ell^2(I)} \le \beta\|x_1 - x_2\|^2_{\mathcal{H}} \quad \forall\, x_1, x_2 \in \mathcal{X}.$$

This is crucial to bound the error of reconstruction in noisy situations.

Under some hypothesis on the structure of the subspaces, Lu and Do [LD08] found necessary and sufficient conditions on Ψ in order for the sampling operator A to be one-to-one and stable when restricted to the union of the subspaces $\mathcal{X}$. These conditions were obtained in two settings. In the euclidean space and in $L^2(\mathbb{R}^d)$. In this latter case the subspaces considered were finitely generated shift-invariant spaces.

There are two technical aspects in the approach of Lu and Do that restrict the applicability of their results in the shift-invariant space case. The first one is due to the fact that the conditions are obtained in terms of Riesz bases of translates of the SISs involved, and it is well known that not every SIS has a Riesz basis of translates (see Example 1.5.11). The second one is that the approach is based upon the sum of every two of the SISs in the union. The conditions on the sampling operator are then obtained using fiberization techniques on that sum. This requires that the sum of each of two subspaces is a closed subspace, which is not true in general.

In this chapter we obtain the conditions for the sampling operator to be one-to-one and stable in terms of *frames* of translates of the SISs instead of orthonormal basis. This extends the previous results to arbitrary SISs and in particular removes the restrictions mentioned above.

We will obtain necessary and sufficient conditions for the stability of the sampling operator A in a union of arbitrary SISs. We will show that, without the assumption of the closedness of the sum of every two of the SISs in the union, we can only obtain sufficient conditions for the injectivity of A.

On the other side, in Chapter 4, using known results from the theory of SISs, we will obtain necessary and sufficient conditions for the closedness of the sum of two shift-invariant spaces. Using this, we will determine families of subspaces on which the conditions for injectivity are necessary and sufficient.

This chapter is organized in the following way: Section 3.2 contains some preliminary results that will be used throughout. In Section 3.3 we set the problem of sampling in a union of subspaces in the general context of an abstract Hilbert space. We also give injectivity and stability conditions for the sampling operator, within this general setting. The case of finite-dimensional subspaces is studied in Section 3.4. Finally in Section 3.5 we analyze the problem for the Hilbert space $L^2(\mathbb{R}^d)$ and sampling in a union of finitely generated shift-invariant spaces.

3.2 Preliminaries

Let us define here an operator which will be useful to develop the sampling theory in a union of subspaces.

Definition 3.2.1. Assume I, J are countable index sets. Suppose $X := \{x_j\}_{j\in J}$ and $Y := \{y_i\}_{i\in I}$ are Bessel sequences in a separable Hilbert space $\mathcal{H}$. Let B_X and B_Y be the analysis operators (see Definition 1.2.9) associated to X and Y respectively. The *cross-correlation* operator is defined by

$$G_{X,Y} : \ell^2(J) \to \ell^2(I), \quad G_{X,Y} := B_Y B_X^*. \tag{3.1}$$

Identifying $G_{X,Y}$ with its matrix representation, we write

$$(G_{X,Y})_{i,j} = \langle x_j, y_i \rangle \quad \forall\, j \in J, \forall\, i \in I.$$

In this chapter we will need the following corollary which is a consequence of Proposition 1.2.13. Its proof is straightforward using that the frame operator of a Parseval frame is the identity operator.

Corollary 3.2.2. *If $X = \{x_j\}_{j\in J}$ is a Parseval frame for a closed subspace V of $\mathcal{H}$, and B_X is the analysis operator associated to X, then the orthogonal projection of $\mathcal{H}$ onto V is*

$$P_V = B_X^* B_X : \mathcal{H} \to \mathcal{H}, \quad P_V h = \sum_{j\in J} \langle h, x_j \rangle\, x_j.$$

3.3 The sampling operator

Let $\mathcal{H}$ be a separable Hilbert space and $V \subseteq \mathcal{H}$ an arbitrary set. Given $\Psi = \{\psi_i\}_{i\in I}$ a Bessel sequence in $\mathcal{H}$, the sampling problem consists of reconstructing a signal $f \in V$ using the data $\{\langle f, \psi_i \rangle\}_{i\in I}$. We first require that the signals are uniquely determined by the data. That is, if we define the *Sampling operator* by

$$A : \mathcal{H} \to \ell^2(I), \quad Af := \{\langle f, \psi_i \rangle\}_{i\in I}, \tag{3.2}$$

we require A to be one-to-one on V. The set Ψ will be called the *Sampling set*.

Note that the sampling operator A is the analysis operator (see Definition 1.2.9) for the sequence Ψ.

Another important property that is usually required for a sampling operator, is stability. This is crucial to bound the error of reconstruction in noisy situations.

The stable sampling condition was first proposed by [Lan67] for the case when V is the Paley-Wiener space. It was then generalized in [LD08] to the case when V is a union of subspaces.

Definition 3.3.1. A sampling operator A is called *stable* on V if there exist two constants $0 < \alpha \le \beta < +\infty$ such that

$$\alpha \|x_1 - x_2\|_{\mathcal{H}}^2 \le \|Ax_1 - Ax_2\|_{\ell^2(I)}^2 \le \beta \|x_1 - x_2\|_{\mathcal{H}}^2 \quad \forall\, x_1, x_2 \in V.$$

When V is a closed subspace, the injectivity and the stability can be expressed in terms of conditions on $P_V\Psi$, where P_V is the orthogonal projection of $\mathcal{H}$ onto V.

Proposition 3.3.2. *Let $\mathcal{H}$ be a Hilbert space, $V \subseteq \mathcal{H}$ a closed subspace and $\Psi = \{\psi_i\}_{i\in I}$ a Bessel sequence in $\mathcal{H}$. If A is the sampling operator associated to Ψ, then we have*

i) *The operator A is* one-to-one *on V if and only if $\{P_V\psi_i\}_{i\in I}$ is complete in V, that is $V = \overline{\operatorname{span}}\{P_V\psi_i\}_{i\in I}$.*

ii) *The operator A is* stable *on V with constants α and β if and only if $\{P_V\psi_i\}_{i\in I}$ is a frame for V with constants α and β.*

Proof. The proof of *i)* is straightforward using that if $f \in V$ then

$$\langle f, P_V\psi_i\rangle = \langle P_V f, \psi_i\rangle = \langle f, \psi_i\rangle.$$

For *ii)* note that for all $f \in V$

$$\|Af\|^2_{\ell^2(I)} = \sum_{i\in I} |\langle f, \psi_n\rangle|^2 = \sum_{i\in I} |\langle P_V f, \psi_n\rangle|^2 = \sum_{i\in I} |\langle f, P_V\psi_n\rangle|^2.$$

□

Remark 3.3.3. Given a closed subspace V in a Hilbert space $\mathcal{H}$, a sequence of vectors $\{\psi_i\}_{i\in I} \subseteq \mathcal{H}$ is called an *outer frame* for V if $\{P_V\psi_i\}_{i\in I}$ is a frame for V. The notion of outer frame was introduced in [ACM04]. See also [FW01] and [LO04] for related definitions. Using this terminology, part ii) of Proposition 3.3.2 says that the sampling operator A is stable if and only if $\{\psi_i\}$ is an outer frame for V.

In what follows we will extend one-to-one and stability conditions for the operator A, to the case of a union of subspaces instead of a single subspace.

If $\{V_\gamma\}_{\gamma\in\Gamma}$ are closed subspaces of $\mathcal{H}$, with Γ an arbitrary index set. Let

$$\mathcal{X} := \bigcup_{\gamma\in\Gamma} V_\gamma.$$

We want to study conditions on Ψ so that the sampling operator A defined by (3.2) is one-to-one and stable on $\mathcal{X}$.

This study continues the one initiated by Lu and Do [LD08] in which they translated the conditions on $\mathcal{X}$ into conditions on the subspaces defined by

$$V_{\gamma,\theta} := V_\gamma + V_\theta = \{x + y \,:\, x \in V_\gamma,\ y \in V_\theta\}. \tag{3.3}$$

Working with the subspaces $V_{\gamma,\theta}$ instead of $\mathcal{X}$, allows to exploit lineal properties of A.

They proved the following proposition, we will include here its proof for the sake of completeness.

Proposition 3.3.4. *[LD08] With the above notation we have,*

i) *The operator A is one-to-one on $\mathcal{X}$ if and only if A is one-to-one on every $V_{\gamma,\theta}$ with $\gamma, \theta \in \Gamma$.*

ii) The operator A is stable for $\mathcal{X}$ with stability bounds α and β, if and only if A is stable for $V_{\gamma,\theta}$ with stability bounds α and β for all $\gamma, \theta \in \Gamma$, i.e.

$$\alpha\|x\|^2_{\mathcal{H}} \leq \|Ax\|^2_{\ell^2(I)} \leq \beta\|x\|^2_{\mathcal{H}} \quad \forall\, x \in V_{\gamma,\theta}, \forall\, \gamma, \theta \in \Gamma.$$

Proof. We first prove part i). Assume that A is one-to-one on $\mathcal{X}$. Given $\gamma, \theta \in \Gamma$, $V_{\gamma,\theta}$ is a subspace. Thus, for proving the injectivity of A on $V_{\gamma,\theta}$ it suffices to show that for $x \in V_{\gamma,\theta}$, $Ax = 0$ implies $x = 0$.

Since $x \in V_{\gamma,\theta}$, there exist $x_1 \in V_\gamma$ and $x_2 \in V_\theta$ such that $x = x_1 + x_2$. Hence $Ax_1 = A(-x_2)$ for $x_1, x_2 \in \mathcal{X}$. Therefore $x_1 = -x_2$, so $x = x_1 + x_2 = 0$.

Suppose now that A is one-to-one on every $V_{\gamma,\theta}$ with $\gamma, \theta \in \Gamma$. Let $x_1, x_2 \in \mathcal{X}$ such that $Ax_1 = Ax_2$. There exist $\gamma, \theta \in \Gamma$ such that $x_1 \in V_\gamma$ and $x_2 \in V_\theta$. So, $A(x_1 - x_2) = 0$ and $x_1 - x_2 \in V_{\gamma,\theta}$. Hence, $x_1 - x_2 = 0$, which implies that $x_1 = x_2$.

Using the same arguments from above the proof of ii) follows easily.

□

The sum of two closed infinite-dimensional subspaces of a Hilbert space is not necessarily closed (see Example 4.4.6). Furthermore, the injectivity of an operator on a subspace does not imply the injectivity on its closure. So, we can not apply Proposition 3.3.2 to the subspaces $V_{\gamma,\theta}$. However, we can obtain a sufficient condition for the injectivity.

Proposition 3.3.5. *If $\{P_{\overline{V}_{\gamma,\theta}}\psi_i\}_{i\in I}$ is complete on $\overline{V}_{\gamma,\theta}$ for every $\gamma, \theta \in \Gamma$, then A is one-to-one on $\mathcal{X}$.*

When the subspaces of the family $\{V_{\gamma,\theta}\}_{\gamma,\theta\in\Gamma}$ are all closed, the condition in Proposition 3.3.5, will be also necessary for the injectivity of A on $\mathcal{X}$. So, a natural question will be, when the sum of two closed subspaces of a Hilbert space is closed. In Chapter 4 we study this problem in several situations.

In the case of the stability, Proposition 3.3.2 can be applied since, by the boundedness of A, we have the following.

Proposition 3.3.6. *Let V be a subspace of $\mathcal{H}$, the operator A is stable for V with constants α and β if and only if it is stable for $\overline{V}$ with constants α and β.*

As a consequence of this, using Propositions 3.3.2 and part ii) of Proposition 3.3.4, we have

Proposition 3.3.7. *A is stable for $\mathcal{X}$ with constants α and β if and only if $\{P_{\overline{V}_{\gamma,\theta}}\psi_i\}_{i\in I}$ is a frame for $\overline{V}_{\gamma,\theta}$ for every $\gamma, \theta \in \Gamma$ with the same constants α and β.*

3.4 Union of finite-dimensional subspaces

In this section we will first obtain conditions on the sequence $\{\psi_i\}_{i\in I}$ for the sampling operator to be one-to-one on a union of finite-dimensional subspaces. We will then analyze the stability requirements. We are interested in expressing these conditions in terms of the generators of the sum of every two subspaces of the union.

3.4.1 The one-to-one condition for the sampling operator

Let $\mathcal{H}$ be a Hilbert space, $\Psi = \{\psi_i\}_{i\in I}$ a Bessel sequence in $\mathcal{H}$, and A the sampling operator associated to Ψ as in (3.2).

Let V be a finite-dimensional subspace of $\mathcal{H}$ and $\Phi = \{\varphi_j\}_{j=1}^m$ a finite frame for V. (Recall that a finite set of vectors from a finite-dimensional subspace is a frame for that subspace if and only if it spans it, see Remark 1.2.12.)

The cross-correlation operator associated to Ψ and Φ (see (3.1)) in this case can be written as,

$$G_{\Phi,\Psi} : \mathbb{C}^m \to \ell^2(I), \quad G_{\Phi,\Psi} = AB_\Phi^*,$$

where $B_\Phi^* : \mathbb{C}^m \to \mathcal{H}$ is the synthesis operator associated to Φ.

The next theorem gives necessary and sufficient conditions on the cross-correlation operator for the sampling operator to be one-to-one on V.

Theorem 3.4.1. *Let $\Psi = \{\psi_i\}_{i\in I}$ be a Bessel sequence for $\mathcal{H}$, V a finite-dimensional subspace of $\mathcal{H}$ and $\Phi = \{\varphi_j\}_{j=1}^m$ a frame for V. Then the following are equivalent:*

i) *Ψ provides a one-to-one sampling operator on V.*

ii) $\ker(G_{\Phi,\Psi}) = \ker(B_\Phi^*)$.

iii) $\dim(\text{range}(G_{\Phi,\Psi})) = \dim(V)$.

Proof. The proof is straightforward using that the range of the operator B_Φ^* is V. □

Remark 3.4.2. Note that the conditions in Theorem 3.4.1 do not depend on the particular chosen frame. That is, if there exists a frame Φ for V, such that $\dim(\text{range}(G_{\Phi,\Psi})) = \dim(V)$, then $\dim(\text{range}(G_{\widetilde{\Phi},\Psi})) = \dim(V)$, for any frame $\widetilde{\Phi}$ for V.

Now we will apply the previous theorem for the case of a union of subspaces.

Let $\{V_\gamma\}_{\gamma\in\Gamma}$ be a collection of finite-dimensional subspaces of $\mathcal{H}$, with Γ an arbitrary index set. Define,

$$\mathcal{X} := \bigcup_{\gamma\in\Gamma} V_\gamma.$$

As before, set $V_{\gamma,\theta} := V_\gamma + V_\theta$.

We obtain the following result which extends the result in [LD08] to the case that the subspaces in the union are described by frames.

Theorem 3.4.3. *Let $\Psi = \{\psi_i\}_{i\in I}$ be a Bessel sequence for $\mathcal{H}$ and for every $\gamma, \theta \in \Gamma$, let $\Phi_{\gamma,\theta}$ be a frame for $V_{\gamma,\theta}$, the following are equivalent:*

i) Ψ *provides a one-to-one sampling operator on* $\mathcal{X}$.

ii) $\dim(\mathrm{range}(G_{\Phi_{\gamma,\theta},\Psi})) = \dim(V_{\gamma,\theta})$ *for all* $\gamma, \theta \in \Gamma$.

Note that if I is a finite set, the problem of testing the injectivity of A on $\mathcal{X}$ reduces to check that the rank of the cross-correlation matrices are equal to the dimension of the subspaces $V_{\gamma,\theta}$.

In this case a lower bound for the cardinality of the sampling set can be established. This is stated in the following corollary from [LD08]. We include a proof of the result based on Theorem 3.4.3.

Corollary 3.4.4. *If the operator A is one-to-one on $\mathcal{X}$ and I is finite, then*

$$\#I \geq \sup_{\gamma,\theta\in\Gamma}(\dim(V_{\gamma,\theta})).$$

Proof. Since I is finite, we have that $\mathrm{range}(G_{\Phi_{\gamma,\theta},\Psi}) \subseteq \mathbb{C}^{\#I}$. Thus, using part ii) of Theorem 3.4.3, we obtain that

$$\dim(V_{\gamma,\theta}) = \dim(\mathrm{range}(G_{\Phi_{\gamma,\theta},\Psi})) \leq \#I, \quad \forall\, \gamma, \theta \in \Gamma.$$

□

3.4.2 The stability condition for the sampling operator

We are now interested in studying conditions for stability of the sampling operator. These conditions will be set in terms of the cross-correlation operator. We will consider Parseval frames to obtain simpler conditions.

Given Hilbert spaces $\mathcal{K}$ and $\mathcal{L}$ and a bounded linear operator $T : \mathcal{K} \to \mathcal{L}$, we denote by $\sigma^2(T)$ the set

$$\sigma^2(T) = \sigma(T^*T).$$

Theorem 3.4.5. *Let $\Psi = \{\psi_i\}_{i\in I}$ be a Bessel sequence for $\mathcal{H}$, V a finite-dimensional subspace of $\mathcal{H}$ and Φ a Parseval frame for V.*

The sequence Ψ provides a stable sampling operator for V with constants α and β if and only if

i) $\dim(\mathrm{range}(G_{\Phi,\Psi})) = \dim(V)$ *and*

ii) $\sigma^2(G_{\Phi,\Psi}) \subseteq \{0\} \cup [\alpha, \beta]$.

Proof. Let $W : \mathcal{H} \to \ell^2(I)$, be the analysis operator associated to $P_V\Psi$. For $x \in \mathcal{H}$, the equation,

$$Wx = \{\langle x, P_V\psi_i\rangle\}_{i\in I} = \{\langle P_V x, \psi_i\rangle\}_{i\in I} = AP_V x,$$

shows that $W = AP_V$.

Since Φ is a Parseval frame for V, by Proposition 3.2.2, $P_V = B_\Phi^* B_\Phi$ then,

$$G_{P_V\Psi} = WW^* = AP_VP_VA^* = AP_VA^* = AB_\Phi^*B_\Phi A^* = G_{\Phi,\Psi}G_{\Phi,\Psi}^*. \tag{3.4}$$

Let us assume first that A is stable for V. Item i) follows from Theorem 3.4.1. Now we prove ii).

Since V is closed (V is finite dimensional) then Proposition 3.3.2 gives that $P_V\Psi := \{P_V\psi_i\}_{i\in I}$ is a frame for V with constants α and β. Using Theorem 1.3.4, we have,

$$\sigma(G_{P_V\Psi}) \subseteq \{0\} \cup [\alpha, \beta].$$

So, by (3.4),

$$\sigma(G_{P_V\Psi}) = \sigma(G_{\Phi,\Psi}G_{\Phi,\Psi}^*) \subseteq \{0\} \cup [\alpha, \beta].$$

Finally, since (see [Rud91])

$$\sigma(G_{\Phi,\Psi}^*G_{\Phi,\Psi}) \subseteq \{0\} \cup \sigma(G_{\Phi,\Psi}G_{\Phi,\Psi}^*),$$

it follows that

$$\sigma^2(G_{\Phi,\Psi}) \subseteq \{0\} \cup [\alpha, \beta].$$

Suppose now that i) and ii) hold. Recall that A is stable for V with stability bounds α, β if and only if $P_V\Psi := \{P_V\psi_i\}_{i\in I}$ is a frame for V with frame bounds α, β.

By Theorem 3.4.1, condition i) implies that the sampling operator is one-to-one on V. Therefore, using Proposition 3.3.2, $P_V\Psi := \{P_V\psi_i\}_{i\in I}$ is complete in V.

