Stability of Algebraic Filters

Alejandro Ribeiro and Alejandro Parada-Mayorga

In this article, we provide a tutorial proof for the first stability result presented in Lecture 12 — stability of algebraic filters. Before proceeding to prove this result, we recall the definition of a stable operator in algebraic signal processing.

Definition 1 (Operator Stability) *Given operators $p(\mathbf{S})$ and $p(\tilde{\mathbf{S}})$ defined on the processing models $(\mathcal{A}, \mathcal{M}, \rho)$ and $(\mathcal{A}, \mathcal{M}, \tilde{\rho})$, we say the operator $p(\mathbf{S})$ is Lipschitz stable if there exist constants $C_0, C_1 > 0$ such that*

$$\left\| p(\mathbf{S})\mathbf{x} - p(\tilde{\mathbf{S}})\mathbf{x} \right\| \leq \left[C_0 \sup_{\mathbf{S}\in\mathcal{S}} \|\mathbf{T}(\mathbf{S})\| + C_1 \sup_{\mathbf{S}\in\mathcal{S}} \left\| D_{\mathbf{T}}(\mathbf{S}) \right\| + \mathcal{O}\left(\|\mathbf{T}(\mathbf{S})\|^2 \right) \right] \left\| \mathbf{x} \right\|, \quad (1)$$

for all $\mathbf{x} \in \mathcal{M}$. In (1), $D_{\mathbf{T}}(\mathbf{S})$ is the Fréchet derivative of the perturbation operator $\mathbf{T}$.

We also recall the definitions of Lipschitz and integral Lipschitz filters.

Definition 2 *Let $p : \mathbb{C} \to \mathbb{C}$ be single variable function. Then, it is said that p is Lipschitz if there exists $L_0 > 0$ such that*

$$|p(\lambda) - p(\mu)| \leq L_0 |\lambda - \mu| \quad (2)$$

for all $\lambda, \mu \in \mathbb{C}$. Additionally, it is said that $p(\lambda)$ is Lipschitz integral if there exists $L_1 > 0$ such that

$$\left| \lambda \frac{dp(\lambda)}{d\lambda} \right| \leq L_1 \text{ for all } \lambda. \quad (3)$$

In what follows, when considering subsets of a commutative algebra $\mathcal{A}$, we denote by $\mathcal{A}_{L_0}$ the subset of elements in $\mathcal{A}$ that are Lipschitz with constant L_0 and by $\mathcal{A}_{L_1}$ the subset of element of $\mathcal{A}$ that are Lipschitz integral with constant L_1. Additionally, for the sake of simplicity we will not make reference to ι.

Our goal is to prove stability of filters p satisfying the above definition, as stated in the following theorem.

Theorem 1 *Let $\mathcal{A}$ be an algebra with one generator element g and let $(\mathcal{M}, \rho)$ be a finite or countable infinite dimensional representation of $\mathcal{A}$. Let $(\mathcal{M}, \tilde{\rho})$ be a perturbed version of $(\mathcal{M}, \rho)$. Then, if $p_{\mathcal{A}} \in \mathcal{A}_{L_0} \cap \mathcal{A}_{L_1}$ the operator $p(\mathbf{S})$ is stable in the sense of Definition 1 with $C_0 = (1+\delta)L_0$ and $C_1 = (1+\delta)L_1$.*

Proof: The proof of this theorem follows immediately from Lemma 1 and Lemma 2 below, whose proofs are deferred to Sections 1 and 2 respectively. Lemma 1 is a result for operators in algebraic models with a single generator, highlighting the role of the Fréchet derivative of the map that relates the operator and its perturbed version. Lemma 2 shows that this term is related to $\mathbf{T}(\mathbf{S})$ and its Fréchet derivative $D_{\mathbf{T}}$.

Lemma 1 *Let $\mathcal{A}$ be an algebra generated by g and let $(\mathcal{M}, \rho)$ be a representation of $\mathcal{A}$ with $\rho(g) = \mathbf{S} \in End(\mathcal{M})$. Let $\tilde{\rho}(g) = \tilde{\mathbf{S}} \in End(\mathcal{M})$ where the pair $(\mathcal{M}, \tilde{\rho})$ is a perturbed version of $(\mathcal{M}, \rho)$. Then, for any $p_{\mathcal{A}} \in \mathcal{A}$ we have*

$$\left\| p(\mathbf{S})\mathbf{x} - p(\tilde{\mathbf{S}})\mathbf{x} \right\| \leq \|\mathbf{x}\| \left(\left\| D_p(\mathbf{S})\left\{ \mathbf{T}(\mathbf{S}) \right\} \right\| + \mathcal{O}\left(\|\mathbf{T}(\mathbf{S})\|^2 \right) \right) \tag{4}$$

where $D_p(\mathbf{S})$ is the Fréchet derivative of p on $\mathbf{S}$.

