



Article

Approximate Diagonal Integral Representations and Eigenmeasures for Lipschitz Operators on Banach Spaces

Ezgi Erdoğan ¹  and Enrique A. Sánchez Pérez ^{2,*} 

¹ Department of Mathematics, Faculty of Art and Science, University of Marmara, Kadıköy, Istanbul 34722, Turkey; ezgi.erdogan@marmara.edu.tr

² Instituto Universitario de Matemática Pura y Aplicada, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain

* Correspondence: easanpce@mat.upv.es

Abstract: A new stochastic approach for the approximation of (nonlinear) Lipschitz operators in normed spaces by their eigenvectors is shown. Different ways of providing integral representations for these approximations are proposed, depending on the properties of the operators themselves whether they are locally constant, (almost) linear, or convex. We use the recently introduced notion of eigenmeasure and focus attention on procedures for extending a function for which the eigenvectors are known, to the whole space. We provide information on natural error bounds, thus giving some tools to measure to what extent the map can be considered diagonal with few errors. In particular, we show an approximate spectral theorem for Lipschitz operators that verify certain convexity properties.

Keywords: eigenmeasure; operators; banach space; eigenvalue; spectral function



Citation: Erdoğan, E.; Sánchez Pérez, E.A. Approximate Diagonal Integral Representations and Eigenmeasures for Lipschitz Operators on Banach Spaces. *Mathematics* **2022**, *10*, 220. <https://doi.org/10.3390/math10020220>

Academic Editors: Alexander Felshtyn and Luigi Rodino

Received: 5 December 2021

Accepted: 10 January 2022

Published: 12 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Diagonalization of operators is deeply tied to the linear character of the operators themselves. Indeed, interest in obtaining the eigenvectors of a given linear map implies that we can use them, together with the corresponding eigenvalues, to obtain an easy representation of the operator. If we think of the finite dimensional case, all symmetric operators are diagonalizable, so that for all of them one can find, after a change of basis, a diagonal form that gives the simplest representation of the map. Thus, a fruitful diagonalization theory has to provide a way to compute the eigenvectors, as well as a structural rule to extend the representation of the map (or at least an approximation) from the eigenvectors to the whole space. This is the point of view from which we develop the research presented in this paper: we consider the set of eigenvalues of a given map and use a given rule to approximate the original map starting from them.

What is fundamental in the present work is that we do not assume linearity, at least in its “full version”. Instead, we will use metric notions, approximating the value of the original endomorphism $T : E \rightarrow E$ in the Banach space E over complex field by using the “diagonal behavior” of functions at other points. The proximity of any other vectors in E to such points will then be used to find suitable approximate expressions. This is why the Lipschitz character of the functions involved is required, since we have to perform the extension/approximation formulas from the set of eigenvectors $ev(T)$ of T to the whole normed space E . On the other hand, to be Lipschitz continuous seems to be a natural requirement for talking on the spectrum of a non-linear operator (see [1], Introduction, p. 3). Much research has been done for finding results on the spectrum of a Lipschitz operator that could be similar to those that hold in the linear case [1].

Of course, this topic is not new [1,2]. A lot of research has been done, from the abstract and the applied points of view, to understand what diagonalization would be for non-linear maps, and a lot of applied procedures to solve concrete issues appearing in relation to this have also been published (which are normally centered on “almost”

linear cases, as the ones that appear in linear processes with non-linear perturbations). Recently, diagonalization of bilinear maps has been attracting attention [3,4], although it is also a classical topic (see [5]). As in the classical works regarding these problems, we will deal with the point spectrum of a given operator, that is, the set of values that satisfy the equation $T(x) = \lambda \cdot x$ for a non-trivial $x \in E$. However, it must be said that the theoretical investigation on this type of problem has focused the attention on the non-linear spectral theory, which is, in general, richer than the study of the equation $f(x) = \lambda \cdot x$. Although the early works on the topic [6,7] were centered on the analysis of the pointwise eigenvalue equation for non-linear maps, after the work of Kachurovskij [8], the theoretical research has been directed to the abstract analysis of the different notions of spectrum that can be formulated for non-linear maps, and concretely for Lipschitz continuous maps (the Rhodius, Neuberger, and Kachurovskij spectra, among others [1] (Chapter 5), that is, the general subject of the present paper. In this sense, in this paper we focus on the point spectrum. We will restrict our attention to functions of real Banach spaces and only consider real eigenvalues. This is an important restriction for the results we present since in many cases we will not be able to obtain complete diagonal representations; of course, even in the linear case complex eigenvalues could appear. The interested reader can find a complete explanation of these subjects in [1] (see also [9]). Some relevant results connecting these notions with the extensions of the concepts of numerical range/numerical index, to non-linear operators, are also known, showing the interest in this relation even in the non-linear case, for example, it is known that the Kachurovskij spectrum (5.9) of a Lipschitz continuous operator is contained in the convex closure of the numerical range [10] (Theorem 2). A lot of research is still been developed in this direction [11,12].