That $P_V\Psi := \{P_V\psi_i\}_{i\in I}$ is a frame sequence is straightforward by ii), (3.4) and Theorem 1.3.4. □

Remark 3.4.6. As in the case of injectivity, we note that the condition of stability does not depend on the chosen Parseval frame. That means, if condition i) and ii) in the previous theorem hold for a Parseval frame Φ for V, then they hold for any Parseval frame $\widetilde{\Phi}$ for V.

Theorem 3.4.5 applied to the union of subspaces gives:

Theorem 3.4.7. *Let $\Psi = \{\psi_i\}_{i\in I}$ be a set of sampling vectors and for every $\gamma, \theta \in \Gamma$, let $\Phi_{\gamma,\theta}$ be a Parseval frame for $V_{\gamma,\theta}$.*

The sequence Ψ provides a stable sampling operator for $\mathcal{X}$ with constants α and β if and only if

i) $\dim(\mathrm{range}(G_{\Phi_{\gamma,\theta},\Psi})) = \dim(V_{\gamma,\theta})$ *for all* $\gamma, \theta \in \Gamma$ *and*

ii) $\sigma^2(G_{\Phi_{\gamma,\theta},\Psi}) \subseteq \{0\} \cup [\alpha, \beta]$ *for all* $\gamma, \theta \in \Gamma$.

For examples and existence of sequences Ψ which verify the conditions of injectivity or stability in a union of finite-dimensional subspaces, we refer the reader to [BD09] and [LD08].

3.5 Sampling in a union of finitely generated shift-invariant spaces

In this section we will consider the case of the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^d)$ and finitely generated shift-invariant spaces (FSISs). That is, we will study sampling in a union of FSISs.

3.5.1 Sampling from a union of FSISs

Our aim is to study the sampling problem for the case in which the signal belongs to the set,

$$\mathcal{X} := \bigcup_{\gamma\in\Gamma} V_\gamma, \tag{3.5}$$

where V_γ are FSISs of $L^2(\mathbb{R}^d)$.

In this setting, since our subspaces are shift-invariant, it is natural and also convenient that the sampling set will be the set of shifts from a fixed collection of functions in $L^2(\mathbb{R}^d)$, that is, the sampling operator will be given by a sequence of integer translates of certain functions.

Given $\Psi := \{\psi_i\}_{i\in I}$ such that $E(\Psi)$ is a Bessel sequence in $L^2(\mathbb{R}^d)$, we define the sampling operator associated to $E(\Psi)$ as

$$A : L^2(\mathbb{R}^d) \to \ell^2(\mathbb{Z}^d \times I), \quad Af = \{\langle f, t_k\psi_i\rangle\}_{i\in I, k\in\mathbb{Z}^d}. \tag{3.6}$$

As we showed in Section 3.3 the conditions on the sampling operator to be one-to-one and stable in a union of subspaces can be established in terms of one-to-one and stability conditions on the sum of every two of the subspaces from the union.

However the condition that we have for the sampling operator to be one-to-one on a subspace, requires that the subspace is closed (Proposition 3.3.2).

Since the sum of two FSISs is not necessarily a closed subspace, the conditions should be imposed on the closure of the sum.

Conditions that guarantee that the sum of two FSISs is closed are described in Chapter 4.

In what follows we will consider, for each $\gamma, \theta \in \Gamma$, the subspaces,

$$\overline{V}_{\gamma,\theta} := \overline{V_\gamma + V_\theta}. \tag{3.7}$$

The following proposition states that the closure of the sum of two SISs is a SIS generated by the union of the generators of the two spaces. Its proof is straightforward.

Proposition 3.5.1. *Let Φ and Φ' be sets in $L^2(\mathbb{R}^d)$, then*

$$\overline{V(\Phi) + V(\Phi')} = V(\Phi \cup \Phi').$$

In particular, if V and V' are FSISs, then $\overline{V+V'}$ is an FSIS and

$$\operatorname{len}(\overline{V+V'}) \leq \operatorname{len}(V) + \operatorname{len}(V').$$

Now, as a consequence of Proposition 3.5.1, for each $\gamma, \theta \in \Gamma$, $\overline{V}_{\gamma,\theta}$ is an FSIS. Then, by Theorem 1.4.4, we can choose, for each $\gamma, \theta \in \Gamma$, a *finite* set

$$\Phi_{\gamma,\theta} = \{\varphi_j^{\gamma,\theta}\}_{j=1}^{m_{\gamma,\theta}}$$

of $L^2(\mathbb{R}^d)$ functions such that,

$$\overline{V}_{\gamma,\theta} = V(\Phi_{\gamma,\theta}),$$

and $E(\Phi_{\gamma,\theta})$ forms a Parseval frame for $\overline{V}_{\gamma,\theta}$.

3.5.2 The one-to-one condition

We now study the conditions that the sampling set must satisfy in order for the operator A defined by (3.6) to be one-to-one on $\mathcal{X}$.

Given a shift-invariant space V, the orthogonal projection onto V, denoted by P_V, commutes with integer translates (see Proposition 1.4.3). Then, part i) of Proposition 3.3.2 can be rewritten as,

Proposition 3.5.2. *Given a shift-invariant space V, $\Psi = \{\psi_i\}_{i\in I}$ such that $E(\Psi)$ is a Bessel sequence in $L^2(\mathbb{R}^d)$ and A the sampling operator associated to $E(\Psi)$. Then the following are equivalent.*

i) The sampling operator A is one-to-one on V.

ii) $E(P_V\Psi) = \{t_k P_V \psi_i\}_{i\in I, k\in\mathbb{Z}^d}$ is complete in V, that is $V = \overline{\operatorname{span}}\, E(P_V\Psi)$.

Since $E(\Psi)$ is a Bessel sequence in $L^2(\mathbb{R}^d)$, by Theorem 1.5.8 we have that $\{\tau\psi_i(\omega)\}_{i\in I}$ is a Bessel sequence in $\ell^2(\mathbb{Z}^d)$ for a.e $\omega \in [0,1)^d$, so we can define (up to a set of measure zero), for $\omega \in [0,1)^d$, the sampling operator related to the fibers:

$$\mathcal{A}(\omega) : \ell^2(\mathbb{Z}^d) \to \ell^2(I),$$

with

$$\mathcal{A}(\omega)(c) = \{\langle c, \tau\psi_i(\omega)\rangle\}_{i\in I}. \tag{3.8}$$

That is, for a fixed $\omega \in [0,1)^d$, we consider the problem of sampling from a union of subspaces in a different setting. The Hilbert space is $\ell^2(\mathbb{Z}^d)$, the sequences of the sampling set are $\{\tau\psi_i(\omega)\}_{i\in I}$, and the subspaces in the union are $J_{V_\gamma}(\omega)$, $\gamma \in \Gamma$.

Since the subspaces $\overline{V}_{\gamma,\theta}$ are FSISs, the fiber spaces $J_{\overline{V}_{\gamma\theta}}(\omega)$ are finite-dimensional. So, the results of Section 3.4 can be applied, and conditions on the fibers can be obtained in order for the operator $\mathcal{A}(\omega)$ to be one-to-one.

We are now going to show that given a finitely generated shift-invariant space V, the operator A is one-to-one on V if and only if for almost every $\omega \in [0,1)^d$, the operator $\mathcal{A}(\omega)$ is one-to-one on the corresponding fiber spaces $J_V(\omega)$ associated to V. Once this is accomplished, we can apply the known conditions for the operator $\mathcal{A}(\omega)$.

Given $\{t_k\varphi_j\}_{j=1,k\in\mathbb{Z}^d}^m$ a Bessel sequence in $L^2(\mathbb{R}^d)$, we have the synthesis operator related to the fibers, that is

$$\mathcal{B}_\Phi^*(\omega) : \mathbb{C}^m \to \ell^2(\mathbb{Z}^d), \quad \mathcal{B}_\Phi^*(\omega)(c_1,\ldots,c_m) = \sum_{j=1}^m c_j\tau\varphi_j(\omega). \tag{3.9}$$

Note that $\mathcal{B}_\Phi^*(\omega)$ is the synthesis operator associated to the set $\tau\Phi(\omega)$, that is $\mathcal{B}_\Phi^*(\omega) = B_{\tau\Phi(\omega)}^*$.

And we will have the cross-correlation operator associated to the fibers

$$\mathcal{G}_{\Phi,\Psi}(\omega) : \mathbb{C}^m \to \ell^2(I), \quad \mathcal{G}_{\Phi,\Psi}(\omega) := \mathcal{A}(\omega)\mathcal{B}_\Phi^*(\omega),$$

$$(\mathcal{G}_{\Phi,\Psi}(\omega))_{i,j} = \langle \tau\varphi_j(\omega), \tau\psi_i(\omega)\rangle \quad \forall\, 1 \le j \le m, i \in I. \tag{3.10}$$

Again we should remark that $\mathcal{G}_{\Phi,\Psi}(\omega)$ is the cross-correlation operator associated to $\tau\Phi(\omega)$ and $\tau\Psi(\omega)$, that is $\mathcal{G}_{\Phi,\Psi}(\omega) = G_{\tau\Phi(\omega),\tau\Psi(\omega)}$.

Theorem 3.5.3. *Let $\Psi = \{\psi_i\}_{i\in I}$ be such that $E(\Psi)$ is a Bessel sequence in $L^2(\mathbb{R}^d)$, V an FSIS generated by a finite set Φ, and A the sampling operator associated to $E(\Psi)$, then the following are equivalent:*

i) Ψ *provides a one-to-one sampling operator for V.*

ii) $\ker(\mathcal{G}_{\Phi,\Psi}(\omega)) = \ker(\mathcal{B}_\Phi^*(\omega))$ *for a.e.* $\omega \in [0,1)^d$.

iii) $\dim(\mathrm{range}(\mathcal{G}_{\Phi,\Psi}(\omega))) = \dim_V(\omega)$ *for a.e.* $\omega \in [0,1)^d$.

For the proof of Theorem 3.5.3 we need the following.

Lemma 3.5.4. *Let V be an FSIS, $\Psi = \{\psi_i\}_{i\in I}$ such that $E(\Psi)$ is a Bessel sequence in $L^2(\mathbb{R}^d)$, and A the sampling operator associated to $E(\Psi)$. Then A is one-to-one on V if and only if $\mathcal{A}(\omega)$ is one-to-one on $J_V(\omega)$ for a.e. $\omega \in [0,1)^d$.*

Proof. Since V is a SIS, by Proposition 3.5.2, A is one-to-one on V if and only if

$$V = \overline{\operatorname{span}}\, E(P_V\Psi). \tag{3.11}$$

By Proposition 1.5.3, equation (3.11) is equivalent to

$$J_V(\omega) = \overline{\operatorname{span}}\{\tau(P_V\psi_i)(\omega) : i \in I\} \quad \text{for a.e. } \omega \in [0,1)^d. \tag{3.12}$$

So, we have proved that A is one-to-one on V if and only if (3.12) holds.

On the other side, given $\omega \in [0,1)^d$, and using Proposition 3.3.2 for the sampling operator $\mathcal{A}(\omega)$ and the space $\mathcal{H} = \ell^2(\mathbb{Z}^d)$, we have that $\mathcal{A}(\omega)$ is one-to-one on $J_V(\omega)$ if and only if

$$J_V(\omega) = \overline{\text{span}}\{P_{J_V(\omega)}(\tau\psi_i(\omega)) : i \in I\}.$$

Then, using Proposition 1.5.4, we conclude that (3.12) holds if and only if $\mathcal{A}(\omega)$ is one-to-one on $J_V(\omega)$, for a.e. $\omega \in [0,1)^d$, which completes the proof of the lemma.

□

Proof of Theorem 3.5.3. Since Φ is a set of generators for V, we have that for a.e. $\omega \in [0,1)^d$, $\tau\Phi(\omega)$ is a set of generators for $J_V(\omega)$.

Now, for a.e. $\omega \in [0,1)^d$ we can apply Theorem 3.4.1 for the sampling operator $\mathcal{A}(\omega)$ and the finite-dimentional subspace $J_V(\omega)$ to obtain the equivalence of the following propositions:

a) $\mathcal{A}(\omega)$ is one-to-one on $J_V(\omega)$.

b) $\ker(\mathcal{G}_{\Phi,\Psi}(\omega)) = \ker(\mathcal{B}^*_\Phi(\omega))$.

c) $\dim(\text{range}(\mathcal{G}_{\Phi,\Psi}(\omega))) = \dim(J_V(\omega)) = \dim_V(\omega)$.

From here the proof follows using Lemma 3.5.4.

□

Note that with the previous theorem we have conditions for A to be one-to-one on $\overline{V}_{\gamma,\theta}$, and since

$$V_{\gamma,\theta} = V_\gamma + V_\theta \subseteq \overline{V}_{\gamma,\theta},$$

we obtain the following corollary.

Corollary 3.5.5. *Let $E(\Psi)$ be a Bessel sequence in $L^2(\mathbb{R}^d)$ for some set of functions Ψ. For every $\gamma, \theta \in \Gamma$, let $\Phi_{\gamma,\theta}$ be a finite set of generators for $\overline{V}_{\gamma,\theta}$. If for each $\gamma, \theta \in \Gamma$,*

$$\dim(\text{range}(\mathcal{G}_{\Phi_{\gamma,\theta},\Psi}(\omega))) = \dim_{\overline{V}_{\gamma,\theta}}(\omega) \quad \textit{for a.e. } \omega \in [0,1)^d,$$

then A is one-to-one on $\mathcal{X}$.

Remark 3.5.6. It is important to note that the injectivity of A on $V_{\gamma,\theta}$ does not imply the injectivity on $\overline{V}_{\gamma,\theta}$, thus, we have only obtained sufficient conditions for A to be one-to-one. This is not a problem in general, because as we will see in the next section, stability implies injectivity in the case of the sampling operator and stability is a common and needed assumption in most sampling applications.

3.5.3 The stability condition

As a consequence of Proposition 3.3.6, we will obtain necessary and sufficient conditions for the stability of A.

As in the previous subsection, using that the orthogonal projection onto a SIS commutes with integer translates, we have the following version of Proposition 3.3.2.

Proposition 3.5.7. *Given V a SIS of $L^2(\mathbb{R}^d)$, $\Psi = \{\psi_i\}_{i\in I}$ such that $E(\Psi)$ is a Bessel sequence in $L^2(\mathbb{R}^d)$ and A the sampling operator associated to $E(\Psi)$. Then the following are equivalent:*

i) The sampling operator A is stable for V with constants α and β.

ii) $E(P_V\Psi)$ is a frame for V with constants α and β.

Now we are able to state the stability theorem. We will use the operator related to the fibers, defined by (3.8), (3.9) and (3.10).

Theorem 3.5.8. *Let $\Psi = \{\psi_i\}_{i\in I}$ be such that $E(\Psi)$ is a Bessel sequence for $L^2(\mathbb{R}^d)$ and A the sampling operator associated to $E(\Psi)$. Let V be an FSIS, and Φ a finite set of functions such that $E(\Phi)$ forms a Parseval frame for V.*

Then $E(\Psi)$ provides a stable sampling operator for V if and only if

i) $\dim(\mathrm{range}(\mathcal{G}_{\Phi,\Psi}(\omega))) = \dim_V(\omega)$ *for a.e. $\omega \in [0,1)^d$ and*

ii) There exist constants $0 < \alpha \leq \beta < \infty$ such that

$$\sigma^2(\mathcal{G}_{\Phi,\Psi}(\omega)) \subseteq \{0\} \cup [\alpha,\beta] \quad \textit{for a.e. } \omega \in [0,1)^d.$$

Proof. Φ is a Parseval frame for V, so, by Theorem 1.5.9, we have that for a.e. $\omega \in [0,1)^d$, $\tau\Phi(\omega)$ is a Parseval frame for $J_V(\omega)$. Since $J_V(\omega)$ is a finite-dimensional space of $\ell^2(\mathbb{Z}^d)$, Theorem 3.4.5 holds for $\mathcal{A}(\omega)$.

So, we only have to prove that A is stable for V with constants α and β if and only if $\mathcal{A}(\omega)$ is stable for $J_V(\omega)$ with constants α and β.

By Proposition 3.5.7, the stability of A in V is equivalent to $E(P_V\Psi)$ being a frame for V with constants α and β. By Theorem 1.5.9, this is equivalent to

$$\{\tau(P_V\psi_i)(\omega)\}_{i\in I}$$

being a frame for $J_V(\omega)$ with constants α and β for a.e. $\omega \in [0,1)^d$.

On the other hand, given $\omega \in [0,1)^d$, the operator $\mathcal{A}(\omega)$ is stable for $J_V(\omega)$, if and only if

$$\{P_{J_V(\omega)}(\tau\psi_i(\omega))\}_{i\in I}$$

is a frame for $J_V(\omega)$ with constants α and β.

The proof can be finished now using first Proposition 1.5.4, i.e.

$$\tau(P_V\psi_i)(\omega) = P_{J_V(\omega)}(\tau\psi_i(\omega)) \quad \text{for a.e. } \omega \in [0,1)^d,$$

and then Theorem 3.4.5. □

Now we apply Theorem 3.5.8 and Proposition 3.3.6 to obtain the following.

Theorem 3.5.9. *Let $\Psi = \{\psi_i\}_{i\in I}$ such that $E(\Psi)$ is a Bessel sequence for $L^2(\mathbb{R}^d)$, and for every $\gamma,\theta \in \Gamma$ let $\Phi_{\gamma,\theta}$ be a Parseval frame for $\overline{V}_{\gamma,\theta}$. Then $E(\Psi)$ provides a stable sampling operator for $\mathcal{X}$ if and only if*

i) $\dim(\text{range}(\mathcal{G}_{\Phi_{\gamma,\theta},\Psi}(\omega))) = \dim_{\overline{V}_{\gamma,\theta}}(\omega)$ *for a.e.* $\omega \in [0,1)^d$, $\forall\, \gamma,\theta \in \Gamma$ *and*

ii) There exist constants $0 < \alpha \leq \beta < \infty$ *such that*

$$\sigma^2(\mathcal{G}_{\Phi_{\gamma,\theta},\Psi}(\omega)) \subseteq \{0\} \cup [\alpha,\beta] \quad \textit{for a.e. } \omega \in [0,1)^d,\ \forall\, \gamma,\theta \in \Gamma.$$

Finally, as in [LD08], we obtain a lower bound for the amount of samples. In contrast to the previous section, we only find bounds for stable operators. We can not say anything about one-to-one operators since we only obtained sufficient conditions for the injectivity.

Proposition 3.5.10. *If the operator A is stable for $\mathcal{X}$ and I is finite, then*

$$\#I \geq \sup_{\gamma,\theta\in\Gamma}(\text{len}(\overline{V}_{\gamma,\theta})).$$

Proof. Since I is finite, it holds that $\text{range}(\mathcal{G}_{\Phi_{\gamma,\theta},\Psi}(\omega)) \subseteq \mathbb{C}^{\#I}$ for a.e. $\omega \in [0,1)^d$. Hence, by Theorem 3.5.9, we have that

$$\dim_{\overline{V}_{\gamma,\theta}}(\omega) = \dim(\text{range}(\mathcal{G}_{\Phi_{\gamma,\theta},\Psi}(\omega))) \leq \#I \quad \text{for a.e. } \omega \in [0,1)^d,\ \forall\, \gamma,\theta \in \Gamma.$$

This shows that, given $\gamma,\theta \in \Gamma$,

$$\text{ess-sup}\,\{\dim_{\overline{V}_{\gamma,\theta}}(\omega) : \omega \in [0,1)^d\} \leq \#I.$$

The proof of the proposition follows using Theorem 1.5.7.

