



Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

journal homepage: www.elsevier.com/locate/jmaa

Regular Articles

Learning optimal smooth invariant subspaces for data approximation

D. Barbieri^{a,*}, C. Cabrelli^b, E. Hernández^a, U. Molter^b^a Universidad Autónoma de Madrid, 28049 Madrid, Spain^b Universidad de Buenos Aires and IMAS-CONICET, Consejo Nacional de Investigaciones Científicas y Técnicas, 1428 Buenos Aires, Argentina

ARTICLE INFO

Article history:

Received 20 November 2023

Available online 25 March 2024

Submitted by M.J. Carro

Keywords:

Invariant subspaces

Data approximation

Paley-Wiener spaces

Optimal subspaces

ABSTRACT

In this article, we consider the problem of approximating a finite set of data (usually huge in applications) by invariant subspaces generated by a small set of smooth functions. The invariance is either by translations under a full-rank lattice or through the action of crystallographic groups. Smoothness is ensured by stipulating that the generators belong to a Paley-Wiener space, which is selected in an optimal way based on the characteristics of the given data. To complete our investigation, we analyze the fundamental role played by the lattice in the process of approximation.

© 2024 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In various signal and image processing applications, we frequently encounter a specific class of signals generated by the measurement of physical phenomena, such as electrocardiograms, audio signals, radar data, medical images, and internet traffic, to name a few. These signals within each class stem from a common source, leading us to expect that they can be effectively characterized by a concise set of parameters.

Furthermore, certain classes of signals exhibit inherent invariance properties, such as translation or rotation invariance, which are particularly prominent in the case of natural images. Consequently, when seeking a simple model or subspace that closely approximates a specific class of signals, it is natural to demand that this subspace also embodies these inherent invariances.

The endeavor to learn an optimal subspace with specific properties from a dataset has been explored in various contexts. This quest was initially addressed in [1], where it was assumed that the class of signals belonged to the space $L^2(\mathbb{R}^d)$. In this early study, an optimal ‘small’ subspace was constructed, which was

* Corresponding author.

E-mail addresses: davide.barbieri@uam.es (D. Barbieri), carlos.cabrelli@gmail.com (C. Cabrelli), eugenio.hernandez@uam.es (E. Hernández), umolter@dm.uba.ar (U. Molter).

required to be invariant under translations along a lattice. The measure of “closeness” was determined by a functional, as we will elaborate on shortly.

Subsequently, in [2,5], the scope was broadened to encompass cases where the approximation subspaces also possessed extra invariance, specifically involving translations on a finer lattice. This approximation problem was also considered in general Banach spaces in [7]. More recently, as demonstrated in [3] a more general scenario was considered, encompassing invariance properties like translations and rotations, such as those associated with crystallographic groups, where wavelet sets were recently considered [11].

In each of these cases, the existence of an optimal subspace was rigorously established, and a formula for selecting the generators was derived. Notably, the generators were demonstrated to belong to the invariant subspace generated by the data. However, these methods did not guarantee that the generators possessed special properties, such as compact support or smoothness.

The aim of the present work is to focus on the smoothness of the generators. The idea is to develop a new strategy that allows one to obtain, in all the considered cases, optimal subspaces with smooth generators. Here the smoothness is measured as a membership to a general Paley-Wiener subspace that will be chosen carefully. However, there is a price to pay in terms of the error for obtaining nice generators. In what follows, we will give a detailed description of the problem and provide the definitions needed.

Given a set of data $\mathcal{F} = \{f_1, \dots, f_m\}$ in a Hilbert space \mathcal{H} , one wishes to minimize the error

$$\mathcal{E}(\mathcal{F}, S) = \sum_{i=1}^m \|f_i - \mathbb{P}_S f_i\|^2, \quad (1)$$

over all subspaces $S \subset \mathcal{H}$ in a specific class, where \mathbb{P}_S stands for the orthogonal projection onto S . In this section, Λ is a full rank lattice of \mathbb{R}^d .

1.1. Case I: Shift-invariant subspaces (SIS)

In [1] the following problem was solved: Given a finite set of data $\mathcal{F} = \{f_1, \dots, f_m\}$ in $L^2(\mathbb{R}^d)$ and a positive integer ℓ , there always exist ℓ functions $\Phi = \{\varphi_1, \dots, \varphi_\ell\}$, called *generators* of the subspace

$$S_\Lambda(\Phi) = \overline{\text{span}}\{t_\lambda \varphi_i : \lambda \in \Lambda, i = 1, \dots, \ell\},$$

such that $S_\Lambda(\Phi)$ is optimal in the sense that it minimizes the error

$$\mathcal{E}(\mathcal{F}, S_\Lambda(\Phi)) = \sum_{i=1}^m \|f_i - \mathbb{P}_{S_\Lambda(\Phi)} f_i\|^2$$

over all other sets Ψ in $L^2(\mathbb{R}^d)$ of ℓ generators. Here $t_\lambda \varphi(x) = \varphi(x - \lambda)$ is the translation of φ by λ . Furthermore, the optimal generators can be chosen within $S_\Lambda(\mathcal{F})$.

1.2. Case II: SIS with extra invariance

There are shift-invariant spaces $S_\Lambda(\Phi)$ in $L^2(\mathbb{R}^d)$ that are also invariant under translations by a closed subgroup M of \mathbb{R}^d containing Λ . Denote by \mathcal{V}_M the collection of all these subspaces that have extra-invariance by M . The result in Case I was extended to this case. It was proved in [5] that given a finite set of data $\mathcal{F} \subset L^2(\mathbb{R}^d)$ there exists a subspace $S_\Lambda(\Phi) \in \mathcal{V}_M$ with $\Phi = \{\phi_1, \dots, \phi_\ell\}$ such that $\mathcal{E}(\mathcal{F}, S_\Lambda(\Phi)) \leq \mathcal{E}(\mathcal{F}, S_\Lambda(\Psi))$ for all $S_\Lambda(\Psi) \in \mathcal{V}_M$ with $\Psi = \{\psi_1, \dots, \psi_\ell\} \subset L^2(\mathbb{R}^d)$.

1.3. Case III: Γ -invariant subspaces

In order to allow for rotations or additional symmetries other than translations, the problem of approximation by subspaces that are invariant by non commutative groups is considered in [3]. There, the groups are of the form $\Gamma = \Lambda \rtimes G$, where G is a finite group that acts faithfully on \mathbb{R}^d by continuous automorphisms that preserve Λ , hence including the crystallographic groups that split.

The result in Case I was extended to this case in order to find the best Γ -invariant space of length ℓ that approximates the given data. This means that we find $\Phi = \{\phi_1, \dots, \phi_\ell\}$ such that

$$\mathcal{E}(\mathcal{F}, S_\Gamma(\Phi)) = \sum_{i=1}^m \|f_i - \mathbb{P}_{S_\Gamma(\Phi)} f_i\|^2 \leq \mathcal{E}(\mathcal{F}, S_\Gamma(\Psi)),$$

for all sets Ψ in $L^2(\mathbb{R}^d)$ of ℓ functions. Here Φ is the set of *generators* of

$$S_\Gamma(\Phi) := \overline{\text{span}}\{t_\lambda R_g \phi_i : g \in G, \lambda \in \Lambda, i = 1, \dots, \ell\},$$

where $R_g(f)(x) = f(g^{-1}x)$ denotes the action of $g \in G$ on $f \in L^2(\mathbb{R}^d)$. In this case, again, the optimal generators can be chosen in $S_\Gamma(\mathcal{F})$.

In this paper we will address the question if it is possible to obtain similar results, but requiring that the approximating subspace consists of smooth functions. Precisely, for each one of the cases I, II, and III above, we ask if it is possible to find the optimal generators $\Phi = \{\phi_1, \dots, \phi_\ell\}$ among those belonging to a Paley-Wiener space $PW_\Omega := \{f \in L^2(\mathbb{R}^d) : \hat{f} \equiv 0 \text{ on } \mathbb{R}^d \setminus \Omega\}$ for a fixed bounded set Ω of prescribed measure. Throughout this paper, \hat{f} denotes de Fourier transform of $f \in L^1(\mathbb{R}^d)$ defined by

$$\hat{f}(\omega) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i \langle x, \omega \rangle} dx,$$

which extends to an isometric isomorphism of $L^2(\mathbb{R}^d)$.

For a given $\Omega \subset \mathbb{R}^d$, for each one of the cases I, II, and III, we prove that optimization can be performed as follows: Project the data \mathcal{F} onto $V = PW_\Omega$ to obtain the data $\mathbb{P}_V \mathcal{F} = \{\mathbb{P}_V f_1, \dots, \mathbb{P}_V f_m\}$, where \mathbb{P}_V denotes the orthogonal projection of $L^2(\mathbb{R}^d)$ onto V , and then minimize $\mathcal{E}(\mathbb{P}_V \mathcal{F}, S)$ over all $S \subseteq V$. It turns out, and this is a general fact in Hilbert spaces, that if S and V are subspaces of a Hilbert space \mathcal{H} and $S \subseteq V$, the error of approximating \mathcal{F} by S coincides with the error of approximating $\mathbb{P}_V \mathcal{F}$ by S plus the error of approximating \mathcal{F} by V . As a consequence, the subspaces that minimize $\mathcal{E}(\mathcal{F}, S)$ over all $S \subseteq V$ coincide with the subspaces that minimize $\mathcal{E}(\mathbb{P}_V \mathcal{F}, S)$ over the same collection. (See Theorem 1 in Section 2.) Then we prove that one of the subspaces that minimize the error can always be chosen inside the invariant subspace generated by the data \mathcal{F} . (See Proposition 4 in Section 2.)

The results outlined in the two previous paragraphs allow us to show the existence of invariant subspaces contained in $V = PW_\Omega$ that minimize the error of approximation of the data \mathcal{F} . (See Theorems 7, 9, and 10 in Section 3.) Observe that the elements of such a minimizing subspace are analytic functions in \mathbb{R}^d since they belong to the Paley-Wiener space PW_Ω .

One could design another strategy to solve this problem. Namely, find the invariant subspace that minimizes the error of approximating \mathcal{F} , and then project this subspace onto $V = PW_\Omega$. It turns out that this method does not always produce a subspace that minimizes the error, as demonstrated by Example 1 in Section 3.

We further show how, for each set of data, the set Ω that define the Paley-Wiener space, can be chosen in an optimal way. In particular, for non-abelian semidirect product groups $\Gamma = \Lambda \rtimes G$, the optimal Ω for the approximation must be invariant under the action of G . See Section 4.