Lemma 2 *Let $\mathcal{A}$ be an algebra with one generator element g and let $(\mathcal{M}, \rho)$ be a finite or countable infinite dimensional representation of $\mathcal{A}$. Let $(\mathcal{M}, \tilde{\rho})$ be a perturbed version of $(\mathcal{M}, \rho)$. If $p_{\mathcal{A}} \in \mathcal{A}_{L_0} \cap \mathcal{A}_{L_1}$, then*

$$\left\| D_p \mathbf{T}(\mathbf{S}) \right\| \leq (1+\delta) \left(L_0 \sup_{\mathbf{S}} \|\mathbf{T}(\mathbf{S})\| + L_1 \sup_{\mathbf{S}} \|D_{\mathbf{T}}(\mathbf{S})\| \right) \tag{5}$$

The stability result is then proved by simply replacing (5) from Lemma 2 into (4) from Lemma 1 and reordering terms.

■

1 Proof of Lemma 1

We say that $p(\mathbf{S})$ as a function of $\mathbf{S}$ is Fréchet differentiable at $\mathbf{S}$ if there exists a bounded linear operator $D_p : \text{End}(\mathcal{M})^m \to \text{End}(\mathcal{M})$ such that [1, 2]

$$\lim_{\|\boldsymbol{\xi}\|\to 0} \frac{\left\|p(\mathbf{S}+\boldsymbol{\xi}) - p(\mathbf{S}) - D_p(\mathbf{S})\left\{\boldsymbol{\xi}\right\}\right\|}{\|\boldsymbol{\xi}\|} = 0 \tag{6}$$

which in Landau notation can be written as

$$p(\mathbf{S}+\boldsymbol{\xi}) = p(\mathbf{S}) + D_p(\mathbf{S})\left\{\boldsymbol{\xi}\right\} + o(\|\boldsymbol{\xi}\|). \tag{7}$$

Calculating the norm in eqn. (7) and applying the triangle inequality we have: $\|p(\mathbf{S}+\boldsymbol{\xi}) - p(\mathbf{S})\| \leq \left\|D_p(\mathbf{S})\left\{\boldsymbol{\xi}\right\}\right\| + \mathcal{O}\left(\|\boldsymbol{\xi}\|^2\right)$ for all $\boldsymbol{\xi} = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_m) \in \text{End}(\mathcal{M})^m$. Now, taking into account that (see [3] pages 69-70)

$$\|D_p(\mathbf{S})\left\{\boldsymbol{\xi}\right\}\| \leq \sum_{i=1}^{m} \left\|D_{p|\mathbf{S}_i}(\mathbf{S})\left\{\boldsymbol{\xi}_i\right\}\right\| \tag{8}$$

we have

$$\|p(\mathbf{S}+\boldsymbol{\xi}) - p(\mathbf{S})\| \leq \sum_{i=1}^{m} \left\|D_{p|\mathbf{S}_i}(\mathbf{S})\left\{\boldsymbol{\xi}_i\right\}\right\| + \mathcal{O}\left(\|\boldsymbol{\xi}\|^2\right),$$

where $D_{p|\mathbf{S}_i}(\mathbf{S})$ is the partial Frechet derivative of $p(\mathbf{S})$ on $\mathbf{S}_i$. Then, taking into account that $\|p(\mathbf{S}+\boldsymbol{\xi})\mathbf{x} - p(\mathbf{S})\mathbf{x}\| \leq \|\mathbf{x}\| \, \|p(\mathbf{S}+\boldsymbol{\xi}) - p(\mathbf{S})\|$ and selecting $\boldsymbol{\xi}_i = \mathbf{T}(\mathbf{S}_i)$ we complete the proof.

2 Proof of Lemma 2

To start, we discuss how to find the Frechet Derivative $D_{p|\mathbf{S}_i}(\mathbf{S})$. First, notice that $p(\mathbf{S}) = \sum_{k_1,\ldots,k_m=0}^{\infty} h_{k_1\ldots k_m} \mathbf{S}_1^{k_1} \ldots \mathbf{S}_m^{k_m} = \sum_{k_i=0}^{\infty} \mathbf{S}_i^{k_i} \mathbf{A}_{k_i}$, where $\mathbf{A}_{k_i} = \sum_{\substack{\{k_j\}=0 \\ j\neq i}}^{\infty} h_{k_1,\ldots,k_m} \prod_{\substack{j=1 \\ j\neq i}}^{m} \mathbf{S}_j^{k_j}$. Then, it follows that

$$p(\mathbf{S}+\boldsymbol{\xi}) - p(\mathbf{S}) = \sum_{k_i=0}^{\infty} \left(\mathbf{S}_i + \boldsymbol{\xi}_i\right)^{k_i} \mathbf{A}_{k_i} - \sum_{k_i=0}^{\infty} \mathbf{S}_i^{k_i} \mathbf{A}_{k_i} \tag{9}$$