Additionally, non-linear eigenvalue problems are still attracting attention, mainly due to the large class of applications that these results find in other mathematical fields (see, for example, [1] (Chapter 12)). The proposed formalism can also be used as an analytic tool for models in geometry and physics, mainly those in which vector-valued integrals allow the description of mechanic properties of materials (moment of inertia, center of mass, ...). A recent review on related topics and applications of non-linear diagonalization can be found in [13].

So, due to these potential applications, in this paper we focus attention on low-dimensional eigenvalue problems for non-linear operators. A lot of references of this kind of research can be found in the scientific literature on this topic. For example, some probabilistic methods for finding estimates of eigenvalues and eigenvectors have been developed. This point of view is interesting for us since our interest lies in measure and integration theory, which supports probabilistic interpretations. This line of research comes from computational mathematics and physics. One relevant example is the so-called stochastic diagonalization, which appeared at the end of the XX Century as a computational tool in quantum mechanics, as explained in [14]. The main idea is to use Monte Carlo methods for the computation of eigenvectors of the Hamiltonian of physical systems; under the label *probabilistic diagonalization*, these ideas can be found in [15] (Section 4) and [16–18]. In general, in real-world applications Monte Carlo sampling provides satisfactory computational results, which allow one to develop sophisticated procedures in the application of quantum mechanics, for example, in quantum chemistry (see, for example, [19–21]). In this paper, we intend to provide some new theoretical ideas for finding approximate representations of Lipschitz maps when some eigenvectors are known and to show some examples and concrete situations for which these formulas can be explicitly given (we call the specific methods for getting them “rules”); we show, for example, an approximate spectral theorem for Lipschitz maps based on convexity in Section 5.

The present paper is divided in six sections. After this Introduction, we describe, in Section 2, our goal: to write the linear case using our tools and some fundamental results for Lipschitz operators. In Section 3, we show some general examples that, in a sense, intend to illustrate the generality of different normal situations that we are considering. Section 4 is devoted to the case of a concrete extension procedure, which we call the “proximity

rule”, and in Section 5 we explain what we call the “convexity rule”. We finish the paper by providing some hints (Section 6) on how effective computations can be performed for finding approximations to eigenvectors and eigenvalues of Lipschitz maps.

As a main theoretical framework, we adopt the usual mathematical context of the non-linear spectral theory; we use [1] as reference book for this. Standard definitions and arguments of measure and integration theory, as well as Banach spaces, will be used [22–24]. Our main reference on Lipschitz maps is [25]. For the aim of simplicity, we will consider only the real case. If E is a Banach space with norm $\| \cdot \|$, we write E^* for its dual space, which in the (finite dimensional) Euclidean space and in the (real) Hilbert space can be identified with E . We will write x for the vectors in E and x^* for the elements of E^* . The unit ball of the dual space B_{E^*} is a convex set that is compact with respect to the weak* topology (Alaoglu’s Theorem; see, for example, [22]). Then, we can consider the space $\mathcal{M}(B_{E^*})$ of regular Borel measures on B_{E^*} , which will play a relevant role in the present paper. Let us recall that, for any $x^* \in E^*$, the Dirac’s delta δ_{x^*} belongs to this space. In the same way, if E is a dual space or E is reflexive, we will write $\mathcal{B}(B_E)$ for the corresponding Borel sets. This notion can be extended to any norm bounded subset of E instead of B_E by including it in a larger ball. As usual, we denote by S_E to the unit sphere of E .

2. Approximation Results for Linear and Non-Linear Maps and First Results for Lipschitz Functions

The aim of this section is to explain the motivation of our investigation and to provide the first general results. Let us start by presenting how we can rewrite the well-known context of diagonalization or linear operators in terms of integral expressions, under the formalism provided by the recently introduced eigenmeasures [26]. We expect to find the best diagonal approximation to a given (linear and continuous) operator $T : E \rightarrow E$. Following the mentioned paper, we introduce first a measure on B_{E^*} . We start by considering a (regular Borel) probability measure $\nu \in \mathcal{P}(B_{E^*})$. Roughly speaking, this gives formal support for those elements of the dual space that we want to keep “active”, i.e., those whose actions on the elements of E we are interested in “keeping running”. In the linear case, this set can actually be quite small; in the case of dual Banach spaces with unconditional basis (or finite-dimensional spaces), we choose the simplest: take an unconditional normalized basis $\{e_k^*\}_{k=1}^{+\infty}$ of the dual space and define ν as

$$\nu(A) = \sum_{k=1}^{+\infty} \frac{1}{2^k} \delta_{e_k^*}, \quad A \in \mathcal{B}(B_{E^*}) \text{ if } E \text{ is infinite dimensional,}$$

and

$$\nu(A) = \frac{1}{n} \sum_{k=1}^n \delta_{e_k^*}, \quad A \in \mathcal{B}(B_{E^*}) \text{ if } E \text{ has dimension } n.$$

Since case E is a (real) separable Hilbert space or a finite dimensional Euclidean space, we choose the same measures, but the basis are assumed to be also *orthonormal*. In the general case, as will be shown later on, this formalism provides a useful tool for choosing renormings for the Hilbert space that is considered by means of an inclusion $x \ni E \mapsto \langle x, \cdot \rangle \in L^2(\nu)$, and so every element x is identified with the function $\langle x, \cdot \rangle$ acting on the elements of the unit ball of the dual space. It is clearly seen that this map is always continuous: we also need it to be injective to obtain an alternate representation of the original space. Throughout the paper, we will write $\| \cdot \|_\nu$ for the norm induced on E by the norm of $L^2(\nu)$, that is, $\|x\|_\nu = (\int_{B_{E^*}} |\langle x, \cdot \rangle|^2 d\nu(\cdot))^{1/2}, x \in E$.