We then analyze the role of the lattices in our approximation problem. Let $A \in GL_n(\mathbb{R})$ and denote by $|A|$ its determinant. We show that the best error of approximating a set of data \mathcal{F} by AA -invariant subspaces of length at most ℓ coincides with the best error of approximation of the dilated data $D_A\mathcal{F} = \{D_A f_1, \dots, D_A f_m\}$ by Λ -invariant subspaces of length at most ℓ . (See Proposition 17 in Section 5.) Here $D_A f(x) = |A|^{\frac{1}{2}} f(Ax)$.

We next compare the best error of approximation of the same set of data \mathcal{F} using different lattices. Denote by $\mathcal{E}^*(\mathcal{F}, \Lambda, \ell)$ the best error of approximation of \mathcal{F} by Λ -invariant subspaces of \mathbb{R}^d of length less than or equal to ℓ . We prove that $\mathcal{E}^*(\mathcal{F}, \Lambda, \ell) \leq \mathcal{E}^*(\mathcal{F}, \mathbb{Z}^d, \ell)$ when Λ is a lattice of \mathbb{R}^d containing \mathbb{Z}^d . We show that this can be an equality even if Λ strictly contains \mathbb{Z}^d , while the reverse inequality could also be true for a lattice Λ that is finer than \mathbb{Z}^d but does not contain it. (See Examples 2, 3 and 4, Section 6.)

We finally analyze how the length of $S_\Lambda(\mathcal{F})$ affects the approximation error. We exhibit a set of data \mathcal{F} and two lattices Λ_1 and Λ_2 such that the length of $S_{\Lambda_1}(\mathcal{F})$ is smaller than the length of $S_{\Lambda_2}(\mathcal{F})$ and also the approximation error with ℓ functions is smaller for Λ_1 than for Λ_2 . (See Example 5 in Section 6.) This example induces to think that the smaller the length of $S_\Lambda(\mathcal{F})$, the smaller the best error of approximation will be. However, Example 6 disproves that.

Acknowledgments. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No 777822. In addition, D. Barbieri and E. Hernández were supported by Grants PID2019-105599GB-I00, PID2022-142202NB-I00 / AEI / 10.13039/501100011033. The research of C. Cabrelli and U. Molter is partially supported by Grants PICT 2018-003399 and 2022-4875 (ANPCyT), PIP 11220150100355 and 202287/22 (CONICET), UBACyT 20020170100430BA and 20020220300154BA (UBA).

2. Projection theorems

By a subspace V of a Hilbert space, we will mean a *closed subspace*, and by \mathbb{P}_V we denote the orthogonal projection onto V .

We start with the following theorem.

Theorem 1. *Let \mathcal{H} be a Hilbert space, $\mathcal{F} = \{f_1, \dots, f_m\} \subset \mathcal{H}$ and let $V \subseteq \mathcal{H}$ be a fixed subspace. Given any subspace $S \subseteq V$ we have*

$$\mathcal{E}(\mathcal{F}, S) = \mathcal{E}(\mathbb{P}_V \mathcal{F}, S) + \mathcal{E}(\mathcal{F}, V). \quad (2)$$

As a consequence, for all subspaces $S, S' \subset V$,

$$\mathcal{E}(\mathcal{F}, S) \leq \mathcal{E}(\mathcal{F}, S') \iff \mathcal{E}(\mathbb{P}_V \mathcal{F}, S) \leq \mathcal{E}(\mathbb{P}_V \mathcal{F}, S').$$

Proof. Suppose $S \subset V$. Then $\mathbb{P}_V \mathbb{P}_S = \mathbb{P}_S = \mathbb{P}_S \mathbb{P}_V$ and $\mathbb{P}_{V^\perp} \mathbb{P}_S = 0$. Hence, for every $f \in \mathcal{H}$

$$\begin{aligned} \|f - \mathbb{P}_S f\|^2 &= \|\mathbb{P}_V(f - \mathbb{P}_S f)\|^2 + \|\mathbb{P}_{V^\perp}(f - \mathbb{P}_S f)\|^2 \\ &= \|\mathbb{P}_V f - \mathbb{P}_S \mathbb{P}_V f\|^2 + \|\mathbb{P}_{V^\perp} f\|^2. \end{aligned}$$

Thus, $\mathcal{E}(\mathcal{F}, S) = \sum_{i=1}^m \|f_i - \mathbb{P}_V f_i\|^2 + \mathcal{E}(\mathbb{P}_V \mathcal{F}, S)$, which implies (2). The last part follow immediately from (2). \square

We will need sometimes the following elementary Lemma about a general fact on Hilbert spaces.

Lemma 2. Let V be a subspace of a Hilbert space \mathcal{H} and let U be a unitary operator in \mathcal{H} . Then

- i) $U\mathbb{P}_V = \mathbb{P}_{U(V)}U$
- ii) $\mathbb{P}_V U = U\mathbb{P}_V \iff U(V) = V$.

Proof. i) The operator $T := U\mathbb{P}_V U^*$ satisfies $T^2 = T$ and $T^* = T$. Therefore, T is an orthogonal projection. The result follows since the range of T is $U(V)$.

ii) Using i) and that U is an isometry we obtain:

$$\mathbb{P}_V U = U\mathbb{P}_V \iff \mathbb{P}_V U = \mathbb{P}_{U(V)}U \iff \mathbb{P}_V = \mathbb{P}_{U(V)} \iff V = U(V). \quad \square$$

If \mathcal{H} is infinite dimensional, and \mathcal{F} is a finite set of data, in [1] the following proposition is proved.

Proposition 3. Let $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq \mathcal{H}$. For an arbitrary subspace $M \subseteq \mathcal{H}$, there always exists a subspace $W \subseteq \text{span}\{\mathcal{F}\}$ such that $\dim(W) \leq \dim(M)$, and

$$\mathcal{E}(\mathcal{F}, W) \leq \mathcal{E}(\mathcal{F}, M). \quad (3)$$

We need here an extension of this property to the case that we want to approximate our data with shift, or Γ -invariant subspaces. Here $\Gamma = \Lambda \rtimes G$ is the semidirect product of a full rank lattice Λ in \mathbb{R}^d and a discrete and countable group G that acts on \mathbb{R}^d by continuous invertible automorphisms. In addition, we also assume that $g\Lambda = \Lambda$ for all $g \in G$, which implies that the Lebesgue measure of \mathbb{R}^d is invariant under the action of G . Since shift-invariant spaces are in particular Γ -invariant for Γ the group of translations on Λ , we state (and prove) the proposition for that case.

Proposition 4. Let Γ be $\Gamma = \Lambda \rtimes G$, and let $\mathcal{C}_\ell(\mathbb{R}^d)$ be the class of Γ -invariant subspaces of $L^2(\mathbb{R}^d)$ of length at most ℓ . Let $\mathcal{F} = \{f_1, \dots, f_m\} \subseteq L^2(\mathbb{R}^d)$. For an arbitrary subspace $M \in \mathcal{C}_\ell(\mathbb{R}^d)$, there always exists a subspace $W \subseteq S_\Gamma(\mathcal{F})$ and $W \in \mathcal{C}_\ell(\mathbb{R}^d)$ such that

$$\mathcal{E}(\mathcal{F}, W) \leq \mathcal{E}(\mathcal{F}, M). \quad (4)$$

Remark 5. The above Proposition shows that the Γ -invariant subspace of $L^2(\mathbb{R}^d)$ generated by at most ℓ elements that best approximates the data \mathcal{F} , can be taken to be generated by elements of $S_\Gamma(\mathcal{F})$.

Proof of Proposition 4. Let $X = S_\Gamma(\mathcal{F}) := \overline{\text{span}}\{T_k R_g f : k \in \Lambda, g \in G, f \in \mathcal{F}\}$. Let $\Psi = \{\psi_1, \dots, \psi_r\}$, ($r \leq \ell$) be such that $M = S_\Gamma(\Psi)$. Define $\varphi_i = \mathbb{P}_X \psi_i$, $i = 1 \dots r$, and $\Phi := \{\varphi_1, \dots, \varphi_r\}$. For $j = 1, \dots, m$

$$\|f_j - \mathbb{P}_{S_\Gamma(\Phi)} f_j\| = \inf_{g \in S_\Gamma(\Phi)} \|f_j - g\|. \quad (5)$$

Also note that since X is Γ -invariant, $\mathbb{P}_X(T_k R_g) = (T_k R_g)\mathbb{P}_X$, $\forall (k, g) \in \Gamma$ as a consequence of Lemma 2, and hence $\mathbb{P}_X \mathbb{P}_{S_\Gamma(\Psi)} f_j \in S_\Gamma(\Phi) \forall j = 1, \dots, m$.

Therefore, since $f_j \in \mathcal{F}$, by equation (5),

$$\|f_j - \mathbb{P}_{S_\Gamma(\Phi)} f_j\| \leq \|f_j - \mathbb{P}_X \mathbb{P}_{S_\Gamma(\Psi)} f_j\| = \|\mathbb{P}_X f_j - \mathbb{P}_X \mathbb{P}_{S_\Gamma(\Psi)} f_j\| \leq \|f_j - \mathbb{P}_{S_\Gamma(\Psi)} f_j\|.$$

Thus, with $W = S_\Gamma(\Phi) \in \mathcal{C}_\ell$,

$$\mathcal{E}(\mathcal{F}, W) = \mathcal{E}(\mathcal{F}, S_\Gamma(\Phi)) = \sum_{j=1}^m \|f_j - \mathbb{P}_{S_\Gamma(\Phi)} f_j\|^2$$

$$\leq \sum_{j=1}^m \|f_j - \mathbb{P}_{S_\Gamma(\Psi)} f_j\|^2 = \mathcal{E}(\mathcal{F}, S_\Gamma(\Psi)) = \mathcal{E}(\mathcal{F}, M). \quad \square$$

We conclude this section by observing that enlarging the search space for the minimum with directions that are orthogonal to the data does not reduce the error of approximation.

Proposition 6. *Let S, V , and W be subspaces of \mathcal{H} . Suppose $V \perp W$, $S \perp W$ and let $S' = S \oplus W$. Then*

$$\mathcal{E}(\mathbb{P}_V \mathcal{F}, S) = \mathcal{E}(\mathbb{P}_V \mathcal{F}, S').$$

Proof.

$$\begin{aligned} \mathcal{E}(\mathbb{P}_V \mathcal{F}, S') &= \sum_{i=1}^m \|\mathbb{P}_V f_i - \mathbb{P}_{S'} \mathbb{P}_V f_i\|^2 = \sum_{i=1}^m \|\mathbb{P}_V f_i - (\mathbb{P}_S + \mathbb{P}_W) \mathbb{P}_V f_i\|^2 \\ &= \sum_{i=1}^m \|\mathbb{P}_V f_i - \mathbb{P}_S \mathbb{P}_V f_i\|^2 = \mathcal{E}(\mathbb{P}_V \mathcal{F}, S). \quad \square \end{aligned}$$

3. Smooth generators

3.1. Approximation by shift-invariant spaces

In this subsection we will consider the case in which we want to approximate our data with a shift-invariant space generated by “smooth” functions. For this, let $\Omega \subset \mathbb{R}^d$ be a bounded measurable set. We will later see how to choose a convenient set Ω depending of the data \mathcal{F} . Throughout this section we will fix a full rank lattice Λ of \mathbb{R}^d and the shift-invariant spaces are invariant along translations by Λ . Thus we will omit the subscript Λ , and simply write $S(\Phi)$ instead of $S_\Lambda(\Phi)$.