for $\boldsymbol{\xi} = (\mathbf{0}, \ldots, \boldsymbol{\xi}_i, \ldots, \mathbf{0})$. Considering the expansion $(\mathbf{S}_i + \boldsymbol{\xi}_i)^{k_i} = \mathbf{S}_i^{k_i} + \boldsymbol{\xi}_i^{k_i} + \sum_{r=1}^{k-1} \boldsymbol{\pi}(r\mathbf{S}_i, (k_i - r)\boldsymbol{\xi}_i)$ for $k_i \geq 2$, eqn. (9) takes the form

$$p(\mathbf{S} + \boldsymbol{\xi}) - p(\mathbf{S}) = \sum_{k_i=1}^{\infty} \sum_{r=1}^{k_i-1} \boldsymbol{\pi}\left(r\boldsymbol{\xi}_i, (k_i - r)\mathbf{S}_i\right) \mathbf{A}_{k_i} + \sum_{k_i=1}^{\infty} \boldsymbol{\xi}_i^{k_i} \mathbf{A}_{k_i}. \tag{10}$$

Separating the linear terms on $\boldsymbol{\xi}_i$ eqn. (10) leads to

$$\begin{aligned} p(\mathbf{S} + \boldsymbol{\xi}) - p(\mathbf{S}) = &\sum_{k_i=1}^{\infty} \boldsymbol{\pi}\left(\boldsymbol{\xi}_i, (k_i - 1)\mathbf{S}_i\right) \mathbf{A}_{k_i} \\ &+ \sum_{k_i=2}^{\infty} \sum_{r=2}^{k_i-1} \boldsymbol{\pi}\left(r\boldsymbol{\xi}_i, (k_i - r)\mathbf{S}_i\right) \mathbf{A}_{k_i} + \sum_{k_i=2}^{\infty} \boldsymbol{\xi}^{k_i} \mathbf{A}_{k_i}. \end{aligned} \tag{11}$$

Therefore, taking into account the definition of Fréchet derivative (see Section 1) it follows that

$$D_{p|\mathbf{S}_i}(\mathbf{S}) \left\{\boldsymbol{\xi}_i\right\} = \sum_{k_i=1}^{\infty} \boldsymbol{\pi}\left(\boldsymbol{\xi}_i, (k_i - 1)\mathbf{S}_i\right) \mathbf{A}_{k_i} \tag{12}$$

We may now proceed to prove Lemma 2.

Proof: [Proof of Lemma 2] Taking into account the definition of the Fréchet derivative, we have

$$\left\| D_{p|\mathbf{S}_i}(\mathbf{S}) \left\{\mathbf{T}(\mathbf{S}_i)\right\} \right\| = \left\| \sum_{k_i=1}^{\infty} \mathbf{A}_{k_i} \boldsymbol{\pi}\left(\mathbf{T}(\mathbf{S}_i), (k_i - 1)\mathbf{S}_i\right) \right\|,$$

and re-organizating terms we have

$$\left\| D_{p|\mathbf{S}_i}(\mathbf{S}) \left\{\mathbf{T}(\mathbf{S}_i)\right\} \right\| = \left\| \sum_{\ell=1}^{\infty} \mathbf{S}_i^{\ell-1} \mathbf{T}(\mathbf{S}_i) \sum_{k_i=\ell}^{\infty} \mathbf{A}_{k_i} \mathbf{S}_i^{k_i-\ell} \right\|.$$

Taking into account that $\mathbf{S}\mathbf{T}_r = \mathbf{T}_{cr}\mathbf{S} + \mathbf{S}\mathbf{P}_r$, it follows that

$$\begin{aligned} &\left\| D_{p|\mathbf{S}_i}(\mathbf{S}) \left\{\mathbf{T}(\mathbf{S}_i)\right\} \right\| = \\ &\left\| \sum_{\ell=1}^{\infty} \left(\mathbf{T}_{0c,i}\mathbf{S}_i^{\ell-1} + \mathbf{S}_i^{\ell-1}\mathbf{P}_{0,i}\right) \sum_{k_i=\ell}^{\infty} \mathbf{A}_{k_i} \mathbf{S}_i^{k_i-\ell} + \sum_{\ell=1}^{\infty} \left(\mathbf{T}_{1c,i}\mathbf{S}_i^{\ell} + \mathbf{S}_i^{\ell-1}\mathbf{P}_{1,i}\mathbf{S}_i\right) \sum_{k=\ell}^{\infty} \mathbf{A}_{k_i} \mathbf{S}^{k_i-\ell} \right\|. \end{aligned} \tag{13}$$