Let us fix now the measure structure on the measurable space $(B_E, \mathcal{B}(B_E))$ that is needed to complete the integral diagonal representation of the approximation to the operator T . Let us recall the definition of weak eigenmeasure (Definition 2 in [26]), adapted for the case of linear maps.

Definition 1. Let E be a reflexive Banach space and let $T : E \rightarrow E$ be an operator. Fix a Borel regular measure ν on B_{E^*} , and let μ be a Borel regular measure on a bounded subset of E . We say that μ is a weak eigenmeasure for T if for each $N \in \mathcal{B}(B_{E^*})$ there is a μ -integrable function $x \mapsto \lambda_N(x)$ such that

$$\int_N \langle T(x), x^* \rangle d\nu = \lambda_N(x) \int_N \langle x, x^* \rangle d\nu, \quad \mu\text{-a.e. } x \in B_E,$$

and the functions $\lambda_N(x)$ are uniformly bounded in N and $x \in E$.

In this paper, we will use a stronger version of this notion, which has been also introduced in [26]. We say that a probability measure satisfying the requirement above is an eigenmeasure if $\lambda_N(x)$ does not depend on N , that is, it has a constant value once x is fixed. Note that the linearity requirement for T is not necessary for this definition. The same notion can be defined in the same way for the case of Lipschitz maps, and in general for any function that satisfies the conditions necessary for all integrals to make sense.

It can be seen that the existence of an eigenmeasure μ implies a diagonal representation μ -a.e. That is, for μ -a.e., all elements $x \in B_E$, once the set N is fixed, we have a diagonal representation of the operator. Indeed, consider the set $ev(T)$ of all the vectors that satisfy that there is a scalar value λ such that $T(x) = \lambda x$: we have that, if μ is an eigenmeasure, $\mu(A \setminus ev(T)) = 0$ for every measurable set $A \subset E$. In this paper, we intend to show to what extent this tool can be used to obtain quasi-diagonal representations of T for the case of linear operators. We will tackle this question in the following subsection for the canonical case, i.e., the Euclidean case in which ν is given by the average of Dirac’s deltas of the elements of an orthonormal basis.

2.1. Motivation: Minimization Procedure and Calculation of the Best Diagonal Approximation for a Linear Operator in a Euclidean Space

In this section, we show the canonical case, which will allow us to explain the geometric meaning of the diagonal approximation of an operator. Let us take the real n -dimensional Euclidean space E . Its dual space can be identified with E , so we do not write the superscript $*$ for any space or any element of the orthonormal basis. Thus, for a given orthogonal basis $\{e_1, \dots, e_n\}$ fix the measure

$$\nu(A) := \frac{1}{n} \sum_{k=1}^n \delta_{e_k}, \quad A \in \mathcal{B}(B_{E^*}).$$

Note that, for this case, we have that for every $x \in E$,

$$\|x\|_\nu = \left(\int_{B_{E^*}} |\langle x, \cdot \rangle|^2 d\nu(\cdot) \right)^{1/2} = \left(\frac{1}{n} \sum_{i=1}^n |x_i|^2 \right)^{1/2} = \sqrt{\frac{1}{n}} \|x\|,$$

where $x_k, k = 1, \dots, n$, are the coordinates of x in the fixed basis. Let us compute the optimal value for every element x_0 of the sphere S_E , which is represented by the Dirac’s measure $x_0 \mapsto \delta_{x_0}$. The stochastic eigenvalue equation is then written, for $x_0 \in S_E$ and every $N \in \mathcal{B}(B_E)$, as

$$\begin{aligned} \frac{1}{n} \sum_{e_k \in N} \langle T(x_0), e_k \rangle &= \int_{B_E} \left(\int_N \langle T(x), x^* \rangle d\nu(x^*) \right) d\delta_{x_0} \\ &= \int_{B_E} \left(\lambda_N(x) \int_N \langle x, x^* \rangle d\nu \right) d\delta_{x_0} = \lambda_N(x_0) \frac{1}{n} \sum_{e_k \in N} \langle x_0, e_k \rangle. \end{aligned}$$

For example, for $N = \{e_k\}, k = 1, \dots, n$, we obtain the equations

$$\langle T(x_0), e_k \rangle = \lambda_k \cdot \langle x_0, e_k \rangle \quad \text{for } \lambda_k = \lambda_{\{e_k\}}(x_0),$$

what gives $\lambda_k = \frac{\langle T(x_0), e_k \rangle}{\langle x_0, e_k \rangle}$ if $\langle x_0, e_k \rangle \neq 0$, and $\lambda_k = 0$ otherwise.