□

We would like to note that based in our results, it is possible to state conditions for the injectivity and stability for the sampling operator in a union of SISs which are not necessarily finitely-generated. For this, condition iii) of Theorem 3.5.3 should be replaced by condition ii).

4

Closedness of the sum of two shift-invariant spaces

4.1 Introduction

In the previous chapter we obtained necessary and sufficient conditions for the stability of the sampling operator A in a union of arbitrary FSISs. We showed that, without the assumption of the closedness of the sum of every two of the FSISs in the union, we could only obtain sufficient conditions for the injectivity of A. An interesting problem that arises as a consequence of this restriction is under which conditions the sum of two SISs is a closed subspace.

For two closed subspaces U and V of an arbitrary Hilbert space $\mathcal{H}$, the conditions on the closedness of the sum of these two spaces is given in terms of the angle between the subspaces. In what follows we will define the notion of Dixmier and Friedrichs angle between subspaces. We refer the reader to [Deu95] for details and proofs.

Throughout this chapter, we will use the symbol $P_U\big|_V$ to denote the restriction of the orthogonal projection P_U to the subspace V.
The orthogonal complement of $U \cap V$ in U will be denoted by

$$U \ominus V := U \cap (U \cap V)^{\perp}.$$

Definition 4.1.1. Let U and V be closed subspaces of $\mathcal{H}$.

a) The *minimal angle* between U and V (or *Dixmier angle*) is the angle in $[0, \frac{\pi}{2}]$ whose cosine is

$$\mathbf{c}_0[U, V] := \sup\{|\langle u, v\rangle| \,:\, u \in U, v \in V, \|u\| \leq 1, \|v\| \leq 1\}.$$

b) The *angle* between U and V (or *Friedrichs angle*) is the angle in $[0, \frac{\pi}{2}]$ whose cosine is

$$\mathbf{c}[U, V] := \sup\{|\langle u, v\rangle| \,:\, u \in U \ominus V, v \in V \ominus U \text{ and } \|u\| \leq 1, \|v\| \leq 1\}.$$

We have the following results concerning both notions of angles between subspaces.

Proposition 4.1.2. *Let U and V be closed subspaces of $\mathcal{H}$.*

i) $\boldsymbol{c}_0[U,V] = \|P_U\big|_V\|_{\text{op}}$.

ii) $\boldsymbol{c}[U,V] = \boldsymbol{c}_0[U \ominus V, V \ominus U]$.

As we have stated before, the Friedrichs angle is closely related with the closedness of the sum of two closed subspaces.

Proposition 4.1.3. *Let U and V be closed subspaces of $\mathcal{H}$. Then $U + V$ is closed if and only if $\boldsymbol{c}[U,V] < 1$.*

In [KKL06] Kim et al. presented a formula for the Dixmier angle between two closed subspaces U, V of a Hilbert space $\mathcal{H}$. This formula is given in terms of the operator norm of an operator formed by the composition of the Gramians and the cross-correlation operator of two sequences X and Y which are frames for U and V respectively. They then use this formula to obtain necessary and sufficient conditions for the closedness of the sum of two SISs in $L^2(\mathbb{R}^d)$.

Following the ideas from [KKL06], in this chapter we will first give a formula for the calculation of the Friedrichs angle between two closed subspaces U, V of a Hilbert space $\mathcal{H}$. Then, we will use it to obtain necessary and sufficient conditions for the closedness of the sum of two SISs in $L^2(\mathbb{R}^d)$. The advantage of using the Friedrichs angle between subspaces instead of the Dixmier angle is that the conditions for the closedness of the sum of two subspaces are computationally simpler than the ones from [KKL06].

Using these results, we will show that it is possible to determine families of subspaces on which the conditions for injectivity of the sampling operator in the union of subspaces are necessary and sufficient.

This chapter is organized as follows. In Section 4.2 we state some preliminary results that will be used throughout. In Section 4.3 we use the notion of Friedrichs angle between subspaces to obtain necessary and sufficient conditions for the closedness of the sum of two closed subspaces of a Hilbert space. We also obtain a formula for the calculation of the Friedrichs angle between two closed subspaces. Finally, in Section 4.4 we provide an expression for the Friedrichs angle between two SISs. Using this, we give necessary and sufficient conditions for the sum of two SISs to be closed.

4.2 Preliminary results

In this section we will introduce the pseudo-inverse of an operator (see [Chr03] for more details).

Definition 4.2.1. Let $\mathcal{H}$ and $\mathcal{K}$ be separable Hilbert spaces, and $T : \mathcal{H} \to \mathcal{K}$ a bounded linear operator with closed range.

We denote by $T^\dagger$, the *pseudo-inverse* of T (or Moore-Penrose inverse) which is defined as follows. Let $R(T)$ be the closed range of T and $\widetilde{T} : \ker(T)^\perp \to R(T)$ the restriction of T to $\ker(T)^\perp$. Since T is injective on $\ker(T)^\perp$, it follows that $\widetilde{T}$ is bijective and has a bounded inverse $\widetilde{T}^{-1} : R(T) \to \ker(T)^\perp$.

The pseudo-inverse of T is defined as the unique extension $T^\dagger$ of $\widetilde{T}^{-1}$ to a bounded operator on $\mathcal{K}$ with the property $\ker(T^\dagger) = R(T)^\perp$.

The pseudo-inverse satisfies the following properties.

Proposition 4.2.2. *Let $\mathcal{H}$ and $\mathcal{K}$ be separable Hilbert spaces, and $T : \mathcal{H} \to \mathcal{K}$ a bounded linear operator with closed range. If $T^\dagger$ is the pseudo-inverse of T, then*

i) $TT^\dagger = P_{range(T)}$.

ii) $(T^\dagger)^* = (T^*)^\dagger$.

iii) $(T^*T)^\dagger = T^\dagger(T^*)^\dagger$.

iv) If T is a positive semi-definite operator, then $T^\dagger$ is also positive semi-definite.

We will need in this chapter the notion of shift-preserving operators and range operator (see [Bow00] for more details).

Definition 4.2.3. Let V be a SIS. A bounded linear operator $T : V \to L^2(\mathbb{R}^d)$ is *shift-preserving* if $Tt_k = t_kT$ for all $k \in \mathbb{Z}^d$, where t_k is the translation by k.

Definition 4.2.4. Assume V is a SIS of $L^2(\mathbb{R}^d)$ with range function J_V. A *range operator* on J_V is a mapping

$$Q : [0,1)^d \to \{\text{bounded operators defined on closed subspaces of } \ell^2(\mathbb{Z}^d)\},$$

so that the domain of $Q(\omega)$ equals $J_V(\omega)$ for a.e. $\omega \in [0,1)^d$.

Q is measurable if $\omega \mapsto Q(\omega)P_{J_V(\omega)}$ is weakly operator measurable, i.e. $\omega \mapsto \langle Q(\omega)P_{J_V(\omega)}a, b\rangle$ is a measurable scalar function for each $a, b \in \ell^2(\mathbb{Z}^d)$.

The following theorem states that there is a correspondence between shift-preserving operators and range operators.

Theorem 4.2.5. *Assume V is a SIS of $L^2(\mathbb{R}^d)$ and J_V is its range function. For every shift preserving operator $T : V \to L^2(\mathbb{R}^d)$ there exists a measurable range operator Q on J_V such that*

$$\tau(Tf)(\omega) = Q(\omega)(\tau f(\omega)) \quad \text{for a.e. } \omega \in [0,1)^d,\ f \in V.$$

The correspondence between T and Q is one-to-one.

Moreover, we have

$$\|T\|_{\text{op}} = \text{ess-sup}\{\|Q(\omega)\|_{\text{op}} : \omega \in [0,1)^d\}.$$

4.3 A formula for the Friedrichs angle

Recall that from Proposition 4.1.3 we have that the sum of two closed subspaces U, V of a Hilbert space $\mathcal{H}$ is closed if and only if the Friedrichs angle satisfies that $\mathbf{c}[U, V] < 1$.

In this section we would like to obtain an easier way of calculating the Friedrichs angle between subspaces. This is achieved in the following theorem which expresses this angle in terms of the operator norm of certain operators associated to frames.

The proof of the theorem was given in [KKL06, Theorem 2.1], but we will include it here for the sake of completeness.

Theorem 4.3.1. *Let U and V be closed subspaces of $\mathcal{H}$. Suppose that X and X' are countable subsets of $\mathcal{H}$ which are frames for $U \ominus V$ and $V \ominus U$ respectively. Then,*

$$\mathbf{c}[U, V] = \|(G_{X'}^{\dagger})^{\frac{1}{2}} G_{X,X'} (G_X^{\dagger})^{\frac{1}{2}}\|_{\mathrm{op}},$$

where G_X and $G_{X'}$ are the Gramian operators and $G_{X,X'}$ is the cross-correlation operator.

Proof. Using part ii) of Proposition 4.1.2 it suffices to show that for U and V closed subspaces of $\mathcal{H}$ it holds that

$$\mathbf{c}_0[U, V] = \|(G_{X'}^{\dagger})^{\frac{1}{2}} G_{X,X'} (G_X^{\dagger})^{\frac{1}{2}}\|_{\mathrm{op}},$$

where X and X' are countable subsets of $\mathcal{H}$ which are frames for U and V respectively.

From Proposition 1.2.10 G_X and $G_{X'}$ have closed ranges, thus, their pseudo-inverses are well-defined.

Let $P := P_V\big|_U$. From Proposition 4.2.2 we have that

$$P = P_V P_U = P_V^* P_U = (B_{X'}^* B_{X'}^{*\dagger})^* B_X^* B_X^{*\dagger} = B_{X'}^{\dagger} B_{X'} B_X^* B_X^{*\dagger} = B_{X'}^{\dagger} G_{X,X'} B_X^{*\dagger}.$$

Then, using part i) of Proposition 4.1.2 and Proposition 4.2.2, we obtain

$$\begin{aligned}
\mathbf{c}_0[U, V]^2 &= \mathbf{c}_0[V, U]^2 = \|P\|_{\mathrm{op}}^2 = \|PP^*\|_{\mathrm{op}} = \|B_{X'}^{\dagger} G_{X,X'} B_X^{*\dagger} B_X^{\dagger} G_{X,X'}^* B_{X'}^{*\dagger}\|_{\mathrm{op}} \\
&= \|B_{X'}^{\dagger} G_{X,X'} (B_X B_X^*)^{\dagger} G_{X,X'}^* B_{X'}^{*\dagger}\|_{\mathrm{op}} = \|B_{X'}^{\dagger} G_{X,X'} G_X^{\dagger} G_{X,X'}^* B_{X'}^{*\dagger}\|_{\mathrm{op}} \\
&= \|B_{X'}^{\dagger} G_{X,X'} (G_X^{\dagger})^{1/2} (G_X^{\dagger})^{1/2} G_{X,X'}^* B_{X'}^{*\dagger}\|_{\mathrm{op}} = \|B_{X'}^{\dagger} G_{X,X'} (G_X^{\dagger})^{1/2}\|_{\mathrm{op}}^2 \\
&= \|(G_X^{\dagger})^{1/2} G_{X,X'}^* B_{X'}^{*\dagger} B_{X'}^{\dagger} G_{X,X'} (G_X^{\dagger})^{1/2}\|_{\mathrm{op}} \\
&= \|(G_X^{\dagger})^{1/2} G_{X,X'}^* G_{X'}^{\dagger} G_{X,X'} (G_X^{\dagger})^{1/2}\|_{\mathrm{op}} \\
&= \|(G_X^{\dagger})^{1/2} G_{X,X'}^* (G_{X'}^{\dagger})^{1/2} (G_{X'}^{\dagger})^{1/2} G_{X,X'} (G_X^{\dagger})^{1/2}\|_{\mathrm{op}} \\
&= \|(G_{X'}^{\dagger})^{1/2} G_{X,X'} (G_X^{\dagger})^{1/2}\|_{\mathrm{op}}^2,
\end{aligned}$$

where we have used that $\|TT^*\|_{\mathrm{op}} = \|T^*T\|_{\mathrm{op}} = \|T\|_{\mathrm{op}}^2$ for a bounded operator T.

□

4.4 Closedness of the sum of two shift-invariant subspaces

As it was stated in Proposition 4.1.3, the closedness of the sum of two subspaces depends on the Friedrichs angle between them. In this section, we provide an expression for the Friedrichs angle between two SISs in terms of the Gramians of the generators. In [KKL06] Kim et al found a similar expression for the Dixmier angle between two SISs.

The main theorem of this part gives necessary and sufficient conditions for the sum of two SISs to be closed. We first state the theorem and then we apply this result to obtain a more general version of Corollary 3.5.5 from Chapter 3. The proof of the theorem will be given at the end of the section.

Theorem 4.4.1. *Let U and V be SISs of $L^2(\mathbb{R}^d)$. Suppose that Φ, Φ' are sets of functions in $L^2(\mathbb{R}^d)$ such that for a.e. $\omega \in [0,1)^d$, $\tau\Phi(\omega)$ and $\tau\Phi'(\omega)$ are frames for $J_{U\ominus V}(\omega)$ and $J_{V\ominus U}(\omega)$ respectively. Then, $U+V$ is closed if and only if*

$$\boldsymbol{c}[U,V] = \text{ess-sup}\,\{\|(\mathcal{G}_{\Phi'}(\omega)^{\dagger})^{\frac{1}{2}}\mathcal{G}_{\Phi,\Phi'}(\omega)(\mathcal{G}_{\Phi}(\omega)^{\dagger})^{\frac{1}{2}}\|_{\text{op}} \,:\, \omega \in [0,1)^d\} < 1. \tag{4.1}$$

Note that, if $V = V(\Phi)$ is an FSIS, we have that $\tau\Phi(\omega)$ is a frame for $J_V(\omega)$ for a.e. $\omega \in [0,1)^d$, even though $E(\Phi)$ is not a frame for V (see Remark 1.5.10). Thus, if U and V are FSISs, condition (4.1) can be checked on any set of generators of the subspaces $U \ominus V$ and $V \ominus U$. At the end of the section we give an example in which we compute the Friedrichs angle between two FSISs.

When the set of functions Φ is finite and $\#(\Phi) = m$, for a fixed $\omega \in [0,1)^d$ the Gramian matrix $\mathcal{G}_\Phi(\omega) \in \mathbb{C}^{m\times m}$ is Hermitian positive semidefinite. Thus, we have that

$$\mathcal{G}_\Phi(\omega) = U(\omega)D(\omega)U^*(\omega),$$

where $U(\omega)$ is a unitary matrix and $D(\omega) = \text{diag}\{\lambda_1(\omega), \dots, \lambda_m(\omega)\}$ with $\lambda_1(\omega) \geq \cdots \geq \lambda_m(\omega) \geq 0$ the eigenvalues of the Gramian matrix.

In [RS95] it was proved that the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_m$ and the entries of the matrix U are measurable functions.

In this case we have that the pseudo-inverse of the Gramian matrix and the square root of the pseudo-inverse are

$$\mathcal{G}_\Phi(\omega)^{\dagger} = U(\omega)D(\omega)^{\dagger}U^*(\omega) \quad \text{and} \quad (\mathcal{G}_\Phi(\omega)^{\dagger})^{\frac{1}{2}} = U(\omega)(D(\omega)^{\dagger})^{\frac{1}{2}}U^*(\omega),$$

where $D(\omega)^{\dagger} = \text{diag}\{\lambda_1(\omega)^{-1}, \dots, \lambda_{d(\omega)}(\omega)^{-1}, 0, \dots, 0\}$ and $d(\omega) = \text{rank}[\mathcal{G}_\Phi(\omega)]$.

In the next theorem we show that imposing certain restrictions on the angle between the subspaces, we obtain necessary and sufficient conditions for the injectivity of the sampling operator in a union of subspaces. This gives a more complete version of Corollary 3.5.5 from Chapter 3.

Theorem 4.4.2. *Let $\Psi = \{\psi_i\}_{i\in I}$ be such that $E(\Psi)$ is a Bessel sequence in $L^2(\mathbb{R}^d)$ and let $\{V_\gamma\}_{\gamma\in\Gamma}$ be FSISs of $L^2(\mathbb{R}^d)$. Suppose condition (4.1) is satisfied for every pair of subspaces V_γ, V_θ with $\gamma, \theta \in \Gamma$.*

If $\Phi_{\gamma,\theta}$ is a finite set of generators for $V_{\gamma,\theta} = V_\gamma + V_\theta$, the following are equivalent:

i) Ψ *provides a one-to-one sampling operator for $\mathcal{X}$.*

ii) $\dim(\operatorname{range}(\mathcal{G}_{\Phi_{\gamma,\theta},\Psi}(\omega))) = \dim_{V_{\gamma,\theta}}(\omega)$ *for a.e. $\omega \in [0,1)^d, \forall\, \gamma, \theta \in \Gamma$.*

Proof. Since condition (4.1) is satisfied for every $\gamma, \theta \in \Gamma$, it holds that the subspaces $V_{\gamma,\theta} = V_\gamma + V_\theta$ are FSISs. The proof of the theorem follows applying Theorem 3.5.3 to these subspaces.

□

In what follows we will give some lemmas which will be needed for the proof of Theorem 4.4.1. The results in these lemmas are interesting by themselves.

The first lemma uses the notion of range function introduced in Definition 1.5.2.

Lemma 4.4.3. *Given U and V SISs of $L^2(\mathbb{R}^d)$. Then the range function*

$$R : [0,1)^d \to \{\text{closed subspaces of } \ell^2(\mathbb{Z}^d)\}, \quad R(\omega) = J_U(\omega) \cap J_V(\omega),$$

is measurable.

Proof. Recall that the measurability of R is equivalent to $\omega \mapsto P_{J_U(\omega)\cap J_V(\omega)}$ being measurable.

It is known (see [VN50]) that given M and N closed subspaces of a separable Hilbert space $\mathcal{H}$, for each $x \in \mathcal{H}$,

$$P_{M\cap N}(x) = \lim_{n\to+\infty} (P_M P_N)^n(x).$$

Note that if we have two measurable functions

$$Q_1, Q_2 : [0,1)^d \to \{\text{orthogonal projections in } \ell^2(\mathbb{Z}^d)\},$$

then the map $\omega \mapsto Q_1(\omega)Q_2(\omega)$ is measurable. For, let F be an arbitrary measurable function from $[0,1)^d$ into $\ell^2(\mathbb{Z}^d)$. Then

$$Q_1(\omega)Q_2(\omega)(F(\omega)) = Q_1(\omega)(Q_2(\omega)(F(\omega))).$$

By Definition 1.5.2, the measurability of $Q_2(\omega)$ implies the vector measurability of $Q_2(\omega)(F(\omega))$. Since $Q_1(\omega)$ is measurable, $Q_1(\omega)Q_2(\omega)(F(\omega))$ is measurable. What shows that $\omega \mapsto Q_1(\omega)Q_2(\omega)$ is measurable.

As a consequence, it holds that for any $n \in \mathbb{N}$ the map $\omega \mapsto (P_{J_U(\omega)}P_{J_V(\omega)})^n$ is measurable, that is, for each $a \in \ell^2(\mathbb{Z}^d)$, $\omega \mapsto (P_{J_U(\omega)}P_{J_V(\omega)})^n(a)$ is measurable. From here the proof follows using that,

$$P_{J_U(\omega)\cap J_V(\omega)}(a) = \lim_{n\to+\infty} (P_{J_U(\omega)}P_{J_V(\omega)})^n(a).$$

□

With the previous lemma we obtain the following property of the fiber spaces.