Applying the triangle inequality and distribuiting the sum we have

$$\left\| D_{p|\mathbf{S}_i}(\mathbf{S})\left\{\mathbf{T}(\mathbf{S}_i)\right\}\right\| \leq \left\|\mathbf{T}_{0c,i}\sum_{\ell=1}^{\infty}\sum_{k_i=\ell}^{\infty}\mathbf{S}_i^{k_i-1}\mathbf{A}_{k_i}\right\| + \left\| D_{p|\mathbf{S}_i}(\mathbf{S})\left\{\mathbf{P}_{0,i}\right\}\right\|$$
$$+\left\|\mathbf{T}_{1c,i}\sum_{\ell=1}^{\infty}\sum_{k_i=\ell}^{\infty}\mathbf{S}^{k_i}\mathbf{A}_{k_i}\right\| + \left\| D_{p|\mathbf{S}_i}(\mathbf{S})\left\{\mathbf{P}_{0,i}\mathbf{S}_i\right\}\right\| \quad (14)$$

Now, we analyze term by term in eqn. (14). For the first term we take into account that $\sum_{\ell=1}^{\infty}\sum_{k_i=\ell}^{\infty}\mathbf{S}_i^{k_i-1}\mathbf{A}_{k_i} = \sum_{k_i=1}^{\infty}k_i\mathbf{A}_{k_i}\mathbf{S}_i^{k_i-1}$ and we apply the product norm property taking into account that the filters belong to $\mathcal{A}_{L_0}$, which leads to

$$\left\|\mathbf{T}_{0c,i}\sum_{\ell=1}^{\infty}\sum_{k_i=\ell}^{\infty}\mathbf{S}_i^{k_i-1}\mathbf{A}_{k_i}\right\| \leq \|\mathbf{T}_{0c,i}\|\left\|\sum_{k_i=1}^{\infty}k_i\mathbf{A}_{k_i}\mathbf{S}_i^{k_i-1}\right\| \leq L_0\|\mathbf{T}_{0,i}\|. \quad (15)$$

For the second term in eqn. (14) we take into account that (see [1] page 84, [2] page 158 and [4] page 386): $\left\| D_{p|\mathbf{S}_i}(\mathbf{S})\left\{\mathbf{P}_{0,i}\right\}\right\| \leq L_0\|\mathbf{P}_{0,i}\|$ if p is Gâteaux differentiable, which is always true because p is Fréchet differentiable and with the fact that $\|\mathbf{P}_{0,i}\| \leq \delta\|\mathbf{T}_{0,i}\|$, we have $\left\| D_{p|\mathbf{S}_i}(\mathbf{S})\left\{\mathbf{P}_{0,i}\right\}\right\| \leq L_0\delta\|\mathbf{T}_{0,i}\|$.

For the third term in eqn. (14), we take into account that $\sum_{\ell=1}^{\infty}\sum_{k_i=\ell}^{\infty}\mathbf{S}_i^{k_i}\mathbf{A}_{k_i} = \sum_{k_i=1}^{\infty}k_i\mathbf{A}_{k_i}\mathbf{S}_i^{k_i}$ and we apply the norm product property taking into account that the filters belong to $\mathcal{A}_{L_1}$, which leads to

$$\left\|\mathbf{T}_{1c,i}\sum_{\ell=1}^{\infty}\sum_{k_i=\ell}^{\infty}\mathbf{S}_i^{k_i}\mathbf{A}_{k_i}\right\| \leq \|\mathbf{T}_{1c,i}\|\left\|\sum_{k_i=1}^{\infty}k_i\mathbf{A}_{k_i}\mathbf{S}_i^{k_i}\right\| \leq L_1\|\mathbf{T}_{1,i}\|. \quad (16)$$

Finally, for the fourth term we use the notation $\tilde{D}(\mathbf{S})\left\{\mathbf{P}_{1,i}\right\} = D_{p|\mathbf{S}_i}(\mathbf{S})\left\{\mathbf{P}_{1,i}\mathbf{S}_i\right\}$. We start taking into account that (see [5] pages 61 and 331) the eigenvalues of the operator $\tilde{D}(\mathbf{S})$ represented as ζ_{pq} are given by

$$\zeta_{pq} = \begin{cases} \frac{p(\lambda_p)-p(\lambda_q)}{\lambda_p-\lambda_q}\lambda_q & \text{if} \quad \lambda_p \neq \lambda_q \\ \lambda_p p'(\lambda_p) & \text{if} \quad \lambda_p = \lambda_q \end{cases}, \quad (17)$$

then, taking into account that the filters belong to $\mathcal{A}_{L_1}$ we have $\|\tilde{D}(\mathbf{S})\| \leq L_1$, therefore $\left\| D_{p|\mathbf{S}_i}(\mathbf{S})\left\{\mathbf{P}_{1,i}\mathbf{S}_i\right\}\right\| = \left\|\tilde{D}(\mathbf{S})\left\{\mathbf{P}_{1,i}\right\}\right\| \leq L_1\|\mathbf{P}_{1,i}\|$. Additionally, with $\|\mathbf{P}_{1,i}\| \leq \delta\|\mathbf{T}_{1,i}\|$ it follows that $\left\| D_{p|\mathbf{S}_i}(\mathbf{S})\left\{\mathbf{P}_{1,i}\mathbf{S}_i\right\}\right\| \leq L_1\delta\|\mathbf{T}_{1,i}\|$.