Let us show an estimate of the best λ for which the equation $T(x) = \lambda(x) x$ holds for a given $x \in E$ in the integral setting that has been fixed above. In other words, we want to minimize the error

$$\varepsilon(x) = \|T(x) - \lambda x\|_v^2 = \int_{B_E} \left| \langle T(x), x^* \rangle - \lambda \langle x, x^* \rangle \right|^2 dv(x^*).$$

The result has a clear geometric meaning: *once the point $x \in S_E$ is fixed, the best diagonal approximation to $T(x)$ is the projection of $T(x)$ on the subspace generated by x . It is given by a direct minimization argument, summarized in the next result.*

Proposition 1. *Consider the measure ν fixed at the beginning of this section for the n -dimensional Euclidean space. Let $x \in E$. Then, the following assertions hold.*

(i) *The best diagonal approximation to T in x is given by $x \mapsto \lambda(x) x$, where*

$$\lambda(x) = \frac{\langle T(x), x \rangle}{\|x\|^2} = \frac{1}{\|x\|^2} \sum_{k=1}^n \langle T(x), e_k \rangle \langle x, e_k \rangle = \frac{1}{\|x\|^2} \sum_{k=1}^n \langle T(x), e_k \rangle x_k.$$

(ii) *The minimal integral error $\varepsilon(x)$ is given by*

$$\varepsilon(x) = \frac{1}{n} \left(\|T(x)\|^2 - \frac{\langle T(x), x \rangle^2}{\|x\|^2} \right).$$

Proof.

(i) Since the basis $\{e_k\}_{k=1}^n$ is orthonormal, we have that the error $\varepsilon(x)$ can be written as

$$\int_{B_E} \left| \langle T(x), x^* \rangle - \lambda \langle x, x^* \rangle \right|^2 dv(x^*) = \frac{1}{n} \sum_{k=1}^n \left| \langle T(x), e_k \rangle - \lambda \langle x, e_k \rangle \right|^2 = \frac{1}{n} \|T(x) - \lambda x\|^2.$$

Therefore, we can rewrite the error as

$$\varepsilon(x)(\lambda) = \frac{1}{n} \left(\langle T(x) - \lambda x, T(x) - \lambda x \rangle \right) = \frac{1}{n} \left(\|T(x)\|^2 - 2\lambda \langle T(x), x \rangle + \lambda^2 \langle x, x \rangle \right),$$

which gives

$$\varepsilon(x)(\lambda) = \frac{1}{n} \left(\|T(x)\|^2 - 2\lambda \langle T(x), x \rangle + \lambda^2 \|x\|^2 \right).$$

Recall that the vector x is fixed. The critical points equation involving the derivative with respect to λ are

$$\frac{d\varepsilon(x)}{d\lambda} = -\frac{2}{n} \langle T(x), x \rangle + \frac{2}{n} \lambda \|x\|^2 = 0,$$

that is, $\lambda = \langle T(x), x \rangle / \|x\|^2$. Since $\frac{d^2\varepsilon(x)}{d\lambda^2} = 2\|x\|^2/n > 0$, this value of λ gives a minimum. Therefore, the best diagonal approximation for $T(x)$ is $x \mapsto (\langle T(x), x \rangle / \|x\|^2) \cdot x$.

(ii) Just writing this value of λ in the expression of the error, we find that the minimal value of the error at x is given by

$$\begin{aligned} \varepsilon(x) &= \frac{1}{n} \left(\|T(x)\|^2 - 2 \frac{\langle T(x), x \rangle \langle T(x), x \rangle}{\|x\|^2} + \frac{\langle T(x), x \rangle^2}{\|x\|^4} \|x\|^2 \right) \\ &= \frac{1}{n} \left(\|T(x)\|^2 - \frac{\langle T(x), x \rangle^2}{\|x\|^2} \right). \quad \square \end{aligned}$$

The best approximation to a diagonal form for any vector x in the space given by $\lambda(x)$ provides a direct way of representation of an operator. Of course, the only vectors for which the error ε equals 0 are the eigenvectors of the operator T ; clearly, we have that a linear operator $T : E \rightarrow E$ is diagonalizable if and only if there is a basis $\{x_k\}_{k=1}^n$ such that $\varepsilon(x_k) = 0$ for all $k = 1, \dots, n$.

Let us remark that, under the assumption of the existence of an orthonormal basis of eigenvectors for T , an integral representation formula is provided: take such a basis $\{e_k\}_{k=1}^n$ and consider an associated eigenmeasure μ that coincides with the measure ν that have been considered above. Then, for every $x_0 \in E$ we have that

$$T(x_0) = \sum_{k=1}^n \lambda(e_k) \langle x_0, e_k \rangle e_k = \int_{B_E} n \langle T(x), x \rangle \langle x_0, x \rangle x d\mu(x),$$

where $\mu(\cdot) = \frac{1}{n} \sum_{i=1}^n \delta_{e_i}(\cdot)$. Therefore, we can write a representation of a diagonalizable operator T as a Bochner integral average associated with the diagonal function $x \mapsto \lambda(x) \cdot x$. This is the linear version of the main idea that we use in the subsequent sections of the paper: even for non-linear cases, an integral approximation to a function can be obtained as an integral average, once a(n) (approximated) eigenmeasure is known.