Lemma 4.4.4. *Let U and V be SISs of $L^2(\mathbb{R}^d)$. Then,*

$$J_{U\ominus V}(\omega) = J_U(\omega) \ominus J_V(\omega) \quad \textit{for a.e. } \omega \in [0,1)^d.$$

Proof. We will first prove that

$$J_{U\cap V}(\omega) = J_U(\omega) \cap J_V(\omega) \quad \text{for a.e. } \omega \in [0,1)^d. \tag{4.2}$$

Let R be the measurable range function defined in Lemma 4.4.3. Since

$$U \cap V = \{f \in L^2(\mathbb{R}^d) \,:\, \tau f(\omega) \in R(\omega) \text{ for a.e. } \omega \in [0,1)^d\},$$

it follows that R is the range function associated to the shift-invariant space $U \cap V$, thus (4.2) holds.

Using (4.2), the proof of the proposition is straightforward as

$$(J_V(\omega))^\perp = J_{V^\perp}(\omega) \quad \text{for a.e. } \omega \in [0,1)^d,$$

for any shift-invariant space V of $L^2(\mathbb{R}^d)$. □

The next lemma follows the ideas from [BG04]. It states that the angle between two shift-invariant spaces is the essential supremum of the angles between the fiber spaces.

Lemma 4.4.5. *Let U and V be SISs of $L^2(\mathbb{R}^d)$. Then,*

$$\boldsymbol{c}[U,V] = \text{ess-sup}\,\{\boldsymbol{c}[J_U(\omega), J_V(\omega)] \,:\, \omega \in [0,1)^d\}.$$

Proof. Given $f \in V$, by Proposition 1.5.4, we have for a.e. $\omega \in [0,1)^d$,

$$\tau(P_U\big|_V f)(\omega) = \tau(P_U P_V f)(\omega) = P_{J_U(\omega)} P_{J_V(\omega)}(\tau f(\omega)) = P_{J_U(\omega)}\big|_{J_V(\omega)}(\tau f(\omega)).$$

By Theorem 4.2.5 this shows that $P_{J_U(\omega)}\big|_{J_V(\omega)}$ is the range operator corresponding to the shift-preserving operator $P_U\big|_V$ in the shift-invariant space V. What implies that

$$\|P_U\big|_V\|_{\text{op}} = \text{ess-sup}\,\{\|P_{J_U(\omega)}\big|_{J_V(\omega)}\|_{\text{op}} \,:\, \omega \in [0,1)^d\}. \tag{4.3}$$

Using (4.3), Proposition 4.1.2 and Lemma 4.4.4, we obtain

$$\begin{aligned}
\mathbf{c}[U,V] &= \mathbf{c}_0[U\ominus V, V\ominus U] = \|P_{U\ominus V}\big|_{V\ominus U}\|_{\text{op}} \\
&= \text{ess-sup}\,\{\|P_{J_{U\ominus V}(\omega)}\big|_{J_{V\ominus U}(\omega)}\|_{\text{op}} \,:\, \omega \in [0,1)^d\} \\
&= \text{ess-sup}\,\{\|P_{J_U(\omega)\ominus J_V(\omega)}\big|_{J_V(\omega)\ominus J_U(\omega)}\|_{\text{op}} \,:\, \omega \in [0,1)^d\} \\
&= \text{ess-sup}\,\{\mathbf{c}_0[J_U(\omega)\ominus J_V(\omega), J_V(\omega)\ominus J_U(\omega)] \,:\, \omega \in [0,1)^d\} \\
&= \text{ess-sup}\,\{\mathbf{c}[J_U(\omega), J_V(\omega)] \,:\, \omega \in [0,1)^d\}.
\end{aligned}$$

□

With the above results, we are able to prove the main theorem of this section.

Proof of Theorem 4.4.1. By Lemma 4.4.4, it follows that $\tau\Phi(\omega)$ and $\tau\Phi'(\omega)$ are frames for $J_U(\omega) \ominus J_V(\omega)$ and $J_V(\omega) \ominus J_U(\omega)$ respectively, for a.e. $\omega \in [0,1)^d$.

Thus, using Theorem 4.3.1, we obtain

$$\mathbf{c}[J_U(\omega), J_V(\omega)] = \|(\mathcal{G}_{\Phi'}(\omega)^{\dagger})^{\frac{1}{2}}\mathcal{G}_{\Phi,\Phi'}(\omega)(\mathcal{G}_{\Phi}(\omega)^{\dagger})^{\frac{1}{2}}\|_{\text{op}} \quad \text{for a.e. } \omega \in [0,1)^d.$$

Hence, by Lemma 4.4.5,

$$\mathbf{c}[U, V] = \text{ess-sup}\,\{\|(\mathcal{G}_{\Phi'}(\omega)^{\dagger})^{\frac{1}{2}}\mathcal{G}_{\Phi,\Phi'}(\omega)(\mathcal{G}_{\Phi}(\omega)^{\dagger})^{\frac{1}{2}}\|_{\text{op}} \,:\, \omega \in [0,1)^d\}. \tag{4.4}$$

The proof of the theorem follows from (4.4) and Proposition 4.1.3. □

Next we provide an example of two shift-invariant spaces whose sum is not closed. In order to prove that, we compute the Friedrichs angle between the subspaces.

Example 4.4.6. Let $\varphi_1 \in L^2(\mathbb{R})$ be given by

$$\widehat{\varphi_1}(\omega) = \begin{cases} \cos(2\pi\omega) & \text{if } 0 \le \omega < 1 \\ \sin(2\pi\omega) & \text{if } 1 \le \omega < 2 \\ 0 & \text{otherwise,} \end{cases}$$

and $\varphi_2, \varphi_3 \in L^2(\mathbb{R})$ satisfying that $\widehat{\varphi_2}(\omega) = \chi_{[2,3)}(\omega)$ and $\widehat{\varphi_3}(\omega) = \chi_{[3,4)}(\omega)$. Define $U = V(\varphi_1, \varphi_2, \varphi_3)$.
Consider now $\varphi_0, \varphi_4 \in L^2(\mathbb{R})$, such that $\widehat{\varphi_0}(\omega) = \chi_{[0,1)}(\omega)$ and $\widehat{\varphi_4}(\omega) = \chi_{[\frac{5}{2},\frac{7}{2})}(\omega)$, set $V = V(\varphi_0, \varphi_4)$.

We will prove that $U + V$ is not closed using Theorem 4.4.1.

Let $\{e_k\}_{k\in\mathbb{Z}}$ be the standard basis for $\ell^2(\mathbb{Z})$. Then, $\tau\varphi_1(\omega) = \cos(2\pi\omega)e_0 + \sin(2\pi\omega)e_1$, $\tau\varphi_2(\omega) = e_2$, $\tau\varphi_3(\omega) = e_3$, $\tau\varphi_0(\omega) = e_0$, $\tau\varphi_4(\omega) = e_3\chi_{[0,\frac{1}{2})}(\omega) + e_2\chi_{[\frac{1}{2},1)}(\omega)$. So, we have that for a.e. $\omega \in [0,1)$,

$$J_U(\omega) \ominus J_V(\omega) = \text{span}\{\tau\varphi_1(\omega), \tau\varphi_5(\omega)\} \quad \text{and} \quad J_V(\omega) \ominus J_U(\omega) = \text{span}\{\tau\varphi_0(\omega)\},$$

where $\widehat{\varphi_5}(\omega) = \chi_{[2,\frac{5}{2})}(\omega) + \chi_{[\frac{7}{2},4)}(\omega)$. Thus, by Lemma 4.4.4, it follows that $U \ominus V = V(\varphi_1, \varphi_5)$ and $V \ominus U = V(\varphi_0)$.

Let $\Phi := \{\varphi_1, \varphi_5\}$ and $\Phi' := \{\varphi_0\}$, then

$$\mathcal{G}_{\Phi'}(\omega) = 1 \quad \mathcal{G}_{\Phi}(\omega) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathcal{G}_{\Phi,\Phi'}(\omega) = (\cos(2\pi\omega), 0).$$

Therefore

$$\begin{aligned} \mathbf{c}[U, V] &= \text{ess-sup}\,\{\|(\mathcal{G}_{\Phi'}(\omega)^{\dagger})^{\frac{1}{2}}\mathcal{G}_{\Phi,\Phi'}(\omega)(\mathcal{G}_{\Phi}(\omega)^{\dagger})^{\frac{1}{2}}\|_{\text{op}} \,:\, \omega \in [0,1)\} \\ &= \text{ess-sup}\,\{|\cos(2\pi\omega)| \,:\, \omega \in [0,1)\} = 1. \end{aligned}$$

Hence, by Theorem 4.4.1, $U + V$ is not closed.

5

Extra invariance of shift-invariant spaces

5.1 Introduction

In Chapter 2 we have studied the problem of finding an FSIS V_0 that best approximates a finite set of functions $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq L^2(\mathbb{R}^d)$. Suppose now that we want to approximate a delayed version of the data $\mathcal{F}$. That is, we would like to approximate the set $t_\alpha \mathcal{F} = \{t_\alpha f_1, \dots, t_\alpha f_m\}$ for some $\alpha \in \mathbb{R}^d$. If the optimal FSIS V_0 for $\mathcal{F}$ is invariant under the translation in α (i.e for any $f \in V_0$, $t_\alpha f \in V_0$), then we will prove in the following proposition that V_0 is also optimal for the data set $t_\alpha \mathcal{F}$.

Proposition 5.1.1. *Let $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq L^2(\mathbb{R}^d)$ and $\alpha \in \mathbb{R}^d$. Assume $\mathcal{L}_k$ is the class of FSISs of length at most k and V_0 is an optimal FSIS for $\mathcal{F}$ in the sense that*

$$\mathcal{E}(\mathcal{F}, V_0) = \inf_{V \in \mathcal{L}_k} \mathcal{E}(\mathcal{F}, V),$$

where $\mathcal{E}$ is as in Definition 2.2.1. If V_0 is invariant under the translation in α, then V_0 is an optimal FSIS for the corrupted data $t_\alpha \mathcal{F}$.

Proof. Due to the α-invariance of V_0, we have that

$$\begin{aligned} \mathcal{E}(\mathcal{F}, V_0) &= \sum_{i=1}^m \|f_i - P_{V_0} f_i\|^2 = \sum_{i=1}^m \|t_\alpha f_i - t_\alpha P_{V_0} f_i\|^2 \\ &= \sum_{i=1}^m \|t_\alpha f_i - P_{V_0} t_\alpha f_i\|^2 = \mathcal{E}(t_\alpha \mathcal{F}, V_0) \end{aligned}$$

For a given $V \in \mathcal{L}_k$, using the preceding and that $\mathcal{E}(\mathcal{F}, V) = \mathcal{E}(t_\alpha \mathcal{F}, t_\alpha V)$, we obtain

$$\begin{aligned} \mathcal{E}(t_\alpha \mathcal{F}, V) &= \mathcal{E}(\mathcal{F}, t_{-\alpha} V) \\ &\geq \mathcal{E}(\mathcal{F}, V_0) = \mathcal{E}(t_\alpha \mathcal{F}, V_0). \end{aligned}$$

Thus,

$$\mathcal{E}(t_\alpha \mathcal{F}, V_0) = \inf_{V \in \mathcal{L}_k} \mathcal{E}(t_\alpha \mathcal{F}, V).$$

□

The previous proposition shows that an FSIS which is optimal for $\mathcal{F}$ and has extra-invariance $\alpha \in \mathbb{R}^d$, is also optimal for the corrupted data $t_\alpha\mathcal{F}$. This fact motivates an important and interesting question regarding SISs which is whether they have the property to be invariant under translations other than integers.

In this chapter we will be interested in characterizing the SISs that are not only invariant under integer translations, but are also invariant under some particular set of translations of $\mathbb{R}^d$.

A limit case is when the space is invariant under translations by every $\alpha \in \mathbb{R}^d$. In this case the space is called *translation invariant*. One example of a translation invariant space is the Paley-Wiener space of functions that are bandlimited to $[-\frac{1}{2}, \frac{1}{2}]$ defined by

$$PW = \left\{ f \in L^2(\mathbb{R}) \ : \ \operatorname{supp}(\widehat{f}) \subseteq \left[-\frac{1}{2}, \frac{1}{2}\right] \right\}.$$

Recall that PW is a principal shift-invariant space generated by the function $\operatorname{sinc}(t)$. This space is translation invariant since if $f \in PW$ and $\alpha \in \mathbb{R}$, we have that $\operatorname{supp}(\widehat{t_\alpha f}) = \operatorname{supp}(e^{-2\pi i\alpha \cdot}\widehat{f}) = \operatorname{supp}(\widehat{f})$, thus $t_\alpha f \in PW$ for every $\alpha \in \mathbb{R}$.

In the same way it is easy to prove that for a measurable set $E \subseteq \mathbb{R}^d$, the spaces

$$\{f \in L^2(\mathbb{R}^d) \ : \ \operatorname{supp}(\widehat{f}) \subseteq E\}, \tag{5.1}$$

are translation invariant. As a matter of fact, Wiener's theorem (see [Hel64], [Sri64]) proves that any closed translation invariant subspaces of $L^2(\mathbb{R}^d)$ if of the form (5.1).

On the other hand, there exist SISs that are only invariant under integer translates. To see this, consider for example the principal shift-invariant space generated by the indicator function $\chi_{[0,1)}$

$$V(\chi_{[0,1)}) = \overline{\operatorname{span}}\{t_k\chi_{[0,1)} \ : \ k \in \mathbb{Z}\}.$$

It is easy to see that this space is only invariant under integer translates.

Let us now define, for a given SIS $V \subseteq L^2(\mathbb{R}^d)$, the *invariance set* associated to V as

$$M := \{x \in \mathbb{R}^d \ : \ t_x f \in V, \ \ \forall\, f \in V\}.$$

So, for the Paley-Wiener space we have that $M = \mathbb{R}$ and for the PSIS $V(\chi_{[0,1)})$, it follows that $M = \mathbb{Z}$.

One question that naturally arises is, for a given SIS V of $L^2(\mathbb{R}^d)$, how is the structure of the invariance set M.

In [ACHKM10] Aldroubi et al. showed that if V is a shift-invariant space, then its *invariance set*, is a closed additive subgroup of $\mathbb{R}$ containing $\mathbb{Z}$. As a consequence, since every additive subgroup of $\mathbb{R}$ is either discrete or dense, there are only two possibilities left for the extra invariance. That is, either V is invariant under translations by the group $(1/n)\mathbb{Z}$, for some positive integer n (and not invariant under any bigger subgroup) or it is translation invariant. They found different characterizations in terms of the Fourier transform, of when a shift invariant space is $(1/n)\mathbb{Z}$-invariant.

The problem that we solve in this chapter is if the characterizations of extra invariance that hold on the line are still valid in several variables. As in the one-dimensional case we will prove that the invariance set M associated to a SIS of $L^2(\mathbb{R}^d)$ is a closed subgroup of $\mathbb{R}^d$ (see Proposition 5.2.1). The main difference here with the one dimensional case, is that there are subgroups of $\mathbb{R}^d$ that are neither discrete nor dense. So, it is no direct that all the characterizations given in [ACHKM10] are still valid in several variables.

We will find necessary and sufficient conditions for a SIS to be invariant under a closed additive subgroup $M \subseteq \mathbb{R}^d$ containing $\mathbb{Z}^d$. In addition our results show the existence of shift-invariant spaces that are *exactly* M-invariant for every closed subgroup $M \subseteq \mathbb{R}^d$ containing $\mathbb{Z}^d$. By 'exactly M-invariant' we mean that they are not invariant under any other subgroup containing M. We apply our results to obtain estimates on the size of the support of the Fourier transform of the generators of the space.

The chapter is organized in the following way: Section 5.2 studies the structure of the invariance set. We review the structure of closed additive subgroups of $\mathbb{R}^d$ in Section 5.3. In Section 5.4 we extend some results, known for shift-invariant spaces in $\mathbb{R}^d$, to M-invariant spaces when M is a closed subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$. The necessary and sufficient conditions for the M-invariance of shift-invariant spaces are stated and proved in Section 5.5. Finally, Section 5.6 contains some applications of our results.

5.2 The structure of the invariance set

For a shift-invariant space $V \subseteq L^2(\mathbb{R}^d)$, we define the invariance set as

$$M := \{x \in \mathbb{R}^d \,:\, t_x f \in V, \;\; \forall\, f \in V\}. \tag{5.2}$$

If Φ is a set of generators for V, it is easy to check that

$$M = \{x \in \mathbb{R}^d \,:\, t_x\varphi \in V, \;\; \forall\, \varphi \in \Phi\}.$$

Our aim in this section is to study the structure of the set M.

Proposition 5.2.1. *Let V be a SIS of $L^2(\mathbb{R}^d)$ and let M be defined as in (5.2). Then M is an additive closed subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$.*

For the proof of this proposition we will need the following lemma. Recall that an additive semigroup is a non-empty set with an associative additive operation.

Lemma 5.2.2. *Let H be a closed semigroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$, then H is a group.*

Proof. Let π be the quotient map from $\mathbb{R}^d$ onto $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. Since H is a semigroup containing $\mathbb{Z}^d$, we have that $H + \mathbb{Z}^d = H$, thus

$$\pi^{-1}(\pi(H)) = \bigcup_{h\in H} h + \mathbb{Z}^d = H + \mathbb{Z}^d = H.$$

This shows that $\pi(H)$ is closed in $\mathbb{T}^d$ and therefore compact.

By [HR63, Theorem 9.16], we have that a compact semigroup of $\mathbb{T}^d$ is necessarily a group, thus $\pi(H)$ is a group and consequently H is a group.

□

Proof of Proposition 5.2.1. Since V is a SIS, $\mathbb{Z}^d \subseteq M$.

We now show that M is closed. Let $x_0 \in \mathbb{R}^d$ and $\{x_n\}_{n\in\mathbb{N}} \subseteq M$, such that $\lim_{n\to\infty} x_n = x_0$. Then

$$\lim_{n\to\infty} \|t_{x_n} f - t_{x_0} f\| = 0.$$

Therefore, $t_{x_0} f \in \overline{V}$. But V is closed, so $t_{x_0} f \in V$.

It is easy to check that M is a semigroup of $\mathbb{R}^d$, hence we conclude from Lemma 5.2.2 that M is a group.

□

Since the invariance set of a SIS is a closed subgroup of $\mathbb{R}^d$, our aim in what follows is to give some characterizations concerning closed subgroups of $\mathbb{R}^d$.

5.3 Closed subgroups of $\mathbb{R}^d$

Throughout this section we describe the additive closed subgroups of $\mathbb{R}^d$ containing $\mathbb{Z}^d$. We first study closed subgroups of $\mathbb{R}^d$ in general.

When two groups G_1 and G_2 are isomorphic we will write $G_1 \approx G_2$. Here and subsequently all the vector subspaces will be real.

5.3.1 General case

We will state in this section, some basic definitions and properties of closed subgroups of $\mathbb{R}^d$, for a detailed treatment and proofs we refer the reader to [Bou74].

Definition 5.3.1. Given M a subgroup of $\mathbb{R}^d$, the *range* of M, denoted by $\mathbf{r}(M)$, is the dimension of the subspace generated by M as a real vector space.

It is known that every closed subgroup of $\mathbb{R}^d$ is either discrete or contains a subspace of at least dimension one (see [Bou74, Proposition 3]).

Definition 5.3.2. Given M a closed subgroup of $\mathbb{R}^d$, there exists a subspace V whose dimension is the largest of the dimensions of all the subspaces contained in M. We will denote by $\mathbf{d}(M)$ the dimension of V. Note that $\mathbf{d}(M)$ can be zero and $0 \le \mathbf{d}(M) \le \mathbf{r}(M) \le d$.