Let $V = PW_\Omega$, be the generalized Paley-Wiener space associated to Ω . i.e. $PW_\Omega = \{f \in L^2(\mathbb{R}^d) : \hat{f} \equiv 0 \text{ on } \mathbb{R}^d \setminus \Omega\}$. For a positive integer ℓ , we denote by $\mathcal{C}_\ell(\mathbb{R}^d)$ the class of *shift-invariant subspaces* of length at most ℓ , and by $\mathcal{C}_\ell(\Omega)$ the elements of $\mathcal{C}_\ell(\mathbb{R}^d)$ included in V . We recall that the *length* of a shift-invariant subspace is the minimal number of generators of the subspace.

In [1] the existence of a best shift-invariant space for given data was proven. If we want to apply the solution of that paper, we cannot guarantee that the generators of the optimal subspace are smooth, i.e. that they belong to V . However, using Theorem 1, we will see that, if instead we solve the theorem of approximation for the data $\mathbb{P}_V \mathcal{F}$, instead of \mathcal{F} , this yields the solution sought after. That is, the optimal subspace for the data $\mathbb{P}_V \mathcal{F}$ is included in V and minimizes the error.

We have the following Theorem.

Theorem 7. *Given data $\mathcal{F} = \{f_1, \dots, f_m\} \subset L^2(\mathbb{R}^d)$, and $\Omega \subset \mathbb{R}^d$, there exists a shift-invariant space $S(\Phi) \in \mathcal{C}_\ell(\Omega)$ such that $\mathcal{E}(\mathcal{F}, S(\Phi)) \leq \mathcal{E}(\mathcal{F}, S(\Psi))$ for all sets $\Psi = \{\psi_1, \dots, \psi_\ell\} \subset PW_\Omega$.*

Proof of Theorem 7. Since any Paley Wiener space is translation invariant for any translation, part ii) of Lemma 2 shows that \mathbb{P}_V commutes with any translation, which in particular implies that the orthogonal projection of any subspace $S(\Psi)$ of $\mathcal{C}_\ell(\mathbb{R}^d)$ onto V is in $\mathcal{C}_\ell(\Omega)$ and $\mathbb{P}_V S(\Psi) = S(\mathbb{P}_V(\Psi))$.

Let now $S^* = S^*(\Phi)$ be an optimal subspace that approximates the projected data $\mathbb{P}_V \mathcal{F}$ with elements of $\mathcal{C}_\ell(\mathbb{R}^d)$. By Proposition 4 we can choose $S^* \subseteq S(\mathbb{P}_V \mathcal{F}) = \mathbb{P}_V S(\mathcal{F}) \subseteq V$. Since S^* minimizes $\mathcal{E}(\mathbb{P}_V \mathcal{F}, S)$ over all $S \in \mathcal{C}_\ell(\mathbb{R}^d)$, then in particular S^* minimizes $\mathcal{E}(\mathbb{P}_V \mathcal{F}, S)$ over all $S \in \mathcal{C}_\ell(\Omega)$. That is,

$$\mathcal{E}(\mathbb{P}_V \mathcal{F}, S^*) = \min_{S \in \mathcal{C}_\ell(\mathbb{R}^d)} \mathcal{E}(\mathbb{P}_V \mathcal{F}, S) = \min_{S \in \mathcal{C}_\ell(\Omega)} \mathcal{E}(\mathbb{P}_V \mathcal{F}, S).$$

Therefore, by Theorem 1,

$$\begin{aligned}\mathcal{E}(\mathcal{F}, S^*) &= \mathcal{E}(\mathbb{P}_V \mathcal{F}, S^*) + \mathcal{E}(\mathcal{F}, V) \\ &= \min_{S \in \mathcal{C}_\ell(\Omega)} \mathcal{E}(\mathbb{P}_V \mathcal{F}, S) + \mathcal{E}(\mathcal{F}, V) = \min_{S \in \mathcal{C}_\ell(\Omega)} \mathcal{E}(\mathcal{F}, S),\end{aligned}$$

which proves the result. \square

Remark 8. For a class of subspaces \mathcal{C} of $L^2(\mathbb{R}^d)$ denote by $\operatorname{argmin}_{S \in \mathcal{C}} \mathcal{E}(\mathcal{F}, S)$ the set of all subspaces $A \in \mathcal{C}$ such that $\mathcal{E}(\mathcal{F}, A) = \min_{S \in \mathcal{C}} \mathcal{E}(\mathcal{F}, S)$. Then we show the following:

$$\left(\operatorname{argmin}_{S \in \mathcal{C}_\ell(\mathbb{R}^d)} \mathcal{E}(\mathbb{P}_V \mathcal{F}, S) \right) \cap \mathcal{C}_\ell(\Omega) = \operatorname{argmin}_{S \in \mathcal{C}_\ell(\Omega)} \mathcal{E}(\mathbb{P}_V \mathcal{F}, S) = \operatorname{argmin}_{S \in \mathcal{C}_\ell(\Omega)} \mathcal{E}(\mathcal{F}, S). \quad (6)$$

The second identity holds by the second part of Theorem 1. For the first identity, we first prove the inclusion \subseteq . If A is in the left hand set of (6), then $A \in \mathcal{C}_\ell(\Omega)$ and

$$\mathcal{E}(\mathbb{P}_V \mathcal{F}, A) = \min_{S \in \mathcal{C}_\ell(\mathbb{R}^d)} \mathcal{E}(\mathbb{P}_V \mathcal{F}, S) \leq \min_{S \in \mathcal{C}_\ell(\Omega)} \mathcal{E}(\mathbb{P}_V \mathcal{F}, S) \leq \mathcal{E}(\mathbb{P}_V \mathcal{F}, A).$$

This shows that $A \in \operatorname{argmin}_{S \in \mathcal{C}_\ell(\Omega)} \mathcal{E}(\mathbb{P}_V \mathcal{F}, S)$.

Let us now prove the opposite inclusion, that is, let us prove that if $S^* \in \operatorname{argmin}_{S \in \mathcal{C}_\ell(\Omega)} \mathcal{E}(\mathbb{P}_V \mathcal{F}, S)$ then $S^* \in \left(\operatorname{argmin}_{S \in \mathcal{C}_\ell(\mathbb{R}^d)} \mathcal{E}(\mathbb{P}_V \mathcal{F}, S) \right) \cap \mathcal{C}_\ell(\Omega)$. To this end assume, by contradiction, that there is an $\tilde{S} \in \mathcal{C}_\ell(\mathbb{R}^d)$ such that $\mathcal{E}(\mathbb{P}_V \mathcal{F}, \tilde{S}) < \mathcal{E}(\mathbb{P}_V \mathcal{F}, S^*)$. Thus, by Proposition 4, there is a $W \in \mathcal{C}_\ell(\Omega)$ such that $\mathcal{E}(\mathbb{P}_V \mathcal{F}, W) \leq \mathcal{E}(\mathbb{P}_V \mathcal{F}, \tilde{S}) < \mathcal{E}(\mathbb{P}_V \mathcal{F}, S^*)$, but by definition of S^* we have that $\mathcal{E}(\mathbb{P}_V \mathcal{F}, S^*) \leq \mathcal{E}(\mathbb{P}_V \mathcal{F}, W)$.

Note that the first identity in (6) would not hold in general if we remove from the left hand side the intersection with $\mathcal{C}_\ell(\Omega)$, as it can be easily seen using Proposition 6.

At this point one can ask what would happen if we find an optimal subspace $S^1 \in \mathcal{C}_\ell(\mathbb{R}^d)$ for approximating the data \mathcal{F} with subspaces of $\mathcal{C}_\ell(\mathbb{R}^d)$ and then project this subspace onto V to obtain a subspace $S_V^1 = \mathbb{P}_V(S^1)$.

Let $S^* = S(\Phi)$ be an optimal subspace for approximating the modified data $\mathbb{P}_V \mathcal{F}$ with subspaces that belong to $\mathcal{C}_\ell(\mathbb{R}^d)$. By Proposition 4, we can choose $S^* \subseteq S(\mathbb{P}_V \mathcal{F}) = \mathbb{P}_V S(\mathcal{F}) \subseteq V$. Recall that $V = PW_\Omega$. Since $S^* \subseteq V$, by Theorem 1,

$$\mathcal{E}(\mathcal{F}, S^*) = \mathcal{E}(\mathbb{P}_V \mathcal{F}, S^*) + \mathcal{E}(\mathcal{F}, V).$$

Since $S_V^1 \subseteq V$, we can see that

$$\mathcal{E}(\mathcal{F}, S^*) = \mathcal{E}(\mathbb{P}_V \mathcal{F}, S^*) + \mathcal{E}(\mathcal{F}, V) \leq \mathcal{E}(\mathbb{P}_V \mathcal{F}, S_V^1) + \mathcal{E}(\mathcal{F}, V) = \mathcal{E}(\mathcal{F}, S_V^1). \quad (7)$$

In some cases, we could have equality in (7). For example, this happens if all the data \mathcal{F} is in V or in V^\perp . On the other hand, we will show in the next example that the inequality in (7) may be strict.

Example 1.¹ Let $\mathcal{F} = \{f_1, f_2\}$ with

$$\hat{f}_1 = \chi_{[-1,0)} + 2\chi_{[1,2)}, \quad \hat{f}_2 = -\chi_{[-1,0)} + 2\chi_{[1,2)}.$$

¹ We thank Inés Armendariz for the idea behind this example.