Putting all these results together into eqn. (14) we reach

$$\left\| D_{p|\mathbf{S}_i}(\mathbf{S})\left\{\mathbf{T}(\mathbf{S}_i)\right\}\right\| \leq (1+\delta)L_0\|\mathbf{T}_{0,i}\| + (1+\delta)L_1\|\mathbf{T}_{1,i}\|$$
$$\leq (1+\delta)\left(L_0 \sup_{\mathbf{S}_i\in\mathcal{S}}\|\mathbf{T}(\mathbf{S}_i)\| + L_1 \sup_{\mathbf{S}_i\in\mathcal{S}}\|D_{\mathbf{T}}(\mathbf{S}_i)\|\right)$$

■

3 Appendix: Spectral Decompositions

In this section we provide the basic notions about spectral or Fourier decompositions in the context of algebraic signal models. We will see how this is derived from decompositions of a given representation in terms of other *irreducible* representations.

We start with some basic definitions.

Definition 3 *Let $(\mathcal{M},\rho)$ be a representation of $\mathcal{A}$. Then, a representation $(\mathcal{U},\rho)$ of $\mathcal{A}$ is a* ***subrepresentation*** *of $(\mathcal{M},\rho)$ if $\mathcal{U} \subseteq \mathcal{M}$ and $\mathcal{U}$ is invariant under all operators $\rho(a)$* for all *$a \in \mathcal{A}$, i.e. $\rho(a)u \in \mathcal{U}$ for all $u \in \mathcal{U}$ and $a \in \mathcal{A}$. A representation $(\mathcal{M} \neq 0,\rho)$ is* ***irreducible*** *or simple if the only subrepresentations of $(\mathcal{M} \neq 0,\rho)$ are $(0,\rho)$ and $(\mathcal{M},\rho)$.*

The class of irreducible representations of an algebra $\mathcal{A}$ is denoted by $\text{Irr}\{\mathcal{A}\}$. Notice that the zero vector space and $\mathcal{M}$ induce themselves subrepresentations of $(\mathcal{M},\rho)$. In order to state a comparison between representations the concept of *homomorphism between representations* is introduced in the following definition.

Definition 4 *Let $(\mathcal{M}_1,\rho_1)$ and $(\mathcal{M}_2,\rho_2)$ be two representations of an algebra $\mathcal{A}$. A homomorphism or* ***interwining operator*** *$\phi : \mathcal{M}_1 \to \mathcal{M}_2$ is a linear operator which commutes with the action of $\mathcal{A}$, i.e.*

$$\phi(\rho_1(a)v) = \rho_2(a)\phi(v). \tag{18}$$

A homomorphism ϕ is said to be an isomorphism of representations if it is an isomorphism of vectors spaces.

Notice from definition 4 a substantial difference between the concepts of isomorphism of vector spaces and isomorphism of representations. In the first case we can consider that two arbitray vector spaces of the same dimension (finite) are isomorphic, while for representations that condition is required but still the condition in eqn. (18) must be satisfied. For instance, as pointed out in [6] all the irreducible 1-dimensional representations of the polynomial algebra $\mathbb{C}[t]$ are non isomorphic.

As we have discussed before, the vector space $\mathcal{M}$ associated to $(\mathcal{M}, \rho)$ provides the space where the signals are modeled. Therefore, it is of central interest to determine whether it is possible or not to *decompose* $\mathcal{M}$ in terms of simpler or smaller spaces consistent with the action of ρ. We remark that for any two representations $(\mathcal{M}_1, \rho_1)$ and $(\mathcal{M}_2, \rho_2)$ of an algebra $\mathcal{A}$, their direct sum is given by the representation $(\mathcal{M}_1 \oplus \mathcal{M}_2, \rho)$ where $\rho(a)(\mathbf{x}_1 \oplus \mathbf{x}_2) = (\rho_1(a)\mathbf{x}_1 \oplus \rho_2(a)\mathbf{x}_2)$. We introduce the concept of indecomposability in the following definition.

Definition 5 *A nonzero representation $(\mathcal{M}, \rho)$ of an algebra $\mathcal{A}$ is said to be* ***indecomposable*** *if it is not isomorphic to a direct sum of two nonzero representations.*

Indecomposable representations provide the *minimum units of information* that can be extracted from signals in a given space when the filters have a specific structure (defined by the algebra) [7]. The following theorem provides the basic building block for the decomposition of finite dimensional representations.