The next theorem establishes that every closed subgroup of $\mathbb{R}^d$ is the direct sum of a subspace and a discrete group.

Theorem 5.3.3. *Let M be a closed subgroup of $\mathbb{R}^d$ such that $\mathbf{r}(M) = r$ and $\mathbf{d}(M) = p$. Let V be the subspace contained in M as in Definition 5.3.2. There exists a basis $\{u_1, \ldots, u_d\}$ for $\mathbb{R}^d$ such that $\{u_1, \ldots, u_r\} \subseteq M$ and $\{u_1, \ldots, u_p\}$ is a basis for V. Furthermore,*

$$M = \Big\{ \sum_{i=1}^{p} t_i u_i + \sum_{j=p+1}^{r} n_j u_j \; : \; t_i \in \mathbb{R},\, n_j \in \mathbb{Z} \Big\}.$$

Corollary 5.3.4. *If M is a closed subgroup of $\mathbb{R}^d$ such that $\mathbf{r}(M) = r$ and $\mathbf{d}(M) = p$, then*

$$M \approx \mathbb{R}^p \times \mathbb{Z}^{r-p}.$$

5.3.2 Closed subgroups of $\mathbb{R}^d$ containing $\mathbb{Z}^d$

We are interested in closed subgroups of $\mathbb{R}^d$ containing $\mathbb{Z}^d$. For their understanding, the notion of dual group is important.

Definition 5.3.5. Let M be a subgroup of $\mathbb{R}^d$. Consider the set

$$M^* := \{x \in \mathbb{R}^d \; : \; \langle x, m \rangle \in \mathbb{Z} \quad \forall\, m \in M\}.$$

Then M^* is a subgroup of $\mathbb{R}^d$ called the *dual group* of M. In particular, $(\mathbb{Z}^d)^* = \mathbb{Z}^d$.

Now we will list some properties of the dual group.

Proposition 5.3.6. *Let M, N be subgroups of $\mathbb{R}^d$.*

i) *M^* is a closed subgroup of $\mathbb{R}^d$.*

ii) *If $N \subseteq M$, then $M^* \subseteq N^*$.*

iii) *If M is closed, then $\mathbf{r}(M^*) = d - \mathbf{d}(M)$ and $\mathbf{d}(M^*) = d - \mathbf{r}(M)$.*

iv) *$(M^*)^* = \overline{M}$.*

Let H be a subgroup of $\mathbb{Z}^d$ with $\mathbf{r}(H) = q$, we will say that a set $\{v_1, \ldots, v_q\} \subseteq H$ is a *basis* for H if for every $x \in H$ there exist unique $k_1, \ldots, k_q \in \mathbb{Z}$ such that

$$x = \sum_{i=1}^{q} k_i v_i.$$

Note that $\{v_1, \ldots, v_d\} \subseteq \mathbb{Z}^d$ is a basis for $\mathbb{Z}^d$ if and only if the determinant of the matrix A which has $\{v_1, \ldots, v_d\}$ as columns is 1 or -1.

Given $B = \{v_1, \ldots, v_d\}$ a basis for $\mathbb{Z}^d$, we will call $\widetilde{B} = \{w_1, \ldots, w_d\}$ a *dual basis* for B if $\langle v_i, w_j \rangle = \delta_{i,j}$ for all $1 \le i, j \le d$.

If we denote by $\widetilde{A}$ the matrix with columns $\{w_1, \ldots, w_d\}$, the relation between B and $\widetilde{B}$ can be expressed in terms of matrices as $\widetilde{A} = (A^*)^{-1}$, where A^* is the transpose of A.

The closed subgroups M of $\mathbb{R}^d$ containing $\mathbb{Z}^d$, can be described with the help of the dual relations. Since $\mathbb{Z}^d \subseteq M$, we have that $M^* \subseteq \mathbb{Z}^d$. So, we need first the characterization of the subgroups of $\mathbb{Z}^d$. This is stated in the following theorem which is proved in [Bou81].

Theorem 5.3.7. *Let H be a subgroup of $\mathbb{Z}^d$ with $\mathbf{r}(H) = q$, then there exist a basis $\{w_1, \ldots, w_d\}$ for $\mathbb{Z}^d$ and unique integers $a_1, \ldots, a_q$ satisfying $a_{i+1} \equiv 0\,(mod.\ a_i)$ for all $1 \le i \le q-1$, such that $\{a_1 w_1, \ldots, a_q w_q\}$ is a basis for H. The integers $a_1, \ldots, a_q$ are called* invariant factors.

Remark 5.3.8. Under the assumptions of the above theorem we obtain

$$\mathbb{Z}^d / H \approx \mathbb{Z}_{a_1} \times \ldots \times \mathbb{Z}_{a_q} \times \mathbb{Z}^{d-q}.$$

We are now able to characterize the closed subgroups of $\mathbb{R}^d$ containing $\mathbb{Z}^d$. The proof of the following theorem can be found in [Bou74], but we include it here for the sake of completeness.

Theorem 5.3.9. *Let $M \subseteq \mathbb{R}^d$. The following conditions are equivalent:*

i) M is a closed subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$ and $\mathbf{d}(M) = d - q$.

ii) There exist a basis $\{v_1, \ldots, v_d\}$ for $\mathbb{Z}^d$ and integers $a_1, \ldots, a_q$ satisfying $a_{i+1} \equiv 0\,(mod.\ a_i)$ for all $1 \le i \le q-1$, such that

$$M = \Big\{ \sum_{i=1}^{q} k_i \frac{1}{a_i} v_i + \sum_{j=q+1}^{d} t_j v_j \,:\, k_i \in \mathbb{Z}, t_j \in \mathbb{R} \Big\}.$$

Furthermore, the integers q and $a_1, \ldots, a_q$ are uniquely determined by M.

Proof. Suppose i) is true. Since $\mathbb{Z}^d \subseteq M$ and $\mathbf{d}(M) = d - q$, we have that $M^* \subseteq \mathbb{Z}^d$ and $\mathbf{r}(M^*) = q$. By Theorem 5.3.7, there exist invariant factors $a_1, \ldots, a_q$ and $\{w_1, \ldots, w_d\}$ a basis for $\mathbb{Z}^d$ such that $\{a_1 w_1, \ldots, a_q w_q\}$ is a basis for M^*.

Let $\{v_1, \ldots, v_d\}$ be the dual basis for $\{w_1, \ldots, w_d\}$.

Since M is closed, it follows from item iv) of Proposition 5.3.6 that $M = (M^*)^*$. So, $m \in M$ if and only if

$$\langle m, a_j w_j \rangle \in \mathbb{Z} \quad \forall\ 1 \le j \le q. \tag{5.3}$$

As $\{v_1, \ldots, v_d\}$ is a basis, given $u \in \mathbb{R}^d$, there exist $u_i \in \mathbb{R}$ such that $u = \sum_{i=1}^d u_i v_i$. Thus, by (5.3), $u \in M$ if and only if $u_i a_i \in \mathbb{Z}$ for all $1 \le i \le q$.

We finally obtain that $u \in M$ if and only if there exist $k_i \in \mathbb{Z}$ and $u_j \in \mathbb{R}$ such that

$$u = \sum_{i=1}^{q} k_i \frac{1}{a_i} v_i + \sum_{j=q+1}^{d} u_j v_j.$$

The proof of the other implication is straightforward.

The integers q and $a_1, \dots, a_q$ are uniquely determined by M since $q = d - \mathbf{d}(M)$ and $a_1, \dots, a_q$ are the invariant factors of M^*.

□

As a consequence of the proof given above we obtain the following corollary.

Corollary 5.3.10. *Let $\mathbb{Z}^d \subseteq M \subseteq \mathbb{R}^d$ be a closed subgroup with $\mathbf{d}(M) = d - q$. If $\{v_1, \dots, v_d\}$ and $a_1, \dots, a_q$ are as in Theorem 5.3.9, then*

$$M^* = \Big\{ \sum_{i=1}^{q} n_i a_i w_i \ : \ n_i \in \mathbb{Z} \Big\},$$

where $\{w_1, \dots, w_d\}$ is the dual basis of $\{v_1, \dots, v_d\}$.

Example 5.3.11. Assume that $d = 3$. If $M = \frac{1}{2}\mathbb{Z} \times \frac{1}{3}\mathbb{Z} \times \mathbb{R}$, then $v_1 = (1,1,0)$, $v_2 = (3,2,0)$ and $v_3 = (0,0,1)$ verify the conditions of Theorem 5.3.9 with the invariant factors $a_1 = 1$ and $a_2 = 6$. On the other hand $v_1' = (1,1,0)$, $v_2' = (3,2,1)$ and $v_3' = (0,0,1)$ verify the same conditions. This shows that the basis in Theorem 5.3.9 is not unique.

Remark 5.3.12. If $\{v_1, \dots, v_d\}$ and $a_1, \dots, a_q$ are as in Theorem 5.3.9, let us define the linear transformation T as

$$T : \mathbb{R}^d \to \mathbb{R}^d, \quad T(e_i) = v_i \quad \forall\, 1 \le i \le d.$$

Then T is an invertible transformation that satisfies

$$M = T\Big(\frac{1}{a_1}\mathbb{Z} \times \cdots \times \frac{1}{a_q}\mathbb{Z} \times \mathbb{R}^{d-q}\Big).$$

If $\{w_1, \dots, w_d\}$ is the dual basis for $\{v_1, \dots, v_d\}$, the inverse of the adjoint of T is defined by

$$(T^*)^{-1} : \mathbb{R}^d \to \mathbb{R}^d, \quad (T^*)^{-1}(e_i) = w_i \quad \forall\, 1 \le i \le d.$$

By Corollary 5.3.10, it is true that

$$M^* = (T^*)^{-1}(a_1\mathbb{Z} \times \cdots \times a_q\mathbb{Z} \times \{0\}^{d-q}).$$

5.4 The structure of principal M-invariant spaces

Throughout this section M will be a closed subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$ and M^* its dual group defined as in the previous section.

Here and subsequently for $\alpha \in \mathbb{R}^d$, we will write the exponential function $e^{-2\pi i\langle\omega,\alpha\rangle}$ as $e_\alpha(\omega)$.

For a set of functions $\Phi \subseteq L^2(\mathbb{R}^d)$, we write $\widehat{\Phi} = \{\widehat{f} : f \in \Phi\}$.

Definition 5.4.1. We will say that a closed subspace V of $L^2(\mathbb{R}^d)$ is *M-invariant* if $t_m f \in V$ for all $m \in M$ and $f \in V$.

Given $\Phi \subseteq L^2(\mathbb{R}^d)$, the M-invariant space generated by Φ is

$$V_M(\Phi) = \overline{\operatorname{span}}(\{t_m\varphi \,:\, m \in M \;,\; \varphi \in \Phi\}).$$

If $\Phi = \{\varphi\}$ we write $V_M(\varphi)$ and we say that $V_M(\varphi)$ is a principal M-invariant space. For simplicity of notation, when $M = \mathbb{Z}^d$, we write $V(\varphi)$ instead of $V_{\mathbb{Z}^d}(\varphi)$.

Principal SISs have been completely characterized by [dBDR94] (see also [dBDVR94],[RS95]) as follows.

Theorem 5.4.2. *Let $f \in L^2(\mathbb{R}^d)$ be given. If $g \in V(f)$, then there exists a $\mathbb{Z}^d$-periodic function η such that $\widehat{g} = \eta\widehat{f}$.*

Conversely, if η is a $\mathbb{Z}^d$-periodic function such that $\eta\widehat{f} \in L^2(\mathbb{R}^d)$, then the function g defined by $\widehat{g} = \eta\widehat{f}$ belongs to $V(f)$.

The aim of this section is to generalize the previous theorem to the M-invariant case. In case that M is discrete, Theorem 5.4.2 follows easily by rescaling. The difficulty arises when M is not discrete.

Theorem 5.4.3. *Let $f \in L^2(\mathbb{R}^d)$ and M a closed subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$. If $g \in V_M(f)$, then there exists an M^*-periodic function η such that $\widehat{g} = \eta\widehat{f}$.*

Conversely, if η is an M^*-periodic function such that $\eta\widehat{f} \in L^2(\mathbb{R}^d)$, then the function g defined by $\widehat{g} = \eta\widehat{f}$ belongs to $V_M(f)$.

Theorem 5.4.3 was proved in [dBDR94] for the lattice case. We adapt their arguments to this more general case.

We will first need some definitions and properties.

By Remark 5.3.12, there exists a linear transformation $T : \mathbb{R}^d \to \mathbb{R}^d$ such that $M = T(\frac{1}{a_1}\mathbb{Z}\times\cdots\times\frac{1}{a_q}\mathbb{Z}\times\mathbb{R}^{d-q})$ and $M^* = (T^*)^{-1}(a_1\mathbb{Z}\times\cdots\times a_q\mathbb{Z}\times\{0\}^{d-q})$, where $q = d - \mathbf{d}(M)$.

We will denote by $\mathcal{D}$ the section of the quotient $\mathbb{R}^d/M^*$ defined as

$$\mathcal{D} = (T^*)^{-1}([0, a_1)\times\cdots\times[0, a_q)\times\mathbb{R}^{d-q}). \tag{5.4}$$

Therefore, $\{\mathcal{D}+m^*\}_{m^*\in M^*}$ forms a partition of $\mathbb{R}^d$.

Given $f, g \in L^2(\mathbb{R}^d)$ we define

$$[f,g](\omega) := \sum_{m^*\in M^*} f(\omega+m^*)\overline{g(\omega+m^*)},$$

where $\omega \in \mathcal{D}$. Note that, as $f, g \in L^2(\mathbb{R}^d)$ we have that $[f,g] \in L^1(\mathcal{D})$, since

$$\begin{aligned}
\int_{\mathbb{R}^d} f(\omega)\overline{g(\omega)}\,d\omega &= \sum_{m^*\in M^*}\int_{\mathcal{D}+m^*} f(\omega)\overline{g(\omega)}\,d\omega \\
&= \sum_{m^*\in M^*}\int_{\mathcal{D}} f(\omega+m^*)\overline{g(\omega+m^*)}\,d\omega \\
&= \int_{\mathcal{D}} [f,g](\omega)\,d\omega. \qquad (5.5)
\end{aligned}$$

From this, it follows that if $f \in L^2(\mathbb{R}^d)$, then $\{f(\omega+m^*)\}_{m^*\in M^*} \in \ell^2(M^*)$ a.e. $\omega \in \mathcal{D}$.

The Cauchy-Schwarz inequality in $\ell^2(M^*)$, gives the following a.e. pointwise estimate

$$|[f,g]|^2 \leq [f,f][g,g] \qquad (5.6)$$

for every $f, g \in L^2(\mathbb{R}^d)$.

Given an M^*-periodic function η and $f, g \in L^2(\mathbb{R}^d)$ such that $\eta f \in L^2(\mathbb{R}^d)$, it is easy to check that

$$[\eta f, g] = \eta[f,g]. \qquad (5.7)$$

The following lemma is an extension to general subgroups of $\mathbb{R}^d$ of a result which holds for the discrete case.

Lemma 5.4.4. *Let $f \in L^2(\mathbb{R}^d)$, M a closed subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$ and $\mathcal{D}$ defined as in (5.4). Then,*

$$V_M(f)^\perp = \{g \in L^2(\mathbb{R}^d) \,:\, [\widehat{f},\widehat{g}](\omega) = 0 \text{ a.e. } \omega \in \mathcal{D}\}.$$

Proof. Since the span of the set $\{t_m f \,:\, m \in M\}$ is dense in $V_M(f)$, we have that $g \in V_M(f)^\perp$ if and only if $\langle \widehat{g}, e_m \widehat{f}\rangle = 0$ for all $m \in M$. As e_m is M^*-periodic, using (5.5) and (5.7), we obtain that $g \in V_M(f)^\perp$ if and only if

$$\int_{\mathcal{D}} e_m(\omega)[\widehat{f},\widehat{g}](\omega)\,d\omega = 0, \qquad (5.8)$$

for all $m \in M$.

At this point, what is left to show is that if (5.8) holds then $[\widehat{f},\widehat{g}](\omega) = 0$ a.e. $\omega \in \mathcal{D}$. For this, taking into account that $[\widehat{f},\widehat{g}] \in L^1(\mathcal{D})$, it is enough to prove that if $h \in L^1(\mathcal{D})$ and $\int_{\mathcal{D}} he_m = 0$ for all $m \in M$ then $h = 0$ a.e $\omega \in \mathcal{D}$.

We will prove the preceding property for the case $M = \mathbb{Z}^q \times \mathbb{R}^{d-q}$. The general case will follow from a change of variables using the description of M and $\mathcal{D}$ given in Remark 5.3.12 and (5.4).

Suppose now $M = \mathbb{Z}^q \times \mathbb{R}^{d-q}$, then $\mathcal{D} = [0,1)^q \times \mathbb{R}^{n-q}$. Take $h \in L^1(\mathcal{D})$, such that

$$\iint_{[0,1)^q \times \mathbb{R}^{n-q}} h(x,y)e^{-2\pi i(kx+ty)}\,dxdy = 0 \quad \forall\, k \in \mathbb{Z}^q,\ t \in \mathbb{R}^{d-q}. \tag{5.9}$$

Given $k \in \mathbb{Z}^q$, define $\alpha_k(y) := \int_{[0,1)^q} h(x,y)e^{-2\pi ikx}\,dx$ for a.e. $y \in \mathbb{R}^{d-q}$. It follows from (5.9) that

$$\int_{\mathbb{R}^{d-q}} \alpha_k(y)e^{-2\pi ity}\,dy = 0 \quad \forall\, t \in \mathbb{R}^{d-q}. \tag{5.10}$$

Since $h \in L^1(\mathcal{D})$, by Fubini's Theorem, $\alpha_k \in L^1([0,1)^q)$. Thus, using (5.10), $\alpha_k(y) = 0$ a.e. $y \in \mathbb{R}^{d-q}$. That is

$$\int_{[0,1)^q} h(x,y)e^{-2\pi ikx}\,dx = 0 \tag{5.11}$$

for a.e. $y \in \mathbb{R}^{d-q}$. Define now $\beta_y(x) := h(x,y)$. By (5.11), for a.e. $y \in \mathbb{R}^{d-q}$ we have that $\beta_y(x) = 0$ for a.e. $x \in [0,1)^q$. Therefore, $h(x,y) = 0$ a.e. $(x,y) \in [0,1)^q \times \mathbb{R}^{d-q}$ and this completes the proof. □

Now we will give a formula for the orthogonal projection onto $V_M(f)$.

Lemma 5.4.5. *Let P be the orthogonal projection onto $V_M(f)$. Then, for each $g \in L^2(\mathbb{R}^d)$, we have $\widehat{Pg} = \eta_g \widehat{f}$, where η_g is the M^*-periodic function defined by*

$$\eta_g := \begin{cases} [\widehat{g},\widehat{f}]/[\widehat{f},\widehat{f}] & \text{on } E_f + M^* \\ 0 & \text{otherwise,} \end{cases}$$

and E_f is the set $\{\omega \in \mathcal{D} : [\widehat{f},\widehat{f}](\omega) \neq 0\}$.

Proof. Let $\widehat{P}$ be the orthogonal projection onto $\widehat{V_M(f)}$. Since $\widehat{Pg} = \widehat{P}\widehat{g}$, it is enough to show that $\widehat{P}\widehat{g} = \eta_g \widehat{f}$.