Let $V = PW_{[-1,1]}$. We compute first the right hand side of (7) with \mathbb{Z} -invariant subspaces of length 1 in $L^2(\mathbb{R})$. To find $S^1 \in \mathcal{C}_1(\mathbb{R})$ we need to find $\varphi_0 \in L^2(\mathbb{R})$ such that

$$\mathcal{E}(\mathcal{F}, S(\varphi_0)) = \min_{\varphi \in L^2(\mathbb{R})} \mathcal{E}(\mathcal{F}, S(\varphi)).$$

By Plancherel Theorem,

$$\mathcal{E}(\mathcal{F}, S(\varphi)) = \|\widehat{f_1} - \widehat{\mathbb{P}_{S(\varphi)} f_1}\|^2 + \|\widehat{f_2} - \widehat{\mathbb{P}_{S(\varphi)} f_2}\|^2. \quad (8)$$

It can be easily seen that, for φ to minimize (8), $\text{supp } \widehat{\varphi}$ must be contained in $[-1, 0) \cup [1, 2)$. Among such functions, the smallest error is attained for φ_0 given by $\widehat{\varphi_0} = \chi_{[1,2)}$. Observe that $\widehat{\mathbb{P}_{S(\varphi_0)} f_1} = 2\chi_{[1,2)} = \widehat{\mathbb{P}_{S(\varphi_0)} f_2}$. Therefore, using (8) we conclude

$$\mathcal{E}(\mathcal{F}, S(\varphi_0)) = \|\chi_{[-1,0)}\|^2 + \|\chi_{[-1,0)}\|^2 = 2.$$

We have shown that $S^1 = S(\varphi_0)$. Hence, $S_V^1 = \mathbb{P}_V(S^1) = \mathbb{P}_V(S(\varphi_0)) = \{0\}$, since $\text{supp } \widehat{\varphi_0} \cap [-1, 0) = \emptyset$. In this case, the right hand side of (7) becomes

$$\mathcal{E}(\mathcal{F}, S_V^1) = \|f_1\|^2 + \|f_2\|^2 = 10.$$

We compute now the left hand side of (7). First find $S^* = S(\varphi_1)$ such that

$$\mathcal{E}(\mathbb{P}_V \mathcal{F}, S^*) = \min_{S(\varphi) \subset V} \mathcal{E}(\mathbb{P}_V \mathcal{F}, S(\varphi)).$$

By taking $\widehat{\varphi_1} = \chi_{[-1,0)}$ (observe that $\varphi_1 \in V = PW_{[-1,1]}$),

$$\begin{aligned} \mathcal{E}(\mathbb{P}_V \mathcal{F}, S(\varphi_1)) &= \|\widehat{\mathbb{P}_V f_1} - (\mathbb{P}_{S(\varphi_1)} \mathbb{P}_V f_1)^\wedge\|^2 + \|\widehat{\mathbb{P}_V f_2} - (\mathbb{P}_{S(\varphi_1)} \mathbb{P}_V f_2)^\wedge\|^2 \\ &= \|\chi_{[-1,0)} - \chi_{[-1,0)}\|^2 + \|\chi_{[-1,0)} - \chi_{[-1,0)}\|^2 = 0. \end{aligned}$$

This shows that $S^* = S(\varphi_1)$. By Theorem 1, the left hand side of (7) becomes

$$\begin{aligned} \mathcal{E}(\mathcal{F}, S^*) &= \mathcal{E}(\mathcal{F}, V) = \|\widehat{f_1} - \widehat{\mathbb{P}_V f_1}\|^2 + \|\widehat{f_2} - \widehat{\mathbb{P}_V f_2}\|^2 \\ &= \|2\chi_{[1,2)}\|^2 + \|2\chi_{[1,2)}\|^2 = 8. \end{aligned}$$

3.2. Approximation by shift-invariant spaces with extra invariance

There are shift-invariant spaces $S_\Lambda(\Psi)$ in $L^2(\mathbb{R}^d)$ that are also invariant under translations by the elements of a closed subgroup M of \mathbb{R}^d containing Λ . In this case we say that $S_\Lambda(\Psi)$ has extra-invariance. Denote by \mathcal{V}_M the set of such subspaces of $L^2(\mathbb{R}^d)$.

Since in Theorem 3.1 of [5] the existence of “best shift-invariant spaces that have extra-invariance” has been proven, if we want to require that the approximating vectors are in $\mathcal{C}_\ell(\Omega)$ we can apply the same proof of the previous theorem, since again, $\mathbb{P}_V S_\Lambda(\Phi) = S_\Lambda(\mathbb{P}_V(\Phi))$. Therefore we have the following theorem.

Theorem 9. *Let M be closed subgroup of \mathbb{R}^d containing Λ . Given data $\mathcal{F} = \{f_1, \dots, f_m\} \subset L^2(\mathbb{R}^d)$, and $\Omega \subset \mathbb{R}^d$, there exists a shift-invariant space $S_\Lambda(\Phi) \in \mathcal{V}_M \cap \mathcal{C}_\ell(\Omega)$ such that $\mathcal{E}(\mathcal{F}, S_\Lambda(\Phi)) \leq \mathcal{E}(\mathcal{F}, S_\Lambda(\Psi))$ for all $S_\Lambda(\Psi) \in \mathcal{V}_M$ and $\Psi = \{\psi_1, \dots, \psi_\ell\} \subset PW_\Omega$.*

3.3. Approximation by Γ -invariant spaces

In this section $\Gamma = \Lambda \rtimes G$ where Λ is a full rank lattice of \mathbb{R}^d and G is a finite group that acts faithfully on \mathbb{R}^d by continuous automorphisms that preserve Λ . In this case the situation is slightly different, since the proof of an analogue of Theorem 7 has to rely on the fact that $\mathbb{P}_V S_\Gamma(\Phi) = S_\Gamma(\mathbb{P}_V(\Phi))$. For an arbitrary set Ω this will not necessarily be true. By Lemma 2, the condition we need to impose on Ω is that it is invariant under the action of the group G , that is $g\Omega = \Omega, \forall g \in G$.

With this extra condition we have again $\mathbb{P}_V S_\Gamma(\Phi) = S_\Gamma(\mathbb{P}_V(\Phi))$. Now, using that in Theorem 5.2 of [3] the existence of “best Γ -invariant spaces” has been proven, if we want to require that the approximating vectors are in $\mathcal{C}_\ell(\Omega)$ we can apply the same proof of Theorem 7, to obtain, the following Theorem.

Theorem 10. *Given data $\mathcal{F} = \{f_1, \dots, f_m\} \subset L^2(\mathbb{R}^d)$, $\Gamma = \mathbb{Z}^d \rtimes G$, with G being a finite group of that acts faithfully on \mathbb{R}^d by continuous automorphisms that preserve \mathbb{Z}^d , and $\Omega \subset \mathbb{R}^d$ such that Ω is invariant under the action of G , there exists a Γ -invariant space $S_\Gamma(\Phi)$, such that $S_\Gamma(\Phi) \in \mathcal{C}_\ell(\Omega)$ and $\mathcal{E}(\mathcal{F}, S_\Gamma(\Phi)) \leq \mathcal{E}(\mathcal{F}, S_\Gamma(\Psi))$ for all sets $\Psi = \{\psi_1, \dots, \psi_\ell\} \subset PW_\Omega$.*

4. Best set Ω

4.1. Shift-invariant case

Since we are looking to find a compact set that *best* fits the data in \mathbb{R}^d , in the sense that the Fourier Transform of the data is inside the set, without loss of generality, we can assume that the data already is in a huge $PW_{[-N, N]^d}$.

Lemma 11. *Let $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$ be a measurable function, and denote by Ω the function from $\mathbb{R}_{\geq 0}$ to the measurable sets of \mathbb{R}^d given by*

$$\Omega(t) = \{\xi \in \mathbb{R}^d : \varphi(\xi) > t\}.$$

For any $M > 0$, set $t^M = \sup\{t : |\Omega(t)| > M\}$ and $t_M = \inf\{t : |\Omega(t)| < M\}$. Let $\Omega_M \subset \mathbb{R}^d$ be any measurable set of measure $|\Omega_M| = M$ satisfying

$$\Omega(t_M) \subset \Omega_M \subset \Omega(t^M). \quad (9)$$

Then

$$\max_{E \subset \mathbb{R}^d : |E| = M} \int_E \varphi(\xi) d\xi = \int_{\Omega_M} \varphi(\xi) d\xi.$$

Proof. Observe first that, since $t \mapsto |\Omega(t)|$ is a monotone decreasing function which goes to zero at infinity, both t^M and t_M exist and are finite for every $M > 0$, with $t^M \leq t_M$. Assume now, by contradiction, that for a given $M > 0$ there exists a measurable $E \subset \mathbb{R}^d$, with $|E| = M$, such that

$$\int_E \varphi(\xi) d\xi > \int_{\Omega_M} \varphi(\xi) d\xi.$$

Denoting by $E_0 = E \setminus (E \cap \Omega_M)$ and by $\Omega_0 = \Omega_M \setminus (\Omega_M \cap E)$, this reads

$$\int_{E_0} \varphi(\xi) d\xi > \int_{\Omega_0} \varphi(\xi) d\xi$$

where $|E_0| = |\Omega_0|$ and $E_0 \cap \Omega_0 = \emptyset$. This implies in particular that there is a subset $E_+ \subset E_0$ of positive measure such that

$$\inf_{E_+} \varphi > \varphi(\xi) \text{ for a.e. } \xi \in \Omega_0. \quad (10)$$

Now, since by the right hand side inclusion of (9) we have $\varphi(\xi) > t^M$ for a.e. $\xi \in \Omega_0$, we have obtained that $\varphi(\xi) > t^M$ for a.e. $\xi \in E_+$, that is, $E_+ \subset \Omega(t^M)$. However, by the left hand side inclusion of (9) we have that the largest values in $\Omega(t^M)$ are already attained in Ω_0 , hence contradicting (10). \square

Theorem 12. Let $\mathcal{F} = \{f_i\}_{i=1}^m \subset L^2(\mathbb{R}^d)$, let $\varphi(\xi) = \sum_{i=1}^m |\widehat{f}_i(\xi)|^2$ and, for any $M > 0$, let Ω_M be as in (9). Then

$$\min_{E \subset \mathbb{R}^d : |E|=M} \sum_{i=1}^m \|\widehat{f}_i(1 - \chi_E)\|^2 = \sum_{i=1}^m \|\widehat{f}_i(1 - \chi_{\Omega_M})\|^2. \quad (11)$$

Proof. The identity (11) reads

$$\sum_{i=1}^m \int_{\mathbb{R}^d} |\widehat{f}_i(\xi)|^2 |1 - \chi_{\Omega_M}(\xi)|^2 d\xi \leq \sum_{i=1}^m \int_{\mathbb{R}^d} |\widehat{f}_i(\xi)|^2 |1 - \chi_E(\xi)|^2 d\xi$$

for all E of measure M . That is, using that $|1 - \chi_E(\xi)|^2 = 1 - \chi_E(\xi)$,

$$\int_{\Omega_M} \left(\sum_{i=1}^m |\widehat{f}_i(\xi)|^2 \right) d\xi \geq \int_E \left(\sum_{i=1}^m |\widehat{f}_i(\xi)|^2 \right) d\xi \quad \forall |E| = M.$$

Hence, the conclusion follows by Lemma 11 \square

Remark 13. Lemma 11 and Theorem 12 are still valid if \mathbb{R}^d is replaced by a σ -finite measure space (X, μ) . Furthermore, Theorem 12 also shows that actually there exists a set E_0 , with $|E_0| = M$ that maximizes $\{\int_E \varphi : E \subset \mathbb{R}^d : |E| = M\}$.