Theorem 2 (Krull-Schmit, [8]) *Any finite dimensional representation of an algebra can be decomposed into a finite direct sum of indecomposable subrepresentations and this decomposition is unique up to the order of the summands and up to isomorphism.*

The uniqueness in this result means that if $(\oplus_{r=1}^{r} V_i, \rho) \cong (\oplus_{j=1}^{s} W_j, \gamma)$ for indecomposable representations $(V_j, \rho_j), (W_j, \gamma_j)$, then $r = s$ and there is a permutation π of the indices such that $(V_i, \rho_i) \cong (W_{\pi(j)}, \gamma_{\pi(j)})$ [8]. Although theorem 2 provides the guarantees for the decomposition of representation in terms of indecomposable representations, it is not applicable when infinite dimensional representations are considered. However, it is possible to overcome this obstacle taking into account that irreducible representations are indecomposable [6, 8], and they can be used then to build representations that are indecomposable. In particular, *irreducibility* plays a central role to decompose the invariance properties of the images of ρ on $\text{End}(\mathcal{M})$ [8]. Representations that allow a decomposition in terms of subrepresentations that are irreducible are called *completely reducible* and its formal description is presented in the following definition.

Definition 6 ([8]) *A representation $(\mathcal{M}, \rho)$ of the algebra $\mathcal{A}$ is said to be* ***completely reducible*** *if $(\mathcal{M}, \rho) = \bigoplus_{i \in I} (\mathcal{U}_i, \rho_i)$ with irreducible subrepresentations $(\mathcal{U}_i, \rho_i)$. The* ***length*** *of $(\mathcal{M}, \rho)$ is given by $|I|$.*

For a given $(\mathcal{U}, \rho_{\mathcal{U}}) \in \text{Irr}\{\mathcal{A}\}$ the sum of all irreducible subrepresentations of (V, ρ_V) that are equivalent (isomorphic) to $(\mathcal{U}, \rho_{\mathcal{U}})$ is represented by $V(\mathcal{U})$ and it is called the $\mathcal{U}$-homogeneous component of (V, ρ_V). This sum is a direct sum, therefore it has a length that is well defined and whose value is called the *multiplicity* of $(\mathcal{U}, \rho_{\mathcal{U}})$ and is represented by $m(\mathcal{U}, V)$ [8]. Additionally, the sum of all irreducible subrepresentations of (V, ρ_V) will be denoted as $\text{soc}\{V\}$. It is possible to see that a given representation (V, ρ_V) is completely reducible if and only if $(V, \rho_V) = \text{soc}\{S\}$ [8]. The connection between $\text{soc}\{V\}$ and $V(\mathcal{U})$ is given by the following proposition.

Proposition 1 (Proposition 1.31 [8]) *Let $(V, \rho_V) \in \mathsf{Rep}\{A\}$. Then $soc\{V\} = \bigoplus_{S \in Irr\{\mathcal{A}\}} V(S)$.*

Now, taking into account that any homogeneous component $V(\mathcal{U})$ is itself a direct sum we have that

$$\text{soc}\{V\} \cong \bigoplus_{S \in \text{Irr}\{\mathcal{A}\}} S^{\oplus m(\mathcal{U}, V)}. \tag{19}$$

Equation (19) provides the building block for the definition of Fourier decompositions in algebraic signal processing [9]. With all these concepts at hand we are ready to introduce the following definition.

Definition 7 (Fourier Decomposition) *For an algebraic signal model $(\mathcal{A}, \mathcal{M}, \rho)$ we say that there is a spectral or Fourier decomposition of $(\mathcal{M}, \rho)$ if*

$$(\mathcal{M}, \rho) \cong \bigoplus_{(\mathcal{U}_i, \phi_i) \in Irr\{\mathcal{A}\}} (\mathcal{U}_i, \phi_i)^{\oplus m(\mathcal{U}_i, \mathcal{M})} \tag{20}$$

where the $(\mathcal{U}_i, \phi_i)$ are irreducible subrepresentations of $(\mathcal{M}, \rho)$. Any signal $\mathbf{x} \in \mathcal{M}$ can be therefore represented by the map Δ given by

$$\begin{aligned} \Delta: \quad \mathcal{M} \to \bigoplus_{(\mathcal{U}_i,\phi_i)\in Irr\{\mathcal{A}\}} (\mathcal{U}_i, \phi_i)^{\oplus m(\mathcal{U}_i, \mathcal{M})} \\ \mathbf{x} \mapsto \hat{\mathbf{x}} \end{aligned} \tag{21}$$

known as the Fourier decomposition of $\mathbf{x}$ and the projection of $\hat{\mathbf{x}}$ in each $\mathcal{U}_i$ are the Fourier components represented by $\hat{\mathbf{x}}(i)$.