We first want to prove that $\eta_g \widehat{f} \in L^2(\mathbb{R}^d)$. Combining (5.5), (5.6) and (5.7)

$$\int_{\mathbb{R}^d} |\eta_g \widehat{f}|^2 = \int_{\mathcal{D}} |\eta_g|^2 [\widehat{f},\widehat{f}] \le \int_{\mathcal{D}} [\widehat{g},\widehat{g}] = \|g\|_{L^2}^2,$$

and so, $\eta_g \widehat{f} \in L^2(\mathbb{R}^d)$. Define the linear map

$$Q : L^2(\mathbb{R}^d) \longrightarrow L^2(\mathbb{R}^d), \quad Q\widehat{g} = \eta_g \widehat{f},$$

which is well defined and has norm not greater than one. We will prove that $Q = \widehat{P}$.

Take $\widehat{g} \in \widehat{V_M(f)}^{\perp} = (V_M(f)^{\perp})^{\wedge}$. Then Lemma 5.4.4 gives that $\eta_g = 0$, hence $Q\widehat{g} = 0$. Therefore, $Q = \widehat{P}$ on $\widehat{V_M(f)}^{\perp}$.

On the other hand, on $E_f + M^*$,

$$\eta_{(t_m f)} = [e_m \widehat{f}, \widehat{f}]/[\widehat{f}, \widehat{f}] = e_m \quad \forall\, m \in M.$$

Since $\widehat{f} = 0$ outside of $E_f + M^*$, we have that $Q(\widehat{t_m f}) = e_m \widehat{f}$. As Q is linear and bounded, and the set $\operatorname{span}\{t_m f \,:\, m \in M\}$ is dense in $V_M(f)$, $Q = \widehat{P}$ on $\widehat{V_M(f)}$.

□

Proof of Theorem 5.4.3. Suppose that $g \in V_M(f)$, then $Pg = g$, where P is the orthogonal projection onto $V_M(f)$. Hence, by Lemma 5.4.5, $\widehat{g} = \eta_g \widehat{f}$.

Conversely, if $\eta \widehat{f} \in L^2(\mathbb{R}^d)$ and η is an M^*-periodic function, then g, the inverse transform of $\eta \widehat{f}$ is also in $L^2(\mathbb{R}^d)$ and satisfies, by (5.7), that $\eta_g = [\eta \widehat{f}, \widehat{f}]/[\widehat{f}, \widehat{f}] = \eta$ on $E_f + M^*$.

On the other hand, since $\operatorname{supp}(\widehat{f}) \subseteq E_f + M^*$, we have that $\eta_g \widehat{f} = \eta \widehat{f}$.

So, $\widehat{Pg} = \eta_g \widehat{f} = \eta \widehat{f} = \widehat{g}$. Consequently, $Pg = g$, and hence $g \in V_M(f)$.

□

5.5 Characterization of the extra invariance

Given M a closed subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$, our goal is to characterize when a SIS V is an M-invariant space. For this, we will construct a partition $\{B_\sigma\}_{\sigma \in \mathcal{N}}$ of $\mathbb{R}^d$, where each B_σ will be an M^*-periodic.

Let Ω be a measurable section of the quotient $\mathbb{R}^d/\mathbb{Z}^d$. Then Ω tiles $\mathbb{R}^d$ by $\mathbb{Z}^d$ translations, that is

$$\mathbb{R}^d = \bigcup_{k \in \mathbb{Z}^d} \Omega + k. \tag{5.12}$$

Now, for each $k \in \mathbb{Z}^d$, consider $(\Omega + k) + M^*$. Although these sets are M^*-periodic, they are not a partition of $\mathbb{R}^d$. So, we need to choose a subset $\mathcal{N}$ of $\mathbb{Z}^d$ such that if $\sigma, \sigma' \in \mathcal{N}$ and $\sigma + M^* = \sigma' + M^*$, then $\sigma = \sigma'$. Thus $\mathcal{N}$ should be a section of the quotient $\mathbb{Z}^d/M^*$.

Given $\sigma \in \mathcal{N}$ we define

$$B_\sigma = \Omega + \sigma + M^* = \bigcup_{m^* \in M^*} (\Omega + \sigma) + m^*. \tag{5.13}$$

Note that, in the notation of Subsection 5.3.2, we can choose the sets Ω and $\mathcal{N}$ as:

$$\Omega = (T^*)^{-1}([0,1)^d), \tag{5.14}$$

and

$$\mathcal{N} = (T^*)^{-1}(\{0, \ldots, a_1 - 1\} \times \ldots \times \{0, \ldots, a_q - 1\} \times \mathbb{Z}^{d-q}), \tag{5.15}$$

where T is as in Remark 5.3.12 and $a_1, \ldots, a_q$ are the invariant factors of M.

Below we give three basic examples of the construction of the partition $\{B_\sigma\}_{\sigma \in \mathcal{N}}$.

Example 5.5.1.

(1) Let $M = \frac{1}{n}\mathbb{Z} \subseteq \mathbb{R}$, then $M^* = n\mathbb{Z}$. We choose $\Omega = [0,1)$ and $\mathcal{N} = \{0,\ldots,n-1\}$. Given $\sigma \in \{0,\ldots,n-1\}$, we have

$$B_\sigma = \bigcup_{m^* \in n\mathbb{Z}} ([0,1)+\sigma) + m^* = \bigcup_{j\in\mathbb{Z}} [\sigma, \sigma+1) + nj.$$

Figure 5.1 illustrates the partition for $n = 4$. In the picture, the black dots represent the set $\mathcal{N}$. The set B_2 is the one which appears in gray.

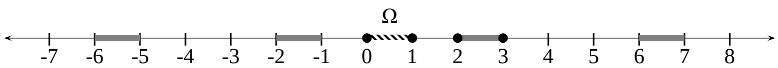

Figure 5.1: Partition of the real line for $M = \frac{1}{4}\mathbb{Z}$.

(2) Let $M = \frac{1}{2}\mathbb{Z} \times \mathbb{R}$, then $M^* = 2\mathbb{Z} \times \{0\}$. We choose $\Omega = [0,1)^2$, and $\mathcal{N} = \{0,1\} \times \mathbb{Z}$.

So, the sets $B_{(i,j)}$ are

$$B_{(i,j)} = \bigcup_{k\in\mathbb{Z}} ([0,1)^2 + (i,j)) + (2k, 0)$$

where $(i,j) \in \mathcal{N}$. See Figure 5.2, where the sets $B_{(0,0)}$, $B_{(1,1)}$ and $B_{(1,-1)}$ are represented by the squares painted in light gray, gray and dark gray respectively. As in the previous figure, the set $\mathcal{N}$ is represented by the black dots.

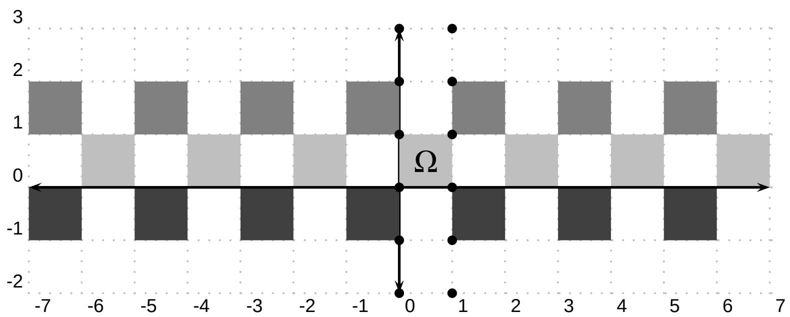

Figure 5.2: Partition of the plane for $M = \frac{1}{2}\mathbb{Z} \times \mathbb{R}$.

(3) Let $M = \{k\frac{1}{3}v_1 + tv_2 \ : \ k \in \mathbb{Z} \text{ and } t \in \mathbb{R}\}$, where $v_1 = (1,0)$ and $v_2 = (-1,1)$. Then, $\{v_1, v_2\}$ satisfy conditions in Theorem 5.3.9. By Corollary 5.3.10, $M^* = \{k3w_1 \ : \ k \in \mathbb{Z}\}$, where $w_1 = (1,1)$ and $w_2 = (0,1)$.

The sets Ω and $\mathcal{N}$ can be chosen in terms of w_1 and w_2 as

$$\Omega = \{tw_1 + sw_2 : \ t, s \in [0,1)\}$$

and

$$\mathcal{N} = \{aw_1 + kw_2 : \ a \in \{0, 1, 2\}, \ k \in \mathbb{Z}\}.$$

This is illustrated in Figure 5.3. In this case the sets $B_{(0,0)}$, $B_{(1,0)}$ and $B_{(1,2)}$ correspond to the light gray, gray and dark gray regions respectively. And again, the black dots represent the set $\mathcal{N}$.

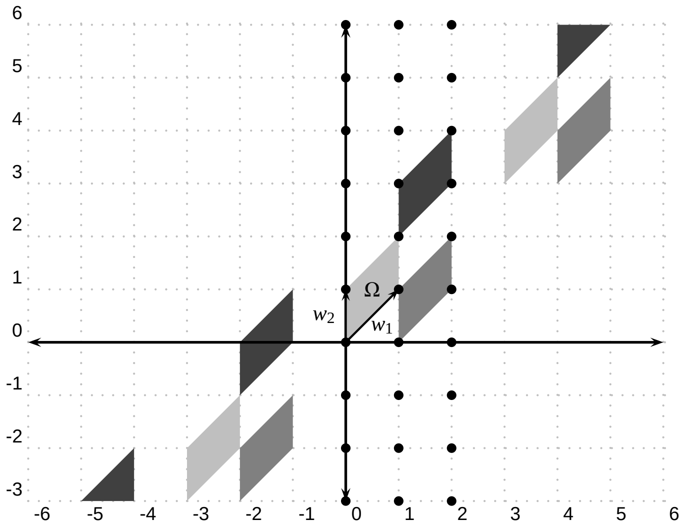

Figure 5.3: Partition for $M = \{k\frac{1}{3}(1, 0) + t(-1, 1) \ : \ k \in \mathbb{Z} \text{ and } t \in \mathbb{R}\}$.

Once the partition $\{B_\sigma\}_{\sigma\in\mathcal{N}}$ is set, for each $\sigma \in \mathcal{N}$, we define the subspaces

$$U_\sigma = \{f \in L^2(\mathbb{R}^d) : \widehat{f} = \chi_{B_\sigma}\widehat{g}, \ \text{with } g \in V\}. \tag{5.16}$$

5.5.1 Characterization of the extra invariance in terms of subspaces

The main theorem of this section characterizes the M-invariance of V in terms of the subspaces U_σ (see (5.16)).

Theorem 5.5.2. *If $V \subseteq L^2(\mathbb{R}^d)$ is a SIS and M is a closed subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$, then the following are equivalent.*

i) V is M-invariant.

ii) $U_\sigma \subseteq V$ for all $\sigma \in \mathcal{N}$.

Moreover, in case any of the above holds, we have that V is the orthogonal direct sum

$$V = \dot{\bigoplus_{\sigma\in\mathcal{N}}} U_\sigma.$$

Below we state a lemma which will be necessary to prove Theorem 5.5.2.

Lemma 5.5.3. *Let V be a SIS and $\sigma \in \mathcal{N}$. Assume that the subspace U_σ defined in (5.16) satisfies $U_\sigma \subseteq V$. Then, U_σ is a closed subspace which is M-invariant and in particular is a SIS.*

Proof. Let us first prove that U_σ is closed. Suppose that $f_j \in U_\sigma$ and $f_j \to f$ in $L^2(\mathbb{R}^d)$. Since $U_\sigma \subseteq V$ and V is closed, f must be in V. Further,

$$\|\widehat{f_j} - \widehat{f}\|_2^2 = \|(\widehat{f_j} - \widehat{f})\chi_{B_\sigma}\|_2^2 + \|(\widehat{f_j} - \widehat{f})\chi_{B_\sigma^c}\|_2^2 = \|\widehat{f_j} - \widehat{f}\chi_{B_\sigma}\|_2^2 + \|\widehat{f}\chi_{B_\sigma^c}\|_2^2.$$

Since the left-hand side converges to zero, we must have that $\widehat{f}\chi_{B_\sigma^c} = 0$ a.e. $\omega \in \mathbb{R}^d$, and $\widehat{f_j} \to \widehat{f}\chi_{B_\sigma}$ in $L^2(\mathbb{R}^d)$. Then, $\widehat{f} = \widehat{f}\chi_{B_\sigma}$. Consequently $f \in U_\sigma$, so U_σ is closed.

Now we show that U_σ is M-invariant. Given $m \in M$ and $f \in U_\sigma$, we will prove that $e_m\widehat{f} \in \widehat{U_\sigma}$. Since $f \in U_\sigma$, there exists $g \in V$ such that $\widehat{f} = \chi_{B_\sigma}\widehat{g}$. Hence,

$$e_m\widehat{f} = e_m(\chi_{B_\sigma}\widehat{g}) = \chi_{B_\sigma}(e_m\widehat{g}). \tag{5.17}$$

If we can find a $\mathbb{Z}^d$-periodic function ℓ_m verifying

$$e_m(\omega) = \ell_m(\omega) \quad \text{a.e.}\, \omega \in B_\sigma, \tag{5.18}$$

then, we can rewrite (5.17) as

$$e_m\widehat{f} = \chi_{B_\sigma}(\ell_m\widehat{g}).$$

By Theorem 5.4.2, $\ell_m\widehat{g} \in \widehat{V(g)} \subseteq \widehat{V}$ and so, $e_m\widehat{f} \in \widehat{U}_\sigma$.

Let us now define the function ℓ_m. Note that, since e_m is M^*-periodic,

$$e_m(\omega + \sigma) = e_m(\omega + \sigma + m^*) \quad \text{a.e. } \omega \in \Omega,\ \forall\, m^* \in M^*. \tag{5.19}$$

For each $k \in \mathbb{Z}^d$, set

$$\ell_m(\omega + k) = e_m(\omega + \sigma) \quad \text{a.e. } \omega \in \Omega. \tag{5.20}$$

It is clear that ℓ_m is $\mathbb{Z}^d$-periodic and combining (5.19) with (5.20), we obtain (5.18).

Note that, since $\mathbb{Z}^d \subseteq M$, the $\mathbb{Z}^d$-invariance of U_σ is a consequence of the M-invariance. □

Proof of Theorem 5.5.2. i)$\Rightarrow$ii): Fix $\sigma \in \mathcal{N}$ and $f \in U_\sigma$. Then $\widehat{f} = \chi_{B_\sigma}\widehat{g}$ for some $g \in V$. Since χ_{B_σ} is an M^*-periodic function, by Theorem 5.4.3, we have that $f \in V_M(g) \subseteq V$, as we wanted to prove.

ii)$\Rightarrow$i): Suppose that $U_\sigma \subseteq V$ for all $\sigma \in \mathcal{N}$. Note that Lemma 5.5.3 implies that U_σ is M-invariant, and we also have that the subspaces U_σ are mutually orthogonal since the sets B_σ are disjoint.

Take $f \in V$. Then, since $\{B_\sigma\}_{\sigma\in\mathcal{N}}$ is a partition of $\mathbb{R}^d$, we can decompose f as $f = \sum_{\sigma\in\mathcal{N}} f^\sigma$ where f^σ is such that $\widehat{f^\sigma} = \widehat{f}\chi_{B_\sigma}$. This implies that $f \in \dot{\bigoplus}_{\sigma\in\mathcal{N}} U_\sigma$ and consequently, V is the orthogonal direct sum

$$V = \dot{\bigoplus_{\sigma\in\mathcal{N}}} U_\sigma.$$

As each U_σ is M-invariant, so is V. □

5.5.2 Characterization of the extra invariance in terms of fibers

The aim of this section is to express the conditions of Theorem 5.5.2 in terms of fibers. We will also give a useful characterization of the M-invariance for an FSIS in terms of the Gramian.

If $f \in L^2(\mathbb{R}^d)$ and $\sigma \in \mathcal{N}$, let f^σ denote the function defined by

$$\widehat{f^\sigma} = \widehat{f}\chi_{B_\sigma}.$$

Let P_σ be the orthogonal projection onto S_σ, where

$$S_\sigma := \{f \in L^2(\mathbb{R}^d) \,:\, \operatorname{supp}(\widehat{f}) \subseteq B_\sigma\}. \tag{5.21}$$

Therefore

$$f^\sigma = P_\sigma f \quad \text{and} \quad U_\sigma = P_\sigma(V) = \{f^\sigma \,:\, f \in V\}. \tag{5.22}$$

Remark 5.5.4. In Lemma 5.5.3 we have proved that the spaces U_σ are closed if they are included in V, but it is important to observe that if this hypothesis is not satisfied, they might not be closed (see Example 5.5.5). More precisely, if V, W are two closed subspaces of a Hilbert space $\mathcal{H}$, then $P_W(V)$ is a closed subspace of $\mathcal{H}$ if and only if $V + W^\perp$ is closed (see [Deu95]). So, as a consequence of (5.22), in the notation of Chapter 4, U_σ will be a closed subspace if and only if the Friedrichs angle satisfies $\mathbf{c}[V, S_\sigma^\perp] < 1$.

We include below an example of a SIS V and a group M for which the subspace U_σ is not closed.

Example 5.5.5. Let $V = V(\varphi)$ where $\varphi = \chi_{[-\frac{1}{2},\frac{1}{2})}$. Consider the discrete group $M = \frac{1}{2}\mathbb{Z}$. If $B_0 = [0, 1) + 2\mathbb{Z}$, we will prove that the subspace $U_0 = \{f \in L^2(\mathbb{R}) : \widehat{f} = \chi_{B_0}\widehat{g},\ \text{with } g \in V\}$ is not closed.

Using the remark from above, it is enough to show that $\mathbf{c}[V, S_0^\perp] = 1$, where $S_0^\perp = S_1 = \{f \in L^2(\mathbb{R}) \,:\, \operatorname{supp}(\widehat{f}) \subseteq B_1\}$ and $B_1 = [0, 1) + 2\mathbb{Z} + 1$.

From Lemma 4.4.5, we have that $\mathbf{c}[V, S_1] = \operatorname{ess-sup}\{\mathbf{c}[J_V(\omega), J_{S_1}(\omega)] \,:\, \omega \in [0, 1)\}$.

Note that $J_V(\omega) = \operatorname{span}\{\tau\varphi(\omega)\}$ and $J_{S_1}(\omega) = \overline{\operatorname{span}}\{e_{2j+1}\}_{j\in\mathbb{Z}}$. So, $J_V(\omega) \cap J_{S_1}(\omega) = \{0\}$. Therefore,

$$\mathbf{c}[J_V(\omega), J_{S_1}(\omega)] = \sup_{j\in\mathbb{Z}} |\operatorname{sinc}(\omega + 2j + 1)|.$$

Then, we obtain that,

$$\mathbf{c}[V, S_1] = \text{ess-sup}\left\{\sup_{j\in\mathbb{Z}} |\text{sinc}(\omega + 2j + 1)| \;:\; \omega \in [0,1)\right\} = 1.$$

Thus, U_0 is not closed.

As we have proved above, the subspaces U_σ might not be closed, so we will need a generalization of the concepts of fiber space and dimension function of SISs (see Section 1.4.1) to this spaces.

Note that although the domain of a range function from Definition 1.5.2 was $[0,1)^d$, it is easy to prove that the same analysis from [Bow00] holds for any measurable section Ω of the quotient $\mathbb{R}^d/\mathbb{Z}^d$.

Let V be a SIS and U_σ be defined as in (5.16). If Φ a countable subset of $L^2(\mathbb{R}^d)$ such that $V = V(\Phi)$, then for $\omega \in \Omega$ we define the subspace $\mathcal{J}_{U_\sigma}(\omega)$ as

$$\mathcal{J}_{U_\sigma}(\omega) = \overline{\text{span}}\{\tau\varphi^\sigma(\omega) \;:\; \varphi \in \Phi\}. \tag{5.23}$$

Note that when U_σ is closed, it is a SIS, so the subspace $\mathcal{J}_{U_\sigma}(\omega)$ is the fiber space $J_{U_\sigma}(\omega)$ defined in Proposition 1.5.3.