4.2. Γ -invariance case

When we optimize with Γ -invariant subspaces, we require the Paley-Wiener space PW_Ω to be Γ -invariant. Since any Paley-Wiener space is translation invariant and $\Gamma = \mathbb{Z}^d \rtimes G$, we only need to require PW_Ω to be G -invariant that is, Ω to be invariant under the action of G (i.e. $g\Omega = \Omega$ for all $g \in G$).

Choose a positive real number M and denote by $\mathcal{M}(G, M) = \{\Omega \subset \mathbb{R}^d : |\Omega| = M, \text{ and } g\Omega = \Omega, \forall g \in G\}$. Thus, given any $M > 0$ we need to find a measurable set $\Omega \in \mathcal{M}(G, M)$ that maximizes $\int_\Omega \varphi$ over the class $\mathcal{M}(G, M)$.

For this, let $\mathcal{F} = \{f_i\}_{i=1}^m \subset L^2(\mathbb{R}^d)$,

$$\varphi(\xi) = \sum_{i=1}^m |\widehat{f}_i(\xi)|^2 \quad (12)$$

and

$$\Phi_G(\xi) = \sum_{g \in G} \sum_{i=1}^m |\widehat{f}_i(g\xi)|^2. \quad (13)$$

Define the equivalence relation \sim_G in \mathbb{R}^d by $x \sim_G y$ if and only if there exist $g \in G$ such that $x = gy$. Choose $D \subset \mathbb{R}^d$ to be a measurable section of the quotient \mathbb{R}^d/\sim_G . Thus, $\mathbb{R}^d = \bigcup_{g \in G} gD$, with the union being disjoint.

The existence of a measurable section D can be obtained in the following way: In Proposition 4.2 in [3] it is proved that there exists a measurable set $P \subset \mathbb{R}^d$ of finite and positive Lebesgue measure such that, up to a measure zero set,

$$\mathbb{R}^d = \bigcup_{k \in \mathbb{Z}^d} \left(\bigcup_{g \in G} gP \right) + k,$$

with the union being disjoint. Using this we have,

$$\mathbb{R}^d = \bigcup_{k \in \mathbb{Z}^d} \left(\bigcup_{g \in G} gP \right) + k = \bigcup_{k \in \mathbb{Z}^d} \bigcup_{g \in G} g(P + g^{-1}k) = \bigcup_{g \in G} g \left(\bigcup_{k \in \mathbb{Z}^d} (P + g^{-1}k) \right).$$

Since $g\mathbb{Z}^d = \mathbb{Z}^d$ for all $g \in G$, we have $\bigcup_{k \in \mathbb{Z}^d} (P + g^{-1}k) = \bigcup_{k \in \mathbb{Z}^d} (P + k)$ and so $\mathbb{R}^d = \bigcup_{g \in G} gD$ with $D = \bigcup_{k \in \mathbb{Z}^d} (P + k)$ which is clearly measurable.

Theorem 14. *Let G be a finite group that leaves \mathbb{Z}^d invariant. Let $M > 0$ and $\mathcal{M} \equiv \mathcal{M}(G, M) = \{\Omega \subset \mathbb{R}^d : |\Omega| = M \text{ and } g\Omega = \Omega \ \forall g \in G\}$. For φ, Φ_G as in (12) and (13), assume that*

$$\Omega_0 \in A = \operatorname{argmax} \left\{ \int_{\Omega} \varphi : \Omega \in \mathcal{M} \right\}$$

and

$$\Sigma_0 \in B = \operatorname{argmax} \left\{ \int_{\Sigma} \Phi_G : \Sigma \subseteq D, |\Sigma| = M/(\#G) \right\}.$$

Then

$$\int_{\Omega_0} \varphi = \int_{\Sigma_0} \Phi_G.$$

Proof. For $\Omega_0 \in A$, consider $\Sigma = \Omega_0 \cap D$. Since $\Omega_0 \in \mathcal{M}$,

$$\Omega_0 = \left(\bigcup_{g \in G} gD \right) \cap \Omega_0 = \bigcup_{g \in G} (gD \cap \Omega_0) = \bigcup_{g \in G} (gD \cap g\Omega_0) = \bigcup_{g \in G} g\Sigma. \quad (14)$$

Therefore,

$$\int_{\Omega_0} \varphi(\xi) d\xi = \int_{\bigcup_{g \in G} g\Sigma} \varphi(\xi) d\xi = \sum_{g \in G} \int_{\Sigma} \varphi(g\omega) d\omega = \int_{\Sigma} \Phi_G(\omega) d\omega. \quad (15)$$

Since G preserves the lattice \mathbb{Z}^d it also preserves the Lebesgue measure of \mathbb{R}^d . From (14) we obtain $|\Sigma| = |\Omega_0|/(\#G) = M/(\#G)$. Hence, as $\Sigma_0 \in B$, we have

$$\int_{\Omega_0} \varphi(\xi) d\xi = \int_{\Sigma} \Phi_G(\omega) d\omega \leq \int_{\Sigma_0} \Phi_G(\omega) d\omega.$$

Now, for $\Sigma_0 \in \widehat{B}$, define $\Omega = \bigcup_{g \in G} g\Sigma_0$. The same computation as in (15) gives

$$\int_{\Sigma_0} \Phi_G(\omega) d\omega = \int_{\Omega} \varphi(\xi) d\xi.$$

Since $|\Omega| = (\#G)|\Sigma_0| = M$ with $g\Omega = \Omega$ for all $g \in G$, and as $\Omega_0 \in A$, we conclude

$$\int_{\Sigma_0} \Phi_G(\omega) d\omega = \int_{\Omega} \varphi(\xi) d\xi \leq \int_{\Omega_0} \varphi(\xi) d\xi. \quad \square$$

5. Lattices

Given a full rank lattice Λ in \mathbb{R}^d , let Ω_Λ be a fundamental set of Λ^\perp (the annihilator of Λ), that is a set such that $\Omega_\Lambda \approx \mathbb{R}^d/\Lambda^\perp$. For $f \in L^2(\mathbb{R}^d)$, denote by

$$\mathcal{T}_\Lambda f(\omega) = \{\widehat{f}(\omega + \gamma)\}_{\gamma \in \Lambda^\perp}, \quad \omega \in \Omega_\Lambda, \quad (16)$$

defining an isometric isomorphism $\mathcal{T}_\Lambda : L^2(\mathbb{R}^d) \rightarrow L^2(\Omega_\Lambda, \ell_2(\Lambda^\perp))$ (see Proposition 1.2 in [4] for the case $\Lambda = \mathbb{Z}^d$ and Proposition 3.3 in [6] for the general case).

For a family $\mathcal{F} = \{f_i\}_{i=1}^m \subset L^2(\mathbb{R}^d)$, denote also its Gramian matrix with respect to Λ by

$$(\mathcal{G}_\Lambda^\mathcal{F})_{ij}(\omega) = \sum_{\gamma \in \Lambda^\perp} \widehat{f}_i(\omega + \gamma) \overline{\widehat{f}_j(\omega + \gamma)}, \quad \omega \in \Omega_\Lambda.$$

We have the following elementary lemma whose proof is left to the reader.

Lemma 15. Let $A \in GL_d(\mathbb{R})$ and denote by $\widehat{A} = (A^t)^{-1}$ and by $|A| = \det(A)$.

- i) Let $\Lambda \subset \mathbb{R}^d$ be a full rank lattice and let Ω_Λ be a fundamental set for $\widehat{\mathbb{R}^d}/\Lambda^\perp$. Then $A\Lambda$ is a full rank lattice and $\widehat{A}\Omega_\Lambda$ is a fundamental set for $\widehat{\mathbb{R}^d}/\widehat{A}\Lambda^\perp$.
- ii) Let $f \in L^2(\mathbb{R}^d)$ and denote by $D_A f(x) = |A|^{\frac{1}{2}} f(Ax)$. Then

$$\widehat{D_A f} = D_{\widehat{A}} \widehat{f}.$$

Proposition 16. Let $\Lambda \subset \mathbb{R}^d$ be a full rank lattice and let Ω_Λ be a fundamental set for $\widehat{\mathbb{R}^d}/\Lambda^\perp$. Let $\mathcal{F} = \{f_i\}_{i=1}^m \subset L^2(\mathbb{R}^d)$, and denote by $D_A \mathcal{F} = \{D_A f_i\}_{i=1}^m$.

- a) for all $f \in L^2(\mathbb{R}^d)$ we have

$$\mathcal{T}_{A\Lambda} f(\widehat{A}\omega) = |A|^{\frac{1}{2}} \mathcal{T}_\Lambda (D_A f)(\omega), \quad \omega \in \Omega_\Lambda.$$

- b) the following covariance relationship for the Gramian holds:

$$\mathcal{G}_{A\Lambda}^\mathcal{F}(\widehat{A}\omega) = |A| \mathcal{G}_\Lambda^{D_A \mathcal{F}}(\omega), \quad \omega \in \Omega_\Lambda.$$

Proof. Both parts are consequence of item ii) of Lemma 15. We only proof b). For $\omega \in \Omega_\Lambda$ we have

$$\left(\mathcal{G}_{A\Lambda}^\mathcal{F}(\widehat{A}\omega) \right)_{ij} = \sum_{\lambda \in \widehat{A}\Lambda^\perp} \widehat{f}_i(\widehat{A}\omega + \lambda) \overline{\widehat{f}_j(\widehat{A}\omega + \lambda)}$$

$$\begin{aligned}
&= \sum_{\gamma \in \Lambda} \widehat{f_i}(\widehat{A}(\omega + \gamma)) \overline{\widehat{f_j}(\widehat{A}(\omega + \gamma))} \\
&= |A| \sum_{\gamma \in \Lambda^\perp} D_{\widehat{A}} \widehat{f_i}(\omega + \gamma) \overline{D_{\widehat{A}} \widehat{f_j}(\omega + \gamma)} \\
&= |A| \sum_{\gamma \in \Lambda^\perp} \widehat{D_A f_i}(\omega + \gamma) \overline{\widehat{D_A f_j}(\omega + \gamma)} = |A| \left(\mathcal{G}_\Lambda^{D_A \mathcal{F}}(\omega) \right)_{ij}. \quad \square
\end{aligned}$$

Recall that, for a general full rank lattice Λ of \mathbb{R}^d , we denote by $\mathcal{E}^*(\mathcal{F}, \Lambda, \ell)$ the minimum error of approximation of the data \mathcal{F} by a Λ -invariant subspace with at most ℓ generators.