Notice that in eqn. (20) there are two sums, one dedicated to the non isomorphic subrepresentations (external) and another one (internal) dedicated to subrepresentations that are isomorphic. In this context, the sum for non isomorphic representations indicates the sum on the *frequencies* of the representation while the sum for isomorphic representations a sum of components associated to a given frequency. It is also worth pointing out that Δ is an interwining operator, therefore, we have that $\Delta(\rho(a)\mathbf{x}) = \rho(a)\Delta(\mathbf{x})$. As pointed out in [10] this can be used to define a convolution operator as $\rho(a)\mathbf{x} = \Delta^{-1}(\rho(a)\Delta(\mathbf{x}))$. The projection of a filtered signal $\rho(a)\mathbf{x}$ on each $\mathcal{U}_i$ is given by $\phi_i(a)\hat{\mathbf{x}}(i)$ and the collection of all this projections is known as the *spectral representation* of the operator $\rho(a)$. Notice that $\phi_i(a)\hat{\mathbf{x}}(i)$ translates to different operations depending on the dimension of $\mathcal{U}_i$. For instance, if $\dim(\mathcal{U}_i) = 1$, $\hat{\mathbf{x}}(i)$ and $\phi_i(a)$ are scalars while if $\dim(\mathcal{U}_i) > 1$ and finite $\phi_i(a)\hat{\mathbf{x}}(i)$ is obtained as a matrix product.

Remark 1 The spectral representation of an operator indicated as $\phi_i(a)\hat{\mathbf{x}}(i)$ and eqns. (20) and (21) highlight one important fact that is essential for the discussion of the results in Section **??**. For a completely reducible representation $(\mathcal{M}, \rho) \in \mathsf{Rep}\{\mathcal{A}\}$ the connection between the algebra $\mathcal{A}$ and the spectral representation is *exclusively* given by $\phi_i(a)$ which is acting on $\hat{\mathbf{x}}(i)$, therefore, it is not possible by the selection of elements or subsets of the algebra to do any modification on the spaces $\mathcal{U}_i$ associated to the irreducible components in eqn.(20). As a consequence, when measuring the similarities between two operators $\rho(a)$ and $\tilde{\rho}(a)$ associated to $(\mathcal{M}, \rho)$ and $(\mathcal{M}, \tilde{\rho})$, respectively, there will be two sources of error. One source of error that can be modified by the selection of $a \in \mathcal{A}$ and another one that will be associated with the differences between spaces $\mathcal{U}_i$ and $\tilde{\mathcal{U}}_i$, which are associated to the direct sum decomposition of $(\mathcal{M}, \rho)$ and $(\mathcal{M}, \tilde{\rho})$, respectively. This point was first elucidated in [11] for the particular case of GNNs, but it is part of a much more general statement that becomes more clear in the language of algebraic signal processing.

Example 1 (Discrete signal processing) In CNNs the filtering is defined by the polynomial algebra $\mathcal{A} = \mathbb{C}[t]/(t^N - 1)$, therefore, in a given layer the spectral representation of the filters is given by

$$\rho(a)\mathbf{x} = \sum_{i=1}^{N} \phi_i \left(\sum_{k=0}^{K-1} h_k t^k \right) \hat{\mathbf{x}}(i)\mathbf{u}_i$$
$$= \sum_{i=1}^{N} \sum_{k=0}^{K-1} h_k \phi_i(t)^k \hat{\mathbf{x}}(i)\mathbf{u}_i = \sum_{i=1}^{N} \sum_{k=0}^{K-1} h_k \left(e^{-\frac{2\pi ij}{N}} \right)^k \hat{\mathbf{x}}(i)\mathbf{u}_i,$$

with $a = \sum_{k=0}^{K-1} h_k t^k$ and where the $\mathbf{u}_i(v) = \frac{1}{\sqrt{N}} e^{\frac{2\pi jvi}{N}}$ are the column vectors of the traditional DFT matrix, while $\phi_i(t) = e^{-\frac{2\pi ji}{N}}$ is the eigenvalue associated to $\mathbf{u}_i$. Here $\hat{\mathbf{x}}$ represents the DFT of $\mathbf{x}$.

Example 2 (Graph signal processing) Taking into account that the filtering in each layer of a GNN is defined by a polynomial algebra, the spectral representation of the filter is given by

$$\rho(a)\mathbf{x} = \sum_{i=1}^{N} \phi_i \left(\sum_{k=0}^{K-1} h_k t^k \right) \hat{\mathbf{x}}(i)\mathbf{u}_i$$
$$= \sum_{i=1}^{N} \sum_{k=0}^{K-1} h_k \phi_i(t)^k \hat{\mathbf{x}}(i)\mathbf{u}_i = \sum_{i=1}^{N} \sum_{k=0}^{K-1} h_k \lambda_i^k \hat{\mathbf{x}}(i)\mathbf{u}_i \quad (22)$$

with $a = \sum_{k=0}^{K-1} h_k t^k$, and where the $\mathbf{u}_i$ are given by the eigenvector decomposition of $\rho(t) = \mathbf{S}$, where $\mathbf{S}$ could be the adjacency matrix or the Laplacian of the graph, while $\phi_i(t) = \lambda_i$ being λ_i the eigenvalue associated to $\mathbf{u}_i$. The projection of $\mathbf{x}$ in each subspace $\mathcal{U}_i$ is given by $\hat{\mathbf{x}}(i) = \langle \mathbf{u}_i, \mathbf{x} \rangle$, and if $\mathbf{U}$ is the matrix of eigenvectors of $\mathbf{S}$ we have the widely known representation $\hat{\mathbf{x}} = \mathbf{U}^T\mathbf{x}$ [12].