2.2. General Finite Dimensional Case: Local Diagonal Approximation for Functions $T : E \rightarrow E$

In this section, we show the general equations for the finite dimensional case, without linearity assumptions: no restriction is imposed on the function T , apart from some weak measurability properties. For the sake of clarity, we revert to the notation E for the original finite dimensional space and E^* for its dual, which is isomorphic to it because it is finite dimensional.

Let us fix again the measures for getting an integral representation of an equivalent norm for E . Take regular Borel probability measures ν and μ on the spaces S_{E^*} and S_E , respectively. Suppose that the probability measure ν separates the points of S_E , that is, it satisfies that if $x \in S_E$ we have $\int_{B_{E^*}} |\langle x, x^* \rangle|^2 d\nu(x^*) \neq 0$. This requirement implies the injectivity of the representation provided by the norm in $L^2(\nu)$ that was explained at the beginning of this section, and it is enough to get that the norm $\|x\|_\nu = (\int_{B_{E^*}} |\langle \cdot, x^* \rangle|^2 d\nu(x^*))^{1/2}$ gives an equivalent norm for the space $(E, \|\cdot\|)$. Let us write $(E_\nu, \|\cdot\|_\nu)$ for the resulting normed space, which inherits a Hilbert space structure from the inclusion $E \hookrightarrow L^2(\nu)$ given by $x \mapsto \langle x, \cdot \rangle$. This is the key for the introduction of stochastic arguments for the estimate of diagonal approximations. However, note that the dual space (and thus its unit ball) that we consider in this section is endowed with the natural norm of $(E, \|\cdot\|)^*$, that is, $\|\cdot\|_{E^*} = \sup_{x \in B_E} |\langle x, \cdot \rangle|$.

Let us start as in the linear case, considering the diagonal error, that is, the measure of how far the operator is at point x of being diagonal. Given a vector $x \in S_{E_\nu}$, the error committed when $T(x)$ is approximated as a line following the direction of x is

$$\varepsilon(x) = \int_{B_{E^*}} \left| \langle T(x), x^* \rangle - \lambda \langle x, x^* \rangle \right|^2 d\nu(x^*).$$

This provides a simplified error formula, which generalizes the geometric idea based on the projection on x of $T(x)$ explained in the previous section. The computations that give the proof are similar to the ones that prove Proposition 1 but consider the duality relation in $L^2(\nu)$ instead of that in the original space E .

Proposition 2. *Let $x \in E$ be fixed. Then, the error $\varepsilon(x)$ attains its minimum value for*

$$\lambda(x) = \frac{1}{\|x\|_\nu^2} \int_{B_{E^*}} \langle T(x), x^* \rangle \langle x, x^* \rangle d\nu(x^*)$$

and has the minimum value for this $\lambda(x)$ given by

$$\begin{aligned} \varepsilon(x) &= \int_{B_{E^*}} |\langle T(x), x^* \rangle|^2 d\nu(x^*) - \left(\frac{1}{\|x\|_v^2} \int_{B_{E^*}} \langle T(x), x^* \rangle \langle x, x^* \rangle d\nu(x^*) \right)^2 \\ &= \|T(x)\|_v^2 - \left(\frac{1}{\|x\|_v^2} \int_{B_{E^*}} \langle T(x), x^* \rangle \langle x, x^* \rangle d\nu(x^*) \right)^2. \end{aligned}$$

We are ready to give the local version of the diagonal approximation for T . Recall that we are searching for an approximation provided by an integral average involving measures on the Borel subsets of E , mainly eigenmeasures. This makes it relevant to obtain integral formulas for subsets A that, in the next sections, will be neighborhoods of the point on which we focus our attention.

Thus, given a measurable subset $A \subset E$, the error committed when the function T is approximated as a diagonal integral average is given by

$$\varepsilon(A) = \int_A \left(\int_{B_{E^*}} |\langle T(x), x^* \rangle - \lambda \langle x, x^* \rangle|^2 d\nu(x^*) \right) d\mu.$$

The proof of the next result also follows the same arguments than the ones of Propositions 1 and 2.

Proposition 3. *Let $A \subseteq E$ be a non- μ -null measurable set. Then, the minimal average error $\varepsilon(A)$ for the diagonal approximation is given by the formula*

$$\varepsilon(A) = \int_A \|T(x)\|_v^2 d\mu(x) - \frac{1}{\int_A \|x\|_v^2 d\mu} \left(\int_A \left(\int_{B_{E^*}} \langle T(x), x^* \rangle \langle x, x^* \rangle d\nu(x^*) \right) d\mu(x) \right)^2,$$

and it is attained for the value

$$\lambda(A) = \frac{1}{\int_A \|x\|_v^2 d\mu} \int_A \left(\int_{B_{E^*}} \langle T(x), x^* \rangle \langle x, x^* \rangle d\nu(x^*) \right) d\mu.$$

Note that, for $A \subset S(E)$, this formula can be rewritten as an integral average, $\lambda(A) = \frac{1}{\mu(A)} \int_A \left(\int_{B_{E^*}} \langle T(x), x^* \rangle \langle x, x^* \rangle d\nu(x^*) \right) d\mu$. This expression suggests the arguments used in the next sections.