Remark 5.5.6. The fibers

$$\tau\varphi^\sigma(\omega) = \{\chi_{B_\sigma}(\omega + k)\widehat{\varphi}(\omega + k)\}_{k\in\mathbb{Z}^d}$$

can be described in a simple way as

$$\chi_{B_\sigma}(\omega + k)\widehat{\varphi}(\omega + k) = \begin{cases} \widehat{\varphi}(\omega + k) & \text{if } k \in \sigma + M^* \\ 0 & \text{otherwise .} \end{cases}$$

Therefore, if $\sigma \neq \sigma'$, $\mathcal{J}_{U_\sigma}(\omega)$ is orthogonal to $\mathcal{J}_{U_{\sigma'}}(\omega)$ for a.e. $\omega \in \Omega$.

Theorem 5.5.7. *Let V be a SIS generated by a countable set $\Phi \subseteq L^2(\mathbb{R}^d)$. The following statements are equivalent.*

i) V is M-invariant.

ii) $\tau\varphi^\sigma(\omega) \in J_V(\omega)$ a.e. $\omega \in \Omega$ for all $\varphi \in \Phi$ and $\sigma \in \mathcal{N}$.

Proof. i)$\Rightarrow$ii): By Theorem 5.5.2, $U_\sigma \subseteq V$ for any $\sigma \in \mathcal{N}$. Using this and (5.22), for a given $\varphi \in \Phi$, we have that $\varphi^\sigma \in V$, so $\tau\varphi^\sigma(\omega) \in J_V(\omega)$.

ii)$\Rightarrow$i): Fix $\sigma \in \mathcal{N}$, we will prove that $U_\sigma \subseteq V$. Let $f \in U_\sigma$, we will show that $\tau f(\omega) \in J_V(\omega)$ for a.e. $\omega \in \Omega$.

For all $\varphi \in \Phi$, $\tau\varphi^\sigma(\omega) \in J_V(\omega)$ a.e. $\omega \in \Omega$, so, it follows that $\mathcal{J}_{U_\sigma}(\omega) \subseteq J_V(\omega)$ a.e. $\omega \in \Omega$. Thus, it is enough to prove that $\tau f(\omega) \in \mathcal{J}_{U_\sigma}(\omega)$ for a.e. $\omega \in \Omega$.

Since $f \in U_\sigma$, there exists $g \in V$ such that $f = g^\sigma$.

The subspace S_σ defined in (5.21) is a SIS, so by Proposition 1.5.4, we obtain

$$\tau f(\omega) = \tau g^\sigma(\omega) = \tau(P_\sigma g)(\omega) = P_{J_\sigma(\omega)}(\tau g(\omega)), \tag{5.24}$$

where $J_\sigma(\omega)$ is the fiber space associated to S_σ.

Since $g \in V$, we have that $\tau g(\omega) \in J_V(\omega) = \overline{\operatorname{span}}\{\tau\varphi(\omega) : \varphi \in \Phi\}$. So,

$$\tau f(\omega) = P_{J_\sigma(\omega)}(\tau g(\omega)) \in P_{J_\sigma(\omega)}(\overline{\operatorname{span}}\{\tau\varphi(\omega) : \varphi \in \Phi\}).$$

The proof follows using that

$$P_{J_\sigma(\omega)}(\overline{\operatorname{span}}\{\tau\varphi(\omega) : \varphi \in \Phi\}) \subseteq \overline{\operatorname{span}}\{\tau\varphi^\sigma(\omega) \ : \ \varphi \in \Phi\} = \mathcal{J}_{U_\sigma}(\omega).$$

□

Now we give a slightly simpler characterization of M-invariance for the finitely generated case.

For $\omega \in \Omega$ by abuse of notation, we will write $\dim_{U_\sigma}(\omega)$ for the dimension of the subspace $\mathcal{J}_{U_\sigma}(\omega)$.

Theorem 5.5.8. *If V is an FSIS generated by Φ, then the following statements are equivalent.*

i) *V is M-invariant.*

ii) *For almost every $\omega \in \Omega$,* $\dim_V(\omega) = \sum_{\sigma\in\mathcal{N}} \dim_{U_\sigma}(\omega)$.

iii) *For almost every $\omega \in \Omega$,* $\operatorname{rank}[\mathcal{G}_\Phi(\omega)] = \sum_{\sigma\in\mathcal{N}} \operatorname{rank}[\mathcal{G}_{\Phi^\sigma}(\omega)]$,

where $\Phi^\sigma = \{\varphi^\sigma \ : \ \varphi \in \Phi\}$.

Proof. i)$\Rightarrow$ii): By Lemma 5.5.3 and Theorem 5.5.2, U_σ is a SIS for each $\sigma \in \mathcal{N}$ and $V = \dot\bigoplus_{\sigma\in\mathcal{N}} U_\sigma$. Then, ii) follows from Proposition 1.5.5.

ii)$\Rightarrow$i): Given $\varphi \in \Phi$, for $\omega \in \Omega$ we have that

$$\tau\varphi(\omega) = \sum_{\sigma\in\mathcal{N}} \tau\varphi^\sigma(\omega).$$

Then, $\tau\varphi(\omega) \in \dot\bigoplus_{\sigma\in\mathcal{N}} \mathcal{J}_{U_\sigma}(\omega)$ for a.e. $\omega \in \Omega$. Note that the orthogonality of the subspaces $\mathcal{J}_{U_\sigma}(\omega)$ is a consequence of Remark 5.5.6.

Since $J_V(\omega) = \overline{\operatorname{span}}\{\tau\varphi(\omega) \ : \ \varphi \in \Phi\}$, it follows that

$$J_V(\omega) \subseteq \dot\bigoplus_{\sigma\in\mathcal{N}} \mathcal{J}_{U_\sigma}(\omega).$$

Using ii), we obtain that $J_V(\omega) = \dot\bigoplus_{\sigma\in\mathcal{N}} \mathcal{J}_{U_\sigma}(\omega)$. This implies that $\tau\varphi^\sigma(\omega) \in J_V(\omega)$ for all $\sigma \in \mathcal{N}$, $\varphi \in \Phi$. The proof follows as a consequence of Theorem 5.5.7.

The equivalence between ii) and iii) follows from Proposition 1.3.5.

□

5.6 Applications of the extra invariance characterizations

In this section we present two applications of the results given before. First, we will estimate the size of the supports of the Fourier transforms of the generators of an FSIS which is also M-invariant. Finally, given M a closed subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$, we will construct a shift-invariant space V which is exactly M-invariant. That is, V will not be invariant under any other closed subgroup containing M.

Theorem 5.6.1. *Let V be an FSIS generated by $\{\varphi_1, \dots, \varphi_\ell\}$, and define*

$$E_j = \{\omega \in \Omega \,:\, \dim_V(\omega) = j\}, \qquad j = 0, \dots, \ell.$$

If V is M-invariant and $\mathcal{D}'$ is any measurable section of $\mathbb{R}^d/M^*$, then

$$|\{y \in \mathcal{D}' \,:\, \widehat{\varphi_h}(y) \neq 0\}| \leq \sum_{j=0}^{\ell} j|E_j| \leq \ell,$$

for each $h = 1, \dots, \ell$.

Proof. The measurability of the sets E_j follows from the results of Helson [Hel64], e.g., see [BK06] for an argument of this type.

Fix any $h \in \{0, \dots, \ell\}$. Note that, as a consequence of Remark 5.5.6, if $J_{U_\sigma}(\omega) = \{0\}$, then $\widehat{\varphi_h}(\omega + \sigma + m^*) = 0$ for all $m^* \in M^*$.

On the other hand, since $\{\Omega + \sigma + m^*\}_{\sigma \in \mathcal{N}, m^* \in M^*}$ is a partition of $\mathbb{R}^d$, if $\omega \in \Omega$ and $\sigma \in \mathcal{N}$ are fixed, there exists a unique $m^*_{(\omega,\sigma)} \in M^*$ such that $\omega + \sigma + m^*_{(\omega,\sigma)} \in \mathcal{D}'$.

So,

$$\{\sigma \in \mathcal{N} \,:\, \widehat{\varphi_h}(\omega + \sigma + m^*_{(\omega,\sigma)}) \neq 0\} \subseteq \{\sigma \in \mathcal{N} \,:\, \dim_{U_\sigma}(\omega) \neq 0\}.$$

Therefore

$$\begin{aligned}
\#\{\sigma \in \mathcal{N} \,:\, \widehat{\varphi_h}(\omega + \sigma + m^*_{(\omega,\sigma)}) \neq 0\} &\leq \#\{\sigma \in \mathcal{N} \,:\, \dim_{U_\sigma}(\omega) \neq 0\} \\
&\leq \sum_{\sigma \in \mathcal{N}} \dim_{U_\sigma}(\omega) \\
&= \dim_V(\omega).
\end{aligned}$$

Consequently, by Fubini's Theorem,

$$\begin{aligned}
|\{y \in \mathcal{D}' \,:\, \widehat{\varphi_h}(y) \neq 0\}| &= \sum_{\sigma \in \mathcal{N}} |\{\omega \in \Omega \,:\, \widehat{\varphi_h}(\omega + \sigma + m^*_{(\omega,\sigma)}) \neq 0\}| \\
&= |\{(\omega, \sigma) \in \Omega \times \mathcal{N} \,:\, \widehat{\varphi_h}(\omega + \sigma + m^*_{(\omega,\sigma)}) \neq 0\}| \\
&= \int_\Omega \#\{\sigma \in \mathcal{N} \,:\, \widehat{\varphi_h}(\omega + \sigma + m^*_{(\omega,\sigma)}) \neq 0\}\, dw \\
&\leq \int_\Omega \dim_V(\omega) dw = \sum_{j=0}^{\ell} j|E_j| \leq \ell.
\end{aligned}$$

□

When M is not discrete, the previous theorem shows that, despite the fact that $\mathcal{D}'$ has infinite measure, the support of $\widehat{\varphi_h}$ in $\mathcal{D}'$ has finite measure.

On the other hand, if M is discrete, the measure of $\mathcal{D}'$ is equal to the measure of the section $\mathcal{D}$ given by (5.4). That is

$$|\mathcal{D}'| = |\mathcal{D}| = a_1 \dots a_d,$$

where $a_1, \dots, a_d$ are the invariant factors. Thus, if $a_1 \dots a_d - \ell > 0$, it follows that

$$|\{y \in \mathcal{D}' \ : \ \widehat{\varphi_h}(y) = 0\}| \geq a_1 \dots a_d - \ell. \tag{5.25}$$

As a consequence of Theorem 5.6.1 we obtain the following corollary.

Corollary 5.6.2. *Let $\varphi \in L^2(\mathbb{R}^d)$ be given. If the SIS $V(\varphi)$ is M-invariant for some closed subgroup M of $\mathbb{R}^d$ such that $\mathbb{Z}^d \subsetneq M$, then $\widehat{\varphi}$ must vanish on a set of infinite Lebesgue measure.*

Proof. Let $\mathcal{D}$ be the measurable section of $\mathbb{R}^d/M^*$ defined in (5.4). Then,

$$\mathbb{R}^d = \bigcup_{m^* \in M^*} \mathcal{D} + m^*,$$

thus

$$|\{y \in \mathbb{R}^d \ : \ \widehat{\varphi}(y) = 0\}| = \sum_{m^* \in M^*} |\{y \in \mathcal{D} + m^* \ : \ \widehat{\varphi}(y) = 0\}|.$$

If M is discrete, by (5.25), we have

$$|\{y \in \mathbb{R}^d \ : \ \widehat{\varphi}(y) = 0\}| \geq \sum_{m^* \in M^*} (|\mathcal{D}| - 1) = +\infty. \tag{5.26}$$

The last equality is due to the fact that M^* is infinite and $|\mathcal{D}| > 1$ (since $M \neq \mathbb{Z}^d$).

If M is not discrete, by Theorem 5.6.1, $|\{y \in \mathcal{D} + m^* \ : \ \widehat{\varphi}(y) = 0\}| = +\infty$, hence $|\{y \in \mathbb{R}^d \ : \ \widehat{\varphi}(y) = 0\}| = +\infty$. □

Another consequence of Theorem 5.6.1 is the following.

Corollary 5.6.3. *If $\varphi \in L^2(\mathbb{R}^d)$ and $V(\varphi)$ is $\mathbb{R}^d$-invariant, then*

$$|\mathrm{supp}(\widehat{\varphi})| \leq 1.$$

Proof. The proof is straightforward applying Theorem 5.6.1 for $M = \mathbb{R}^d$.

□

The converse of the previous corollary is not true. To see this consider the function $\varphi \in L^2(\mathbb{R}^d)$ such that $\widehat{\varphi} = \chi_{[0,1)^{d-1}\times[0,\frac{1}{2})} + \chi_{[0,1)^{d-1}\times[1,\frac{3}{2})}$. If $V := V(\varphi)$ were $\mathbb{R}^d$-invariant, by Theorem 5.5.8, we would have that $\operatorname{rank}[\mathcal{G}_\varphi(\omega)] = \sum_{j\in\mathbb{Z}^d} \operatorname{rank}[\mathcal{G}_{\varphi^j}(\omega)]$ for a.e. $\omega \in [0,1)^d$, with $\widehat{\varphi^j} = \chi_{([0,1)^d+j)}\widehat{\varphi}$. However, for $\omega \in [0,1)^{d-1} \times [0,\frac{1}{2})$ we obtain that $\operatorname{rank}[\mathcal{G}_\varphi(\omega)] = 1$ and $\operatorname{rank}[\mathcal{G}_{\varphi^0}(\omega)] = 1 = \operatorname{rank}[\mathcal{G}_{\varphi^{e_d}}(\omega)]$, with $e_d = (0,\dots,0,1)$. Thus, V can not be $\mathbb{R}^d$-invariant.

The following remark states that the converse of Corollary 5.6.3 is true if we impose some conditions on the generator φ.

Remark 5.6.4. Let $\varphi \in L^2(\mathbb{R}^d)$ such that $\operatorname{supp}(\mathcal{G}_\varphi) = [0,1)^d$. If $|\operatorname{supp}(\widehat{\varphi})| \leq 1$, then $V(\varphi)$ is $\mathbb{R}^d$-invariant.

Proof. Using the decomposition from the previous section, for each $j \in \mathbb{Z}^d = \mathcal{N}$ we have the fibers $\tau\varphi^j(\omega) = \widehat{\varphi}(\omega + j)e_j$, where $\{e_j\}$ is the canonical basis for $\ell^2(\mathbb{Z}^d)$.

By Theorem 5.5.8 $V(\varphi)$ is $\mathbb{R}^d$-invariant if and only if for almost every $\omega \in [0,1)^d$, $\operatorname{rank}[\mathcal{G}_\varphi(\omega)] = \sum_{j\in\mathbb{Z}^d} \operatorname{rank}[\mathcal{G}_{\varphi^j}(\omega)]$. Since $\operatorname{supp}(\mathcal{G}_\varphi) = [0,1)^d$, we have that $\operatorname{rank}[\mathcal{G}_\varphi(\omega)] = 1$ for almost every $\omega \in [0,1)^d$.

As $\mathcal{G}_{\varphi^j}(\omega) = |\widehat{\varphi}(\omega + j)|^2$, we obtain that

$$\operatorname{rank}[\mathcal{G}_{\varphi^j}(\omega)] = \begin{cases} 1 & \text{if } \widehat{\varphi}(\omega + j) \neq 0 \\ 0 & \text{if } \widehat{\varphi}(\omega + j) = 0 \end{cases}$$

Thus, $V(\varphi)$ is $\mathbb{R}^d$-invariant if and only if for almost every $\omega \in [0,1)^d$ there exists one and only one $j \in \mathbb{Z}^d$ such that $\widehat{\varphi}(\omega + j) \neq 0$.

For a given $\omega \in [0,1)^d$, the existence of such a $j \in \mathbb{Z}^d$ is a consequence of the fact that $\operatorname{supp}(\mathcal{G}_\varphi) = [0,1)^d$. To prove the uniqueness, we will show that for $j \in \mathbb{Z}^d$, the sets

$$N_j := \{\omega \in [0,1)^d \ : \ \widehat{\varphi}(\omega + j) \neq 0\} = (\operatorname{supp}(\widehat{\varphi}) \cap ([0,1)^d + j)) - j$$

satisfy that $|N_i \cap N_j| = 0$ for all $i \neq j$.

Since $\operatorname{supp}(\mathcal{G}_\varphi) = \bigcup_{j\in\mathbb{Z}^d} N_j$, we obtain

$$\begin{aligned} 1 &= |\operatorname{supp}(\mathcal{G}_\varphi)| = \Big|\bigcup_{j\in\mathbb{Z}^d} N_j\Big| \leq \sum_{j\in\mathbb{Z}^d} |N_j| \\ &= \sum_{j\in\mathbb{Z}^d} |\operatorname{supp}(\widehat{\varphi}) \cap ([0,1)^d + j)| \\ &= \Big|\bigcup_{j\in\mathbb{Z}^d} (\operatorname{supp}(\widehat{\varphi}) \cap ([0,1)^d + j))\Big| \\ &= |\operatorname{supp}(\widehat{\varphi})| \leq 1. \end{aligned}$$

Thus,

$$\Big|\bigcup_{j\in\mathbb{Z}^d} N_j\Big| = \sum_{j\in\mathbb{Z}^d} |N_j|.$$

So, $|N_i \cap N_j| = 0$ for all $i \neq j$. Therefore, $V(\varphi)$ is $\mathbb{R}^d$-invariant.

□

Observe that by Remark 1.5.14, if $\varphi \in L^2(\mathbb{R}^d)$ satisfies that $\{t_k\varphi : k \in \mathbb{Z}^d\}$ is a Riesz basis for $V(\varphi)$, then $\mathrm{supp}(\mathcal{G}_\varphi) = [0,1)^d$. So, as a consequence of the previous remark, if $\varphi \in L^2(\mathbb{R}^d)$ is such that $\{t_k\varphi : k \in \mathbb{Z}^d\}$ is a Riesz basis for $V(\varphi)$ and $|\mathrm{supp}(\widehat{\varphi})| \leq 1$, then $V(\varphi)$ is $\mathbb{R}^d$-invariant.

5.6.1 Exact invariance

Given M be a closed subgroup of $\mathbb{R}^d$, we will say that a subspace $V \subseteq L^2(\mathbb{R}^d)$ is *exactly* M*-invariant* if it is an M-invariant space that is *not* invariant under any vector outside M.

Note that due to of Proposition 5.2.1, an M-invariant space is exactly M-invariant if and only if it is not invariant under any closed subgroup M' containing M.

It is known that on the real line, the SIS generated by a function φ with compact support can only be invariant under integer translations. That is, it is exactly $\mathbb{Z}$-invariant. The following proposition extends this result to $\mathbb{R}^d$.

Proposition 5.6.5. *If a nonzero function $\varphi \in L^2(\mathbb{R}^d)$ has compact support, then $V(\varphi)$ is exactly $\mathbb{Z}^d$-invariant.*

Proof. The proof is a straightforward consequence of Corollary 5.6.2.

□

Note that the compactness of the support of $\varphi \in L^2(\mathbb{R}^d)$ is not a necessary condition for the exactly $\mathbb{Z}^d$-invariance of $V(\varphi)$. To see this, consider the function $\varphi \in L^2(\mathbb{R}^d)$ such that $\widehat{\varphi} = \chi_{[0,2)^d}$. Since $\mathrm{supp}(\widehat{\varphi}) = [0,2)^d$, it follows that φ is not compactly supported.