Example 3 (Group signal processing) Considering the Fourier decomposition on general groups [13, 14, 15], we obtain the spectral representation of the algebraic filters as

$$\boldsymbol{a} * \mathbf{x} = \sum_{u,h \in G} \boldsymbol{a}(uh^{-1}) \sum_{i,j,k} \frac{d_k}{|G|} \hat{\mathbf{x}} \left(\boldsymbol{\varphi}^{(k)} \right)_{i,j} \boldsymbol{\varphi}_{i,j}^{(k)}(h)hu,$$

where $\hat{\mathbf{x}}(\boldsymbol{\varphi}^{(k)})$ represents the Fourier components associated to the kth irreducible representation with dimension d_k and $\boldsymbol{\varphi}^{(k)}$ is the associated unitary element. We can see that the kth element in this decomposition is $\sum_{i,j} \mathbf{x}(\varphi^{(k)})_{i,j} \sum_{u,h} \frac{d_k}{|G|} \boldsymbol{a}(uh^{-1})\boldsymbol{\varphi}^{(k)}_{i,j}(h)hu$.

Example 4 (Graphon signal processing) According to the spectral theorem [16, 17], it is possible to represent the action of a compact normal operator S as $S\mathbf{x} = \sum_i \lambda_i \langle \boldsymbol{\varphi}_i, \mathbf{x}\rangle \boldsymbol{\varphi}_i$ where λ_i and $\boldsymbol{\varphi}_i$ are the eigenvalues and eigenvectors of S, respectively, and $\langle \cdot \rangle$ indicates an inner product. Then, the spectral representation of the filtering of a signal in the layer ℓ is given by

$$\rho_\ell\left(p(t)\right)\mathbf{x} = \sum_i p(\lambda_i)\langle \mathbf{x}, \boldsymbol{\varphi}_i\rangle \boldsymbol{\varphi}_i = \sum_i \phi_i(p(t))\hat{\mathbf{x}}_i \boldsymbol{\varphi}_i,$$

where $\phi_i(p(t)) = p(\lambda_i)$.

References

[1] Y. Benyamini and J. Lindenstrauss. *Geometric Nonlinear Functional Analysis*. Number v. 48, no. 1 in American Mathematical Society colloquium publications. American Mathematical Society, 2000.

[2] J. Lindenstrauss, D. Preiss, and J. Tišer. *Frechet Differentiability of Lipschitz Functions and Porous Sets in Banach Spaces*. Annals of Mathematics Studies. Princeton University Press, 2012.

[3] M.S. Berger. *Nonlinearity and Functional Analysis: Lectures on Nonlinear Problems in Mathematical Analysis*. Pure and Applied Mathematics. Elsevier Science, 1977.

[4] Joram Lindenstrauss, David Preiss, and Jaroslav Tivser. Fréchet differentiability of lipschitz functions via a variational principle. *Journal of the European Mathematical Society*, 12:385–412, 2010.

[5] N.J. Higham. *Functions of Matrices: Theory and Computation*. Other Titles in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 2008.

[6] P.I. Etingof, O. Golberg, S. Hensel, T. Liu, A. Schwendner, D.V. , and E. Yudovina. *Introduction to Representation Theory*. Student mathematical library. American Mathematical Society, 2011.

[7] M. Barot. *Introduction to the Representation Theory of Algebras*. Springer International Publishing, 2014.

[8] M. Lorenz. *A Tour of Representation Theory*. Graduate studies in mathematics. American Mathematical Society, 2018.

[9] M. Puschel and J. M. F. Moura. Algebraic signal processing theory: Foundation and 1-d time. *IEEE Transactions on Signal Processing*, 56(8):3572–3585, Aug 2008.

[10] Markus Pschel and Jos M. F. Moura. Algebraic signal processing theory, 2006.

[11] Fernando Gama, Joan Bruna, and Alejandro Ribeiro. Stability properties of graph neural networks, 2019.

[12] F. Gama, A. G. Marques, G. Leus, and A. Ribeiro. Convolutional neural network architectures for signals supported on graphs. *IEEE Transactions on Signal Processing*, 67(4):1034–1049, 2019.

[13] B. Steinberg. *Representation Theory of Finite Groups: An Introductory Approach*. Universitext. Springer New York, 2011.

[14] A. Terras. *Fourier Analysis on Finite Groups and Applications*. London Mathematical Society Student Texts. Cambridge University Press, 1999.

[15] W. Fulton, W.F.J. Harris, and J. Harris. *Representation Theory: A First Course*. Graduate Texts in Mathematics. Springer New York, 1991.

[16] J.B. Conway. *A Course in Functional Analysis*. Graduate Texts in Mathematics. Springer New York, 1994.

[17] C.D. Aliprantis. *An Invitation to Operator Theory*. Graduate studies in mathematics. American Mathematical Society, 2002.