2.3. Approximation Formulas Based on Eigenmeasures: Bochner Integration for Lipschitz Maps

Let us present the integral approximation procedure that can be constructed by adapting the previous framework for the case of Lipschitz functions. Thus, let us describe now how to get an approximation based on eigenmeasures in this case. In [26], it is shown, for the case of bilinear maps, how the calculus of eigenmeasures can be used to deal with non-linear diagonalization. Here, we assume that such an eigenmeasure μ is given for a certain non-linear map $f : E \rightarrow E$, where E is a (finite dimensional) Euclidean space, and we show how we can use it to obtain an approximation to f . In order to do this, we change the context of scalar integration used before, which has been useful for the computation of the error and to obtain the best diagonal pointwise approximation for a given map, by the vector-valued Bochner integration.

The idea is to consider f as a Bochner (locally) integrable function (see, for example, [23], Chapter II). In order to do this, if $f : E \rightarrow E$ is a strongly measurable function (see [23], Chapter II), we consider the vector-valued function $A \subset E \rightarrow E$ given by $x \mapsto f(x)$ as a Bochner integrable function with respect to the measure μ . Integration of such a function provides a (countably additive) vector measure, which assures convergence of the limits of the integrals on sequences of disjoint sets (see [23,24]). In order to get the desired

integral representation, we only need f to be bounded in the sets A , which will be a certain class of neighborhoods of the points in the Euclidean space E .

Definition 2. Let $f : E \rightarrow E$ be a Lipschitz function, and let μ be an eigenmeasure for f . We define the eigenmeasure approximation of a strongly measurable Lipschitz function for a given $\epsilon > 0$, under the requirement $\mu(B_\epsilon(x_0)) > 0$, by the formula

$$\widehat{f}_\epsilon(x_0) := \frac{1}{\mu(B_\epsilon(x_0))} \int_{B_\epsilon(x_0)} f(x) d\mu(x), \quad x_0 \in E.$$

Let us write $\lambda_f(\cdot)$ and $\lambda_{\widehat{f}_\epsilon}(\cdot)$ for the best approximated eigenvalue fixed for the functions f and \widehat{f}_ϵ , respectively.

We need a previous lemma for the computation of the error committed when we approximate the value of $f(x)$ by its projection on the subspace defined by other vector x_0 . It is similar to Proposition 2, but in this case we are considering two different points, x_0 and x .

Lemma 1. Let $x_0, x \in E$ and $f : E \rightarrow E$ satisfying all the requirements of integrability explained above. The best value of λ that minimizes the error

$$\varepsilon_{x_0}(x) = \|f(x) - \lambda x_0\|_v^2 = \int_{B_E} |\langle f(x) - \lambda x_0, x^* \rangle|^2 d\nu(x^*)$$

committed when we approximate $f(x)$ by its projection on the subspace generated by x_0 is given by

$$\lambda_{x_0}(x) = \frac{1}{\|x_0\|_v^2} \int_{B_E} \langle x_0, x^* \rangle \langle f(x), x^* \rangle d\nu(x^*).$$

Proof. The computations are again similar to the ones that prove Proposition 1. We have that

$$\begin{aligned} \varepsilon_{x_0}(x) &= \int_{B_E} |\langle f(x) - \lambda x_0, x^* \rangle|^2 d\nu(x^*) \\ &= \|f(x)\|_v^2 - 2\lambda \int_{B_E} \langle x_0, x^* \rangle \langle f(x), x^* \rangle d\nu(x^*) + \lambda^2 \int_{B_E} \langle x_0, x^* \rangle^2 d\nu(x^*), \end{aligned}$$

and so, by computing the value of the derivative with respect to λ to be equal to 0, we obtain

$$\lambda_{x_0}(x) = \frac{1}{\|x_0\|_v^2} \int_{B_E} \langle x_0, x^* \rangle \langle f(x), x^* \rangle d\nu(x^*).$$

□

Remark 1. Lemma 1 allows one to introduce a new concept, in the line of the integral means that the eigenmeasures formalism provides. If $x_0 \in E$, we can consider the integral average $\lambda_{x_0}^\epsilon$ of the pointwise diagonal approximation to f , which is given by

$$\begin{aligned} \lambda_{x_0}^\epsilon &= \frac{1}{\mu(B_\epsilon(x_0))} \int_{B_\epsilon(x_0)} \lambda_{x_0}(x) d\mu \\ &= \frac{1}{\|x_0\|_v^2 \mu(B_\epsilon(x_0))} \int_{B_\epsilon(x_0)} \int_{B_E} \langle x_0, x^* \rangle \langle f(x), x^* \rangle d\nu(x^*) d\mu(x), \end{aligned}$$

where the function $\lambda_{x_0}(x)$ is given in Lemma 1.