Proposition 17. *For any full rank lattice Λ of \mathbb{R}^d , any set of data $\mathcal{F} = \{f_1, \dots, f_m\} \subset L^2(\mathbb{R}^d)$, and any $A \in GL_n(\mathbb{R})$,*

$$\mathcal{E}^*(\mathcal{F}, A\Lambda, \ell) = \mathcal{E}^*(D_A \mathcal{F}, \Lambda, \ell).$$

Proof. Let $S = S_{A\Lambda}(\{\psi_1, \dots, \psi_\ell\})$ be a general $A\Lambda$ -invariant subspace of $L^2(\mathbb{R}^d)$ of length ℓ . Let $S' = S_\Lambda(\{D_A \psi_1, \dots, D_A \psi_\ell\})$ be the Λ -invariant subspace generated by $\{D_A \psi_1, \dots, D_A \psi_\ell\}$. By part *i*) of Lemma 2,

$$D_A \mathbb{P}_S = \mathbb{P}_{D_A(S)} D_A.$$

Since for any ψ_j , $j = 1, 2, \dots, \ell$, and any $\gamma \in \Lambda$,

$$D_A T_{A\gamma} \psi_j(x) = |A|^{1/2} \psi_j(A(x - \gamma)) = T_\gamma D_A \psi_j(x),$$

we conclude that $D_A(S) = S'$. Therefore,

$$D_A \mathbb{P}_S = \mathbb{P}_{S'} D_A.$$

It follows that,

$$\begin{aligned}
\mathcal{E}(\mathcal{F}, S) &= \sum_{i=1}^m \|f_i - \mathbb{P}_S f_i\|^2 = \sum_{i=1}^m \|D_A f_i - D_A \mathbb{P}_S f_i\|^2 \\
&= \sum_{i=1}^m \|D_A f_i - \mathbb{P}_{S'} D_A f_i\|^2 = \mathcal{E}(D_A \mathcal{F}, S').
\end{aligned}$$

This shows that for any $A\Lambda$ -invariant subspace S of $L^2(\mathbb{R}^d)$ of length ℓ there is a Λ -invariant subspace of $L^2(\mathbb{R}^d)$, namely S' , such that $\mathcal{E}(\mathcal{F}, S) = \mathcal{E}(D_A \mathcal{F}, S')$. Therefore, $\mathcal{E}^*(D_A \mathcal{F}, \Lambda, \ell) \leq \mathcal{E}^*(\mathcal{F}, A\Lambda, \ell)$. The reverse inequality is proved similarly, since any Λ -invariant subspace $V = S_\Lambda(\{\varphi_1, \dots, \varphi_\ell\})$ of $L^2(\mathbb{R}^d)$ of length ℓ produces an $A\Lambda$ invariant subspace $V' = S_{A\Lambda}(\{D_{A^{-1}} \varphi_1, \dots, D_{A^{-1}} \varphi_\ell\}) = D_{A^{-1}}(V)$ for which $\mathcal{E}(D_A \mathcal{F}, V) = \mathcal{E}(\mathcal{F}, V')$. \square

Remark 18. A different proof of Proposition 17 can be given using part b) of Proposition 16 and Theorem 2.3 in [1] (generalized to include full-rank lattices of \mathbb{R}^d in Theorem 5.2 in [3]).

6. Examples

Let $\mathcal{F} = \{f_1, \dots, f_m\}$ be a finite set of data in $L^2(\mathbb{R}^d)$ and let Λ be a full rank lattice of \mathbb{R}^d . Proposition 17 shows that the best error of approximation of the modified data $D_A\mathcal{F} = \{D_A f_1, \dots, D_A f_m\}$ by Λ -invariant subspaces of length ℓ coincides with the best error of approximation of the original data \mathcal{F} by $A\Lambda$ -invariant subspaces also with length ℓ .

The situation is different if we try to compare the best error of approximation of the same data with subspaces that are invariant by two different lattices. We will show with one and two dimensional examples that the comparison depends on how the lattices relate to each other and also how the data behaves with respect to the lattices.

The simplest situation is when we consider the lattice \mathbb{Z}^d and a full rank lattice Λ containing \mathbb{Z}^d . The set

$$\{T_\gamma \psi_i : \gamma \in \Lambda, i = 1, \dots, \ell\},$$

clearly contains the set

$$\{T_k \psi_i : k \in \mathbb{Z}^d, i = 1, \dots, \ell\}.$$

Therefore, if $\Psi = \{\psi_1, \dots, \psi_\ell\}$,

$$S_{\mathbb{Z}^d}(\Psi) \subseteq S_\Lambda(\Psi). \quad (17)$$

For a given function $f \in L^2(\mathbb{R}^d)$, the distance between f and each one of the subspaces in (17) is smaller for the bigger subspace. Therefore, for any data set \mathcal{F} ,

$$\mathcal{E}(\mathcal{F}, S_\Lambda(\Psi)) \leq \mathcal{E}(\mathcal{F}, S_{\mathbb{Z}^d}(\Psi)).$$

Minimizing over all collections Ψ of ℓ functions of $L^2(\mathbb{R}^d)$ we deduce

$$\mathcal{E}^*(\mathcal{F}, \Lambda, \ell) \leq \mathcal{E}^*(\mathcal{F}, \mathbb{Z}^d, \ell). \quad (18)$$

6.1. One dimensional examples

Rewriting (18) for $d = 1$ and $N \in \mathbb{N}$, we have

$$\mathcal{E}^*(\mathcal{F}, \mathbb{Z}/N, \ell) \leq \mathcal{E}^*(\mathcal{F}, \mathbb{Z}, \ell). \quad (19)$$

In this section we shall show that, if $N \geq 2$, in some situations the above inequality is strict, but in others is not, depending on the data \mathcal{F} we want to approximate.

In the next examples we shall use the following result:

Lemma 19. *Let Λ be a full rank lattice in \mathbb{R}^d and $\psi \in L^2(\mathbb{R}^d)$.*

- If $f \in S_\Lambda(\psi)$, there exists a measurable Λ^\perp -periodic function m in \mathbb{R}^d such that $\widehat{f}(\omega) = m(\omega)\widehat{\psi}(\omega)$ a.e. $\omega \in \mathbb{R}^d$.*
- If m is a measurable Λ^\perp -periodic function in \mathbb{R}^d such that $m\widehat{\psi} \in L^2(\mathbb{R}^d)$, the function f defined by $\widehat{f}(\omega) = m(\omega)\widehat{\psi}(\omega)$ a.e. $\omega \in \mathbb{R}^d$ satisfies $f \in S_\Lambda(\psi)$.*

This result can be found in Theorem 2.14 in [8] for the case $\Lambda = \mathbb{Z}^d$. Their proof can be modified to give the Lemma.

Example 2. Let $\mathcal{F} = \{f_1, f_2\}$ where $\widehat{f}_1 = \chi_{[0,1/2)}$ and $\widehat{f}_2 = \chi_{[1/2,1)}$. Take $\varphi \in L^2(\mathbb{R})$ given by $\widehat{\varphi} = \chi_{[0,1)}$. Let

$$m_1(\omega) = \sum_{k \in \mathbb{Z}} \chi_{[0,1/2)}(\omega + k) \quad \text{and} \quad m_2(\omega) = \sum_{k \in \mathbb{Z}} \chi_{[1/2,1)}(\omega + k).$$

The functions $m_j, j = 1, 2$ are measurable and \mathbb{Z} -periodic. Moreover, $\widehat{f}_1 = m_1 \widehat{\varphi}$ and $\widehat{f}_2 = m_2 \widehat{\varphi}$. By b) of Lemma 19 both f_1 and f_2 belong to $S_{\mathbb{Z}}(\varphi)$. Hence, $\mathcal{E}(\mathcal{F}, S_{\mathbb{Z}}(\varphi)) = 0$ and also $\mathcal{E}^*(\mathcal{F}, \mathbb{Z}, 1) = 0$. By (19), $\mathcal{E}^*(\mathcal{F}, \mathbb{Z}/2, 1) = 0$ and (19) is an equality in this case.

Example 3. Let $\mathcal{F} = \{f_1, f_2\}$ where $\widehat{f}_1 = \chi_{[0,1/2)}$ and $\widehat{f}_2 = \chi_{[1,3/2)}$. Take $\varphi \in L^2(\mathbb{R})$ given by $\widehat{\varphi} = \chi_{[0,2)}$. Let

$$m_1(\omega) = \sum_{k \in \mathbb{Z}} \chi_{[0,1/2)}(\omega + 2k) \quad \text{and} \quad m_2(\omega) = \sum_{k \in \mathbb{Z}} \chi_{[1,3/2)}(\omega + 2k).$$

The functions $m_j, j = 1, 2$ are measurable and $2\mathbb{Z}$ -periodic. Moreover, $\widehat{f}_1 = m_1 \widehat{\varphi}$ and $\widehat{f}_2 = m_2 \widehat{\varphi}$. By b) of Lemma 19 both f_1 and f_2 belong to the $S_{\mathbb{Z}/2}(\varphi)$. As in the previous example, $\mathcal{E}^*(\mathcal{F}, \mathbb{Z}/2, 1) = 0$.

On the other hand $\mathcal{E}^*(\mathcal{F}, \mathbb{Z}, 1) > 0$. To prove this assume that $\mathcal{E}^*(\mathcal{F}, \mathbb{Z}, 1) = 0$. By Theorem 2.1 in [1] there exists $\varphi \in L^2(\mathbb{R})$ such that $\mathcal{E}(\mathcal{F}, S_{\mathbb{Z}}(\varphi)) = \mathcal{E}^*(\mathcal{F}, \mathbb{Z}, 1) = 0$. For this to hold we must have that f_1 and f_2 must belong to $S_{\mathbb{Z}}(\varphi)$. By part a) of Lemma 19, there exist \mathbb{Z} -periodic functions η_1 and η_2 such that for $\omega \in \mathbb{R}$,

$$\chi_{[0,1/2)}(\omega) = \widehat{f}_1(\omega) = \eta_1(\omega) \widehat{\varphi}(\omega) \quad \text{and} \quad \chi_{[1,3/2)}(\omega) = \widehat{f}_2(\omega) = \eta_2(\omega) \widehat{\varphi}(\omega). \quad (20)$$

For $\omega \in [0, 1/2)$, the first equation in (20) gives $\widehat{\varphi}(\omega) \neq 0$ for all $\omega \in [0, 1/2)$. The second equation in (20) produces $\eta_2(\omega) = 0$ for all $\omega \in [0, 1/2)$. Since η_2 is \mathbb{Z} -periodic we conclude that $\eta_2(\omega) = 0$ for $\omega \in [1, 3/2)$. This contradicts the second equation in (20).