We claim that $V := V(\varphi)$ is exactly $\mathbb{Z}^d$-invariant. On the contrary, assume that V is M-invariant for some closed subgroup M of $\mathbb{R}^d$ such that $\mathbb{Z}^d \subsetneq M$.

For $1 \leq j \leq d$ let $e_j \in \mathbb{R}^d$ be the canonical vectors. Since $\mathbb{Z}^d \subsetneq M$, it follows that $M^* \subsetneq \mathbb{Z}^d$. Thus there exists $1 \leq j \leq d$ such that $e_j \notin M^*$.

We have that $e_j \neq 0$ in $\mathcal{N} = \mathbb{Z}^d/M^*$. Let φ^0, φ^{e_j} be the functions defined by

$$\widehat{\varphi^0} = \chi_{B_0 \cap [0,2)^d} \text{ and } \widehat{\varphi^{e_j}} = \chi_{B_{e_j} \cap [0,2)^d},$$

where $B_\sigma = [0,1)^d + \sigma + M^*$ for $\sigma \in \mathcal{N}$.

Since $[0,1)^d \subseteq B_0 \cap [0,2)^d$ and $[0,1)^d + e_j \subseteq B_{e_j} \cap [0,2)^d$, it follows that $\widehat{\varphi^0}(\omega) = \widehat{\varphi^{e_j}}(\omega + e_j) = 1$ for all $\omega \in [0,1)^d$. Thus,

$$\mathrm{rank}[\mathcal{G}_{\varphi^0}(\omega)] = 1 = \mathrm{rank}[\mathcal{G}_{\varphi^{e_j}}(\omega)], \quad \text{for } \omega \in [0,1)^d.$$

On the other hand, $\mathrm{rank}[\mathcal{G}_\varphi(\omega)] = 1$ for $\omega \in [0,1)^d$. So, by Theorem 5.5.8, V can not be M-invariant.

The next theorem shows the existence of SISs that are exactly M-invariant for every closed subgroup M of $\mathbb{R}^d$ containing $\mathbb{Z}^d$.

Theorem 5.6.6. *For each closed subgroup M of $\mathbb{R}^d$ containing $\mathbb{Z}^d$, there exists a shift-invariant space of $L^2(\mathbb{R}^d)$ which is exactly M-invariant.*

Proof. Let M be a subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$. We will construct a principal shift-invariant space that is exactly M-invariant.

Suppose that $0 \in \mathcal{N}$ and take $\varphi \in L^2(\mathbb{R}^d)$ satisfying $\operatorname{supp}(\widehat{\varphi}) = B_0$, where B_0 is defined as in (5.13). Let $V = V(\varphi)$.

Then, $U_0 = V$ and $U_\sigma = \{0\}$ for $\sigma \in \mathcal{N}$, $\sigma \neq 0$. So, as a consequence of Theorem 5.5.2, it follows that V is M-invariant.

Now, if M' is a closed subgroup such that $M \subsetneq M'$, we will show that V can not be M'-invariant.

Since $M \subseteq M'$, $(M')^* \subseteq M^*$. Consider a section $\mathcal{H}$ of the quotient $M^*/(M')^*$ containing the origin. Then, the set given by

$$\mathcal{N}' := \{\sigma + h \,:\, \sigma \in \mathcal{N},\, h \in \mathcal{H}\},$$

is a section of $\mathbb{Z}^d/(M')^*$ and $0 \in \mathcal{N}'$.

If $\{B'_\gamma\}_{\gamma\in\mathcal{N}'}$ is the partition defined in (5.13) associated to M', for each $\sigma \in \mathcal{N}$ it holds that $\{B'_{\sigma+h}\}_{h\in\mathcal{H}}$ is a partition of B_σ, since

$$B_\sigma = \Omega + \sigma + M^* = \bigcup_{h\in\mathcal{H}} \Omega + \sigma + h + (M')^* = \bigcup_{h\in\mathcal{H}} B'_{\sigma+h}. \tag{5.27}$$

We will show now that $U'_0 \not\subseteq V$, where U'_0 is the subspace defined in (5.16) for M'. Let $g \in L^2(\mathbb{R}^d)$ such that $\widehat{g} = \widehat{\varphi}\chi_{B'_0}$. Then $g \in U'_0$. Moreover, since $\operatorname{supp}(\widehat{\varphi}) = B_0$, by (5.27), $\widehat{g} \neq 0$.

Suppose that $g \in V$, then $\widehat{g} = \eta\widehat{\varphi}$ where η is a $\mathbb{Z}^d$-periodic function. Since $M \subsetneq M'$, there exists $h \in \mathcal{H}$ such that $h \neq 0$. By (5.27), $\widehat{g}$ vanishes in B'_h. Then, the $\mathbb{Z}^d$-periodicity of η implies that $\eta(y) = 0$ a.e. $y \in \mathbb{R}^d$. So $\widehat{g} = 0$, which is a contradiction.

This shows that $U'_0 \not\subseteq V$. Therefore, V is not M'-invariant.

□

5.7 Extension to LCA groups

We would like to remark here that the characterizations of the extra invariance for shift-invariant spaces are still valid for the general context of locally compact abelian (LCA) groups (see [ACP10a]). This is important in order to obtain general conditions that can be applied to different cases, as for example the case of the classic groups such as the d-dimensional torus $\mathbb{T}^d$, the discrete group $\mathbb{Z}^d$, and the finite group $\mathbb{Z}_d$.

Although in [ACP10a] we developed all the necessary theory to the complete understanding of the problem, we will not include here all the results obtained in that paper since we would need a lot of notation and technical aspects which are not congruent with the general line of these thesis. However, in this section, we will give a brief description of the problem for the LCA context and the general results obtained for this case.

Assume G is an LCA group and K is a closed subgroup of G. For $y \in G$ let us denote by t_y the translation operator acting on $L^2(G)$. That is, $t_y f(x) = f(x - y)$ for $x \in G$ and $f \in L^2(G)$. A closed subspace V of $L^2(G)$ satisfying that $t_k f \in V$ for every $f \in V$ and every $k \in K$ is called K-invariant. In the case that G is $\mathbb{R}^d$ and K is $\mathbb{Z}^d$ the subspace V is the classical shift-invariant space. The structure of these spaces for the context of general LCA groups has been studied in [KT08, CP10]. Independently of their mathematical interest, they are very important in applications. They provide models for many problems in signal and image processing.

In [ACP10a] we study necessary and sufficient conditions in order that an H-invariant space $V \subseteq L^2(G)$ is M-invariant, where $H \subseteq G$ is a countable uniform lattice and M is any closed subgroup of G satisfying that $H \subseteq M \subseteq G$. As a consequence of our results we proved that for each closed subgroup M of G containing the lattice H, there exists an H-invariant space V that is exactly M-invariant. That is, V is not invariant under any other subgroup M' containing M. We also obtained estimates on the support of the Fourier transform of the generators of the H-invariant space, related to its M-invariance.

Bibliography

[Ach03] D. Achlioptas. *Database-friendly random projections: Johnson-Lindenstrauss with binary coins*. J. Comput. System Sci., vol. 66 (2003), no. 4, pp. 671–687. Special issue on PODS 2001 (Santa Barbara, CA). 26, 27

[AM04] P. Agarwal and N. Mustafa. *k-means projective clustering*. Proc. Principles of Database Systems (PODS04) (ACM Press, 2004). 16

[ACHM07] A. Aldroubi, C. Cabrelli, D. Hardin and U. Molter. *Optimal shift invariant spaces and their Parseval frame generators*. Appl. Comput. Harmon. Anal., vol. 23 (2007), no. 2, pp. 273–283. xiv, 16, 19

[ACHKM10] A. Aldroubi, C. Cabrelli, C. Heil, K. Kornelson and U. Molter. *Invariance of a shift-invariant space*. J. Fourier Anal. Appl., vol. 16 (2010), no. 1, pp. 60–75. xviii, 58, 59

[ACM08] A. Aldroubi, C. Cabrelli and U. Molter. *Optimal non-linear models for sparsity and sampling*. J. Fourier Anal. Appl., vol. 14 (2008), no. 5-6, pp. 793–812. xvi, 15, 16, 17, 20, 22, 23

[ACM04] A. Aldroubi, C. Cabrelli and U. M. Molter. *Wavelets on irregular grids with arbitrary dilation matrices and frame atoms for* $L^2(\mathbb{R}^d)$. Appl. Comput. Harmon. Anal., vol. 17 (2004), no. 2, pp. 119–140. 38

[AG01] A. Aldroubi and K. Gröchenig. *Nonuniform sampling and reconstruction in shift-invariant spaces*. SIAM Rev., vol. 43 (2001), no. 4, pp. 585–620 (electronic). xiv, 9

[AT10] A. Aldroubi and R. Tessera. *On the existence of optimal subspace clustering models*. Found. Comput. Math., (2010). Article in Press. http://arxiv.org/abs/1008.4811. 17

[AC09] M. Anastasio and C. Cabrelli. *Sampling in a union of frame generated subspaces*. Sampl. Theory Signal Image Process., vol. 8 (2009), no. 3, pp. 261–286. 22

[ACP10a] M. Anastasio, C. Cabrelli and V. Paternostro. *Extra invariance of shift-invariant spaces on LCA groups*. J. Math. Anal. Appl., vol. 370 (2010), no. 2, pp. 530–537. xix, 78, 79

[AV06] R. I. Arriaga and S. Vempala. *An algorithmic theory of learning: robust concepts and random projection*. In *40th Annual Symposium on Foundations of Computer Science (New York, 1999)*, pp. 616–623 (IEEE Computer Soc., Los Alamitos, CA, 1999). 31

[Bha97] R. Bhatia. *Matrix analysis*, *Graduate Texts in Mathematics*, vol. 169 (Springer-Verlag, New York, 1997). 18, 30

[BD09] T. Blumensath and M. E. Davies. *Sampling theorems for signals from the union of finite-dimensional linear subspaces*. IEEE Trans. Inform. Theory, vol. 55 (2009), no. 4, pp. 1872–1882. xvii, 42

[Bou74] N. Bourbaki. *Éléments de mathématique* (Springer, Berlin, 1974). Topologie générale. Chapitres 5 à 10. [Algebra. Chapters 5–10]. Reprint of the 1974 original. 60, 62

[Bou81] N. Bourbaki. *Éléments de mathématique*, *Lecture Notes in Mathematics*, vol. 864 (Masson, Paris, 1981). Algèbre. Chapitres 4 à 7. [Algebra. Chapters 4–7]. 62

[Bow00] M. Bownik. *The structure of shift-invariant subspaces of* $L^2(\mathbf{R}^n)$. J. Funct. Anal., vol. 177 (2000), no. 2, pp. 282–309. 8, 10, 13, 51, 72

[BG04] M. Bownik and G. Garrigós. *Biorthogonal wavelets, MRA's and shift-invariant spaces*. Studia Math., vol. 160 (2004), no. 3, pp. 231–248. 55

[BK06] M. Bownik and N. Kaiblinger. *Minimal generator sets for finitely generated shift-invariant subspaces of* $L^2(\mathbb{R}^n)$. J. Math. Anal. Appl., vol. 313 (2006), no. 1, pp. 342–352. 74

[CP10] C. Cabrelli and V. Paternostro. *Shift-invariant spaces on LCA groups*. J. Funct. Anal., vol. 258 (2010), no. 6, pp. 2034–2059. 79

[CRT06] E. J. Candès, J. Romberg and T. Tao. *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*. IEEE Trans. Inform. Theory, vol. 52 (2006), no. 2, pp. 489–509. xiv, xvi

[CT06] E. J. Candes and T. Tao. *Near-optimal signal recovery from random projections: universal encoding strategies?* IEEE Trans. Inform. Theory, vol. 52 (2006), no. 12, pp. 5406–5425. xiv, xvi

[CL09] G. Chen and G. Lerman. *Foundations of a multi-way spectral clustering framework for hybrid linear modeling*. Found. Comput. Math., vol. 9 (2009), no. 5, pp. 517–558. xvi, 22

[Chr03] O. Christensen. *An introduction to frames and Riesz bases*. Applied and Numerical Harmonic Analysis (Birkhäuser Boston Inc., Boston, MA, 2003). 2, 7, 8, 14, 50

[Con90] J. B. Conway. *A course in functional analysis*, *Graduate Texts in Mathematics*, vol. 96 (Springer-Verlag, New York, 1990), 2nd ed. 7, 9, 35

[dBDR94] C. de Boor, R. A. DeVore and A. Ron. *Approximation from shift-invariant subspaces of* $L_2(\mathbf{R}^d)$. Trans. Amer. Math. Soc., vol. 341 (1994), no. 2, pp. 787–806. 8, 64

[dBDVR94] C. de Boor, R. A. DeVore and A. Ron. *The structure of finitely generated shift-invariant spaces in* $L_2(\mathbf{R}^d)$. J. Funct. Anal., vol. 119 (1994), no. 1, pp. 37–78. 8, 12, 64

[DRVW06] A. Deshpande, L. Rademacher, S. Vempala and G. Wang. *Matrix approximation and projective clustering via volume sampling*. Theory Comput., vol. 2 (2006), pp. 225–247. 16

[Deu95] F. Deutsch. *The angle between subspaces of a Hilbert space*. In *Approximation theory, wavelets and applications (Maratea, 1994)*, *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, vol. 454, pp. 107–130 (Kluwer Acad. Publ., Dordrecht, 1995). 49, 71

[Don06] D. L. Donoho. *Compressed sensing*. IEEE Trans. Inform. Theory, vol. 52 (2006), no. 4, pp. 1289–1306. xiv, xvi

[EM09] Y. C. Eldar and M. Mishali. *Robust recovery of signals from a structured union of subspaces*. IEEE Trans. Inform. Theory, vol. 55 (2009), no. 11, pp. 5302–5316. xvii, 22

[EV09] E. Elhamifar and R. Vidal. *Sparse subspace clustering*. Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, (2009), pp. 2790–2797. 22

[FW01] H. G. Feichtinger and T. Werther. *Atomic systems for subspaces*. Proceedings SampTA 2001 (L. Zayed, Orlando, FL, 2001). 38

[FB96] P. Feng and Y. Bresler. *Spectrum-blind minimum-rate sampling and reconstruction of multiband signals* (1996) pp. 1688–1691. xv

[Hei11] C. Heil. *A basis theory primer*. Applied and Numerical Harmonic Analysis (Birkhäuser/Springer, New York, 2011), expanded ed. 2

[Hel64] H. Helson. *Lectures on invariant subspaces* (Academic Press, New York, 1964). 8, 10, 58, 74

[HSWW10a] E. Hernández, H. Šikić, G. Weiss and E. Wilson. *On the properties of the integer translates of a square integrable function*. In *Harmonic analysis and partial differential equations*, *Contemp. Math.*, vol. 505, pp. 233–249 (Amer. Math. Soc., Providence, RI, 2010). 8, 14

[HR63] E. Hewitt and K. A. Ross. *Abstract harmonic analysis. Vol. I: Structure of topological groups. Integration theory, group representations*. Die Grundlehren der mathematischen Wissenschaften, Bd. 115 (Academic Press Inc., Publishers, New York, 1963). 60

[HL05] J. A. Hogan and J. D. Lakey. *Non-translation-invariance in principal shift-invariant spaces*. In *Advances in analysis*, pp. 471–485 (World Sci. Publ., Hackensack, NJ, 2005). xvii

[JL84] W. B. Johnson and J. Lindenstrauss. *Extensions of Lipschitz mappings into a Hilbert space*. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, *Contemp. Math.*, vol. 26, pp. 189–206 (Amer. Math. Soc., Providence, RI, 1984). 26

[KT08] R. A. Kamyabi Gol and R. R. Tousi. *The structure of shift invariant spaces on a locally compact abelian group*. J. Math. Anal. Appl., vol. 340 (2008), no. 1, pp. 219–225. 79

[Kan01] K. Kanatani. *Motion segmentation by subspace separation and model selection*. Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on, vol. 2 (2001), pp. 586–591. 22

[KM02] K. Kanatani and C. Matsunaga. *Estimating the number of independent motions for multibody motion segmentation*. 5th Asian Conference on Computer Vision, (2002), pp. 7–12. 22

[KKL06] H. O. Kim, R. Y. Kim and J. K. Lim. *Characterization of the closedness of the sum of two shift-invariant spaces*. J. Math. Anal. Appl., vol. 320 (2006), no. 1, pp. 381–395. 50, 52, 53

[Lan67] H. J. Landau. *Sampling, data transmission, and the nyquist rate*. Proceedings of the IEEE, vol. 55 (1967), no. 10, pp. 1701–1706. 37

[LO04] S. Li and H. Ogawa. *Pseudoframes for subspaces with applications*. J. Fourier Anal. Appl., vol. 10 (2004), no. 4, pp. 409–431. 38

[LD08] Y. M. Lu and M. N. Do. *A theory for sampling signals from a union of subspaces*. IEEE Trans. Signal Process., vol. 56 (2008), no. 6, pp. 2334–2345. xiv, xv, xvi, xvii, xix, 15, 22, 36, 37, 38, 40, 41, 42, 48

[MT82] N. Megiddo and A. Tamir. *On the complexity of locating linear facilities in the plane*. Oper. Res. Lett., vol. 1 (1981/82), no. 5, pp. 194–197. 23

[RS95] A. Ron and Z. Shen. *Frames and stable bases for shift-invariant subspaces of $L_2(\mathbf{R}^d)$*. Canad. J. Math., vol. 47 (1995), no. 5, pp. 1051–1094. 8, 53, 64

[Rud91] W. Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics (McGraw-Hill Inc., New York, 1991), 2nd ed. 42

[Sar08] T. Sarlós. *Improved approximation algorithms for large matrices via random projection*. Foundations of Computer Science, 2006. FOCS '06. 47th Annual IEEE Symposium on, (2006), pp. 143–152. 31

[Sch07] E. Schmidt. *Zur Theorie der linearen und nicht linearen Integralgleichungen Zweite Abhandlung*. Math. Ann., vol. 64 (1907), no. 2, pp. 161–174. 18

[Sri64] T. P. Srinivasan. *Doubly invariant subspaces*. Pacific J. Math., vol. 14 (1964), pp. 701–707. 58

[Sun05] W. Sun. *Sampling theorems for multivariate shift invariant subspaces*. Sampl. Theory Signal Image Process., vol. 4 (2005), no. 1, pp. 73–98. xiv, 9, 10

[TB97] L. N. Trefethen and D. Bau, III. *Numerical linear algebra* (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997). 23

[VMB02] M. Vetterli, P. Marziliano and T. Blu. *Sampling signals with finite rate of innovation*. IEEE Trans. Signal Process., vol. 50 (2002), no. 6, pp. 1417–1428. xv, xvi

[VMS05] R. Vidal, Y. Ma and S. Sastry. *Generalized principal component analysis (gpca)*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27 (2005), no. 12, pp. 1945–1959. 22

[VN50] J. von Neumann. *Functional Operators. II. The Geometry of Orthogonal Spaces*. Annals of Mathematics Studies, no. 22 (Princeton University Press, Princeton, N. J., 1950). 54

[Wal92] G. G. Walter. *A sampling theorem for wavelet subspaces*. IEEE Trans. Inform. Theory, vol. 38 (1992), no. 2, part 2, pp. 881–884. xiv, 9

[ZS99] X. Zhou and W. Sun. *On the sampling theorem for wavelet subspaces*. J. Fourier Anal. Appl., vol. 5 (1999), no. 4, pp. 347–354. xiv, 9, 10

Index