This value does not allow one to find a new eigenvector but provides the natural way of controlling how far a given vector x_0 is from being an eigenvector for the function f . Next result covers the relevant information about the behavior of the eigenmeasure approximation introduced in this section: expected error, both in the pointwise approximation to

the function f and in the corresponding eigenvalues, and bounds for the distance among the value of the approximating at a given point x_0 and its optimal diagonal form. Note that even for a linear and continuous operator, it may happen that the set of eigenvectors is not sufficient to represent the operator by eigenmeasures: a rotation in \mathbb{R}^2 does not have eigenvectors.

Theorem 1. *Let ν be a regular Borel probability measure on $(B_E, \mathcal{B}(B_{E^*}))$. Consider a Lipschitz map $f : E \rightarrow E$ with Lipschitz constant $Lip(f)$ and assume that $ev(f)$ is non-trivial. Consider an eigenmeasure μ for f , and suppose that f is a strongly measurable function with respect to μ .*

Fix a vector $x_0 \in E$, and let $\epsilon > 0$ such that $\mu(B_\epsilon(x_0)) > 0$. Then,

- (1) $\|\widehat{f}_\epsilon(x_0) - f(x_0)\|_\nu \leq Lip(f) \cdot \epsilon.$
- (2) $|\lambda_f(x_0) - \lambda_{\widehat{f}_\epsilon}(x_0)| \leq \frac{Lip(f)}{\|x_0\|} \cdot \epsilon.$
- (3) *The distance with respect to the diagonal approximation to f , $\|\widehat{f}_\epsilon(x_0) - \lambda_{x_0}^\epsilon x_0\|_\nu$, is bounded by $\frac{1}{\mu(B_\epsilon(x_0))} \left(\int_{B_\epsilon(x_0)} (\|f(x)\|_\nu^2 - \lambda_{x_0}^2(x) \cdot \|x_0\|^2) d\mu(x) \right).$*

Proof. Again a rewriting of the arguments around Theorem 1 gives the result. Let us show them step by step.

- (1) It is just a consequence of the integral domination provided by the Lipschitz inequality for f . Indeed,

$$\begin{aligned} \|\widehat{f}_\epsilon(x_0) - f(x_0)\|_\nu &= \left\| \frac{1}{\mu(B_\epsilon(x_0))} \int_{B_\epsilon(x_0)} (f(x) - f(x_0)) d\mu(x) \right\| \\ &\leq \frac{1}{\mu(B_\epsilon(x_0))} \sup_{x \in B_\epsilon(x_0)} \|f(x) - f(x_0)\| \mu(B_\epsilon(x_0)) \leq Lip(f) \cdot \epsilon. \end{aligned}$$

- (2) Let us recall that for any function f , and at any point x_0 , the optimal value for λ is $\lambda_f(x_0) = \int_{B_E} \langle f(x_0), x^* \rangle \langle x_0, x^* \rangle d\nu(x^*) / \|x_0\|_\nu^2$. We have to use the property of the Bochner integration that assures that, if ϕ is a Bochner integrable function with respect to a measure η , $\langle \int \phi d\eta, x^* \rangle = \int \langle \phi, x^* \rangle d\eta$ (see, for example, [27] (Lemma 11.44), or [23] (Chapter II)). Consequently,

$$\begin{aligned} \lambda_{\widehat{f}_\epsilon}(x_0) &= \frac{1}{\|x_0\|_\nu^2} \left(\int_{B_E} \left\langle \frac{1}{\mu(B_\epsilon(x_0))} \int_{B_\epsilon(x_0)} f(x) d\mu(x), x^* \right\rangle \langle x_0, x^* \rangle d\nu(x^*) \right) \\ &= \frac{1}{\|x_0\|_\nu^2 \mu(B_\epsilon(x_0))} \int_{B_E} \int_{B_\epsilon(x_0)} \langle f(x), x^* \rangle \langle x_0, x^* \rangle d\mu(x) d\nu(x^*). \end{aligned}$$

Therefore, using Fubini’s Theorem, we get

$$\begin{aligned} |\lambda_f(x_0) - \lambda_{\widehat{f}_\epsilon}(x_0)| &= \left| \frac{1}{\|x_0\|_\nu^2 \mu(B_\epsilon(x_0))} \left(\int_{B_\epsilon(x_0)} \int_{B_E} \langle f(x_0), x^* \rangle \langle x_0, x^* \rangle d\nu(x^*) d\mu(x) \right. \right. \\ &\quad \left. \left. - \int_{B_E} \int_{B_\epsilon(x_0)} \langle f(x), x^* \rangle \langle x_0, x^* \rangle d\mu(x) d\nu(x^*) \right) \right| = \frac{1}{\|x_0\|_\nu^2 \mu(B_\epsilon(x_0))} \\ &\quad \times \left| \int_{B_\epsilon(x_0)} \int_{B_E} \langle f(x_0) - f(x), x^* \rangle \langle x_0, x^* \rangle d\nu(x^*) d\mu(x) \right| \\ &\leq \frac{1}{\|x_0\|_\nu^2} \sup_{x \in B_\epsilon(x_0)} \left(\int_{B_E} |\langle f(x_0) - f(x), x^* \rangle|^2 d\nu(x^*) \right)^{1/2} \left(\int_{B_E} |\langle x_0, x^* \rangle|^2 d\nu(x^*) \right)^{1/2} \\ &\leq \frac{1}{\|x_0\|_\nu^2} \sup_{x \in B_\epsilon(x_0)} \|f(x) - f(x_0)\|_\nu \cdot \|x_0\|_\nu \leq \frac{1}{\|x_0\|_\nu} \cdot Lip(f) \cdot \epsilon. \end{aligned}$$