Example 4. We will show that for each $h \in \mathbb{R}$ irrational there exists a set of data $\mathcal{F}_h \subset L^2(\mathbb{R})$ such that

$$0 = \mathcal{E}^*(\mathcal{F}_h, h\mathbb{Z}, 1) < \mathcal{E}^*(\mathcal{F}_h, \mathbb{Z}, 1).$$

Let $f = \chi_{[-1/2, 1/2]}$ and $\mathcal{F}_h = \{f, t_h f\}$. It is clear that f and $t_h f$ belong to $S_{h\mathbb{Z}}(f)$, the $h\mathbb{Z}$ -invariant subspace generated by f . Therefore $\mathcal{E}^*(\mathcal{F}_h, h\mathbb{Z}, 1) = 0$.

We want to show that $\mathcal{E}^*(\mathcal{F}_h, \mathbb{Z}, 1) > 0$. If $\mathcal{E}^*(\mathcal{F}_h, \mathbb{Z}, 1) = 0$, by Theorem 2.1 in [1] there exists $\varphi \in L^2(\mathbb{R})$ such that $\mathcal{E}(\mathcal{F}_h, S_{\mathbb{Z}}(\varphi)) = \mathcal{E}^*(\mathcal{F}_h, \mathbb{Z}, 1) = 0$. For this to hold we must have that f and $t_h f$ must belong to $S_{\mathbb{Z}}(\varphi)$. By part a) of Lemma 19, there exist \mathbb{Z} -periodic functions η_1 and η_2 such that for a.e. $\omega \in \mathbb{R}$,

$$\widehat{f}(\omega) = \eta_1(\omega) \widehat{\varphi}(\omega) \quad \text{and} \quad \widehat{t_h f}(\omega) = \eta_2(\omega) \widehat{\varphi}(\omega).$$

Since $\widehat{f}(\omega) = \text{Sinc}(\omega)$ and $\widehat{t_h f}(\omega) = e^{-2\pi i h \omega} \text{Sinc}(\omega)$, we have

$$\text{Sinc}(\omega) = \eta_1(\omega) \widehat{\varphi}(\omega) \quad \text{and} \quad e^{-2\pi i h \omega} \text{Sinc}(\omega) = \eta_2(\omega) \widehat{\varphi}(\omega).$$

Let Ω be the set where $\text{Sinc}(\omega) \neq 0$ which satisfies $|\mathbb{R} \setminus \Omega| = 0$. On this set Ω

$$e^{-2\pi i h \omega} = \frac{\eta_2(\omega)}{\eta_1(\omega)}.$$

This is impossible because the function in the right hand side of the above equation is \mathbb{Z} -periodic while the function in the left is not.

Observe that the same proof works for any $f \in L^2(\mathbb{R})$ such that the support of \widehat{f} is \mathbb{R} .

6.2. Two dimensional examples with an isometry

Now we will look at two examples in which we allow for different lattices – not necessarily a refinement of one another. The examples show that lattices play a significant role in the computation of the error.

In order to present them, we need to recall the notion of *range function* which was introduced by Helson and Lowdenslager in [10,9] and developed by Bownik in [4]. Let Λ be a full rank lattice in \mathbb{R}^d and let $\Omega_\Lambda \subset \widehat{\mathbb{R}^d}$ be a Borel section of $\widehat{\mathbb{R}^d}/\Lambda^\perp$. Recall the definition of the isometric isomorphism \mathcal{T}_Λ given in equation (16). A range function is a map

$$\mathcal{J} : \Omega_\Lambda \rightarrow \{\text{closed subspaces of } \ell_2(\Lambda^\perp)\}.$$

A range function \mathcal{J} is said to be measurable if the family $\mathcal{P}_{\mathcal{J}(\omega)} \in \mathcal{B}(\ell_2(\Lambda^\perp))$ of orthogonal projections onto $\mathcal{J}(\omega)$ is measurable.

A closed subspace V of $L^2(\mathbb{R}^d)$ is Λ -invariant if and only if there exists a measurable range function \mathcal{J} such that

$$V = \{f \in L^2(\mathbb{R}^d) : \mathcal{T}_\Lambda f(\omega) \in \mathcal{J}(\omega) \text{ a.e } \omega \in \Omega_\Lambda\}.$$

Moreover, if $V = S_\Lambda(\mathcal{A})$ for some countable set \mathcal{A} of $L^2(\mathbb{R}^d)$ the measurable range function associated to V is given by

$$\mathcal{J}(\omega) = \overline{\text{span}}\{\mathcal{T}_\Lambda \varphi(\omega) : \varphi \in \mathcal{A}\} \quad \text{a.e } \omega \in \Omega_\Lambda.$$

Good references for this result are Proposition 1.5 in [4] and Theorem 3.10 in [6].

Let $\ell(S_\Lambda(\mathcal{A}))$ denote the length of $S_\Lambda(\mathcal{A})$, that is the smallest number of generators of $S_\Lambda(\mathcal{A})$. It then can be seen that

$$\ell(S_\Lambda(\mathcal{A})) = \text{ess sup}_{\omega \in \Omega_\Lambda} \dim(\mathcal{J}(\omega)).$$

Example 5. Let $\Lambda = \mathbb{Z}^2$ and $\Lambda_1 = R\mathbb{Z}^2$ be a rotation of \mathbb{Z}^2 by $\frac{\pi}{6}$. Let $\gamma := \left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$ and let $\Omega = [0, 1) \times [0, 1)$, and $\Omega_1 = R\Omega$ be fundamental domains for Λ and Λ_1 respectively. Observe that since R is unitary, the dual lattice of Λ_1 is itself. Let $B_1, B_2 \subset \Omega$ be the following circles (see Fig. 1):

$$B_1 = B\left((0, 0), \frac{1}{25}\right) + \left(\frac{2}{25}, \frac{3}{10}\right)$$

$$B_2 = B_1 + \gamma = B\left((0, 0), \frac{1}{25}\right) + \left(\frac{\sqrt{3}}{2} + \frac{2}{25}, \frac{1}{2} + \frac{3}{10}\right).$$

This example is based on the fact that $B_2 = B_1 + \gamma$, but for the lattice \mathbb{Z}^2 both belong to the same fundamental domain while for $R\mathbb{Z}^2$ they don't. That is, $B_1 \in \Omega \cap R\Omega$, while $B_2 \in \Omega \setminus R\Omega$.

Let $\mathcal{F} = \{f_1, f_2\} \subset L^2(\mathbb{R}^2)$ be defined by $\widehat{f}_i = \chi_{B_i}$, $i = 1, 2$. Hence, for the lattice $\Lambda = \mathbb{Z}^2$,

$$\mathcal{T}_\Lambda[f_1](\omega) = \begin{cases} e_{(0,0)} & \omega \in B_1 \\ 0 & \omega \in \Omega \setminus B_1 \end{cases} \quad \mathcal{T}_\Lambda[f_2](\omega) = \begin{cases} e_{(0,0)} & \omega \in B_2 \\ 0 & \omega \in \Omega \setminus B_2 \end{cases}.$$

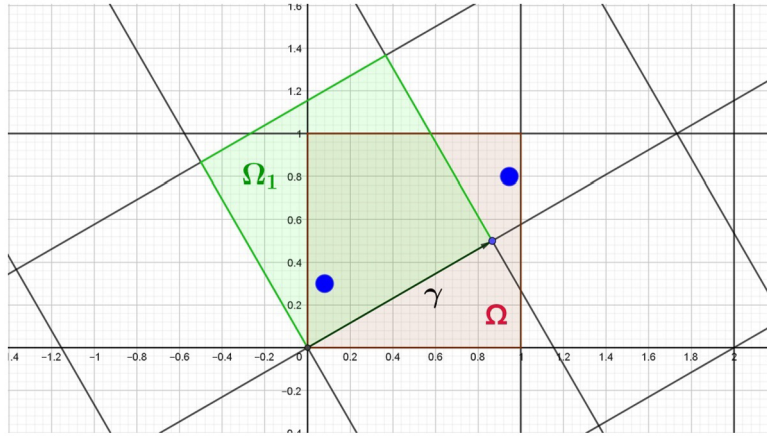


Fig. 1. The sets of Example 5.

This means that, for *a.e.* $\omega \in \Omega$, the dimension of $\mathcal{J}(\omega)$ is less or equal than 1, and the length of $S(\mathcal{F})$ is $\ell(S(\mathcal{F})) = 1$.

If $\Phi = \{\varphi\}$ with

$$\mathcal{T}_{\Lambda}[\varphi](\omega) = \begin{cases} e_{(0,0)} & \omega \in B_1 \cup B_2 \\ 0 & \omega \in \Omega \setminus (B_1 \cup B_2) \end{cases},$$

we have that $\|f_1 - \mathbb{P}_{S_{\Lambda}(\varphi)}(f_1)\| = \|f_2 - \mathbb{P}_{S_{\Lambda}(\varphi)}(f_2)\| = 0$ a.e. $\omega \in \Omega$ and hence $S(\varphi)$ approximates the data \mathcal{F} perfectly.

However, if instead of the lattice $\Lambda = \mathbb{Z}^2$ we take the lattice $\Lambda_1 = R\mathbb{Z}^2$, then

$$\mathcal{T}_{\Lambda_1}[f_1](\omega) = \begin{cases} e_{(0,0)} & \omega \in B_1 \\ 0 & \omega \in \Omega_1 \setminus B_1 \end{cases} \quad \mathcal{T}_{\Lambda_1}[f_2](\omega) = \begin{cases} e_{(1,0)} & \omega \in B_1 \\ 0 & \omega \in \Omega_1 \setminus B_1 \end{cases}$$

So, for $\omega \in B_1$, the dimension of $\mathcal{J}(\omega)$ is 2; hence the length of $S_{\Lambda_1}(\mathcal{F})$ is 2 and therefore, if we want to approximate with a single function, the error can never be 0. Hence, $\mathcal{E}^*(\mathcal{F}, \mathbb{Z}^2, 1) < \mathcal{E}^*(\mathcal{F}, R\mathbb{Z}^2, 1)$.

This example induces us to believe that to minimize the error of approximation we should search for the lattice Λ that minimizes the length of $S_{\Lambda}(\mathcal{F})$. As we show in the next example, this is not necessarily true.

Example 6. Let the lattices and fundamental domains be as in Example 5, and let $\gamma_1 = \left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$ and $\gamma_2 = 2 \cdot \left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$, Consider the following sets (see Fig. 2):

$$B_1 = B \left((0, 0), \frac{1}{25} \right) + \left(\frac{2}{25}, \frac{3}{10} \right)$$

$$B_2 = B_1 + \gamma_1$$

$$B_3 = B_1 + \gamma_2$$

$$Q_1 = \text{square of side } \frac{1}{25} \text{ with lower left vertex at } \left(\frac{1}{5}, \frac{1}{25} \right)$$

$$Q_2 = Q_1 + (1, 0).$$

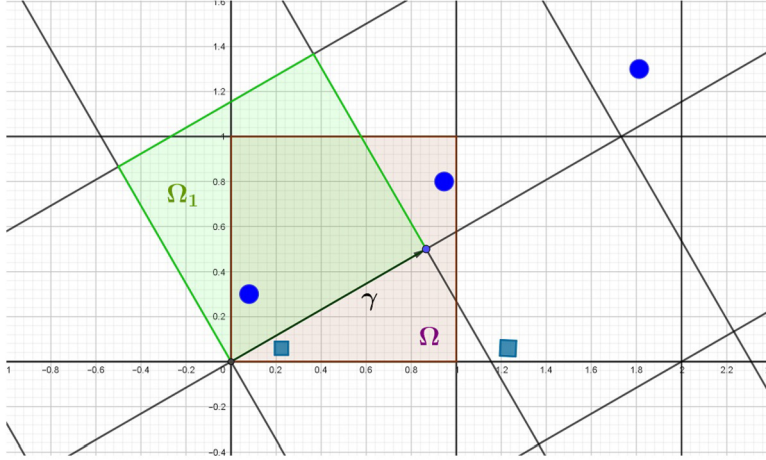


Fig. 2. The sets of Example 6.

Let $\mathcal{F} = \{f_1, f_2, f_3, f_4, f_5\} \subset L^2(\mathbb{R}^2)$ be defined by

$$\begin{aligned} \hat{f}_1 &= \chi_{B_1}, \quad \hat{f}_2 = \chi_{B_1} + \varepsilon \cdot \chi_{B_2}, \quad \hat{f}_3 = \chi_{B_1} + \varepsilon \cdot \chi_{B_3}, \\ \hat{f}_4 &= \chi_{Q_1}, \quad \hat{f}_5 = \chi_{Q_2}. \end{aligned}$$

Hence, for the lattice \mathbb{Z}^2 , and $\Omega = [0, 1) \times [0, 1)$,

$$\begin{aligned} \mathcal{T}_{\mathbb{Z}^2}[f_1](\omega) &= \begin{cases} e_{(0,0)} & \omega \in B_1 \\ 0 & \omega \in \Omega \setminus B_1 \end{cases} \\ \mathcal{T}_{\mathbb{Z}^2}[f_2](\omega) &= \begin{cases} e_{(0,0)} & \omega \in B_1 \\ \varepsilon \cdot e_{(0,0)} & \omega \in B_2 \\ 0 & \omega \in \Omega \setminus (B_1 \cup B_2) \end{cases} \\ \mathcal{T}_{\mathbb{Z}^2}[f_3](\omega) &= \begin{cases} e_{(0,0)} & \omega \in B_1 \\ \varepsilon \cdot e_{(1,1)} & \omega \in B_3 - (1, 1) \\ 0 & \omega \in \Omega \setminus ((B_3 - (1, 1)) \cup B_1) \end{cases} \\ \mathcal{T}_{\mathbb{Z}^2}[f_4](\omega) &= \begin{cases} e_{(0,0)} & \omega \in Q_1 \\ 0 & \omega \in \Omega \setminus Q_1 \end{cases} \\ \mathcal{T}_{\mathbb{Z}^2}[f_5](\omega) &= \begin{cases} e_{(1,0)} & \omega \in Q_1 \\ 0 & \omega \in \Omega \setminus Q_1 \end{cases} \end{aligned}$$

This means, that for almost every ω the dimension of $\mathcal{J}(\omega)$ is less or equal than 2, and on Q_1 the dimension is 2, so, the length of $S_{\mathbb{Z}^2}(\mathcal{F})$ is 2.

Suppose now that we want to approximate the data with a single generator. For the lattice \mathbb{Z}^2 we can represent f_1, f_2 and f_3 with zero error with any φ such that

$$\mathcal{T}_{\mathbb{Z}^2}[\varphi](\omega)\chi_{\Omega \setminus Q_1}(\omega) = e_{(0,0)}\chi_{B_1}(\omega) + \varepsilon e_{(0,0)}\chi_{B_2}(\omega) + \varepsilon e_{(1,1)}\chi_{B_3 - (1,1)}(\omega).$$

For $\omega \in Q_1$, the Gramian of $\{f_4, f_5\}$ with respect to the lattice \mathbb{Z}^2 is the 2×2 identity matrix since the vectors $e_{(0,0)}$ and $e_{(1,0)}$ are orthonormal in $\ell_2(\mathbb{Z}^2)$. By Theorem 2.3 in [1],

$$\mathcal{E}^*(\mathcal{F}, \mathbb{Z}^2, 1) = \mathcal{E}^*(\{f_4, f_5\}, \mathbb{Z}^2, 1) = \int_{Q_1} 1 d\omega = |Q_1|.$$

On the other hand, if instead of the lattice \mathbb{Z}^2 we take the lattice $\Lambda = R\mathbb{Z}^2$, where R is a rotation of $\frac{\pi}{6}$, with fundamental domain $\Omega_1 = R\Omega$, we get

$$\begin{aligned}\mathcal{T}_\Lambda[f_1](\omega) &= \begin{cases} e_{(0,0)} & \omega \in B_1 \\ 0 & \omega \in \Omega_1 \setminus B_1 \end{cases} \\ \mathcal{T}_\Lambda[f_2](\omega) &= \begin{cases} e_{(0,0)} + \varepsilon \cdot e_{(1,0)} & \omega \in B_1 \\ 0 & \omega \in \Omega_1 \setminus B_1 \end{cases} \\ \mathcal{T}_\Lambda[f_3](\omega) &= \begin{cases} e_{(0,0)} + \varepsilon \cdot e_{(2,0)} & \omega \in B_1 \\ 0 & \omega \in \Omega_1 \setminus B_1 \end{cases} \\ \mathcal{T}_\Lambda[f_4](\omega) &= \begin{cases} e_{(0,-1)} & \omega \in Q_1 + (-\frac{1}{2}, \frac{\sqrt{3}}{2}) \\ 0 & \omega \in \Omega_1 \setminus (Q_1 + (-\frac{1}{2}, \frac{\sqrt{3}}{2})) \end{cases} \\ \mathcal{T}_\Lambda[f_5](\omega) &= \begin{cases} e_{(1,-1)} & \omega \in Q_1 - \gamma_1 + (-\frac{1}{2}, \frac{\sqrt{3}}{2}) \\ 0 & \omega \in \Omega_1 \setminus (Q_1 - \gamma_1 + (-\frac{1}{2}, \frac{\sqrt{3}}{2})) \end{cases}\end{aligned}$$

Hence, in B_1 we have three linearly independent vectors, and on every other point only one or none, so the length of $S_\Lambda(\mathcal{F})$ is 3.

For the lattice Λ we can represent f_4 and f_5 with zero error taking any φ such that

$$\mathcal{T}_\Lambda[\varphi](\omega)\chi_{\Omega \setminus B_1}(\omega) = e_{(0,-1)}\chi_{Q_1 + (-\frac{1}{2}, \frac{\sqrt{3}}{2})}(\omega) + e_{(1,-1)}\chi_{Q_1 - \gamma_1 + (-\frac{1}{2}, \frac{\sqrt{3}}{2})}(\omega).$$

For $\omega \in B_1$, we take φ such that $\mathcal{T}_\Lambda[\varphi](\omega)\chi_{B_1}(\omega) = e_{(0,0)}\chi_{B_1}(\omega)$. Then,

$$\mathcal{E}(\{f_1, f_2, f_3\}, S_\Lambda(\varphi)) = \int_{B_1} (\|\varepsilon \cdot e_{(1,0)}\|_{\ell_2(\Lambda)}^2 + \|\varepsilon \cdot e_{(2,0)}\|_{\ell_2(\Lambda)}^2) d\omega = 2\varepsilon^2|B_1|.$$

Thus,

$$\mathcal{E}^*(\mathcal{F}, \Lambda, 1) = \mathcal{E}^*(\{f_1, f_2, f_3\}, \Lambda, 1) \leq 2\varepsilon^2|B_1|.$$

So, for a sufficiently small ε , the set \mathcal{F} of this example can be better approximated in the second lattice, even though the length is bigger.

References

- [1] A. Aldroubi, C. Cabrelli, D. Hardin, U. Molter, Optimal shift-invariant spaces and their Parseval frame generators, *Appl. Comput. Harmon. Anal.* 23 (2007) 273–283.
- [2] A. Aldroubi, I. Krishtal, R. Tessera, H. Wang, Principal shift-invariant spaces with extra invariance nearest to observed data, *Collect. Math.* 63 (2012) 393–401.
- [3] D. Barbieri, C. Cabrelli, E. Hernández, U. Molter, Approximation by group invariant subspaces, *J. Math. Pures Appl.* 142 (2020) 76–100.
- [4] M. Bownik, The structure of shift-invariant subspaces of $L^2(\mathbb{R}^n)$, *J. Funct. Anal.* 177 (2) (2000) 282–309.
- [5] C. Cabrelli, C.A. Mosquera, Subspaces with extra invariance nearest to observed data, *Appl. Comput. Harmon. Anal.* 41 (2) (2016) 660–676.
- [6] C. Cabrelli, V. Paternostro, Shift-invariant spaces on LCA groups, *J. Funct. Anal.* 258 (2010) 2034–2059.
- [7] H.H. Cuenya, F.E. Levis, Existence of optimal subspaces in reflexive Banach spaces, *Ann. Funct. Anal.* 6 (2015) 69–77.

- [8] C. de Boor, R. DeVore, A. Ron, Approximation by shift-invariant subspaces of $L_2(\mathbb{R}^d)$, *Trans. Am. Math. Soc.* 341 (1994) 787–806.
- [9] H. Helson, *Lectures on Invariant Subspaces*, Academic Press, New York/London, 1964.
- [10] H. Helson, D. Lowdenslager, Prediction theory and Fourier series in several variables, *Acta Math.* 99 (1958) 165–202.
- [11] K.D. Merrill, Wavelet sets for crystallographic groups, in: *Excursions in Harmonic Analysis*, vol. 6, Springer, 2021, pp. 185–198.