(3) Using Hölder inequality, we have that

$$\begin{aligned} \|\widehat{f}_\epsilon(x_0) - \lambda_{x_0}^\epsilon x_0\|_v &= \int_{B_\epsilon(x_0)} \left| \left\langle \frac{1}{\mu(B_\epsilon(x_0))} \int_{B_\epsilon(x_0)} f(x) d\mu(x), x^* \right\rangle - \lambda_{x_0}^\epsilon \langle x_0, x^* \rangle \right|^2 dv(x^*) \\ &= \int_{B_E} \left| \frac{1}{\mu(B_\epsilon(x_0))} \int_{B_\epsilon(x_0)} \langle f(x), x^* \rangle d\mu(x) - \frac{1}{\mu(B_\epsilon(x_0))} \int_{B_\epsilon(x_0)} \langle \lambda_{x_0}(x)x_0, x^* \rangle d\mu \right|^2 dv(x^*) \\ &\leq \int_{B_E} \frac{1}{\mu(B_\epsilon(x_0))} \int_{B_\epsilon(x_0)} |\langle f(x) - \lambda_{x_0}(x)x_0, x^* \rangle|^2 d\mu(x) dv(x^*). \end{aligned}$$

Thus, applying the equation for the best approximation provided by Lemma 1 and doing some standard calculations, we obtain that the minimal expression for this bound is given by

$$\|\widehat{f}_\epsilon(x_0) - \lambda_{x_0}^\epsilon \cdot x_0\|_v^2 \leq \frac{1}{\mu(B_\epsilon(x_0))} \left(\int_{B_\epsilon(x_0)} \|f(x)\|_v^2 d\mu - \int_{B_\epsilon(x_0)} \lambda_{x_0}^2(x) d\mu(x) \cdot \|x_0\|_v^2 \right).$$

□

Remark 2. The eigenmeasure approximation gives a direct formula to compute diagonal-based representation of operators. Note that, the smaller ϵ is, the better the approximation, but ϵ cannot in general converge to 0. Indeed, in this case it may happen that $ev(f) \cap B_\epsilon(x_0)$ is empty, and so by the definition of eigenmeasure, $\mu(ev(f) \cap B_\epsilon(x_0)) = 0$, which implies that the integral formula that gives $\widehat{f}_\epsilon(x_0)$ equals 0.

Notice that this provides a way to approximate a Lipschitz function f once the set of eigenvectors of f is known, at least partially. Since we are thinking about computational applications supported by these ideas, we also need a numerical way of finding an approximation to the set $ev(f)$. One of the most straightforward methods would be the minimization of the functions $|f(x) - \lambda(x)x|$, which can be performed by sampling methods and Monte Carlo approximations of the integrals involved. Although explanations of these procedures and the algorithms involved are not part of the content of this paper, some hints are given in the last section.

3. Diagonalizable Non-Linear Operators: Examples and Situations

Let E be a Banach space. Suppose that $T : E \rightarrow E$ is a (non-necessarily linear) map such that the set of its eigenvectors $ev(T)$ is not empty. For every non-empty subset $S \subset ev(T)$, consider the restriction $T_S : S \rightarrow E$. We will say that a procedure for extending the value of the function T from $ev(T)$ to the whole space E is an *extension rule* (ER) if for every non-empty subset $S \subset ev(T)$, (ER) produces a unique extension $\widehat{T}_S^{ER} : E \rightarrow E$ of T_S (that is, $\widehat{T}_S^{ER}(x)$ is well-defined for all $x \in E$, and $\widehat{T}_S^{ER}|_S = T_S$).

Thus, we can establish that a map T is *ER-diagonalizable* if there is an increasing sequence of subsets $(S_k)_{k=1}^{+\infty}$ of $ev(T)$ such that for every $x \in E$,

$$\lim_n \widehat{T}_{S_k}^{ER}(x) = T(x).$$

Application of the usual linear extension rule for linear operators gives that such equality can be achieved exactly. However, in general we cannot expect to find a complete representation and only approximate expressions can be obtained. We will look for such approximations in the subsequent sections, after showing some basic examples in the present one.

3.1. Linear and Quasi-Linear Rules

Let us explain a general setting for using the usual framework of spectral representation of linear operators. In the finite dimensional case, diagonalization of a linear map is essentially given by linearity, once a basis of eigenvectors is obtained. If the operator allows a diagonal representation, the extension rule is then provided by the linear combination of

eigenvectors, which define a basis of the space. Let us see how this can be extended to the nonlinear case.

Let $T : E \rightarrow E$, where E is a Banach space with an unconditional (or finite) normalized basis $\mathcal{B} := \{b_1, \dots, b_k, \dots : k = 1, \dots, n\}$, where $n \in \mathbb{N} \cup \{+\infty\}$. Assume that $\{\tau b_k : \tau \in \mathbb{R}\} \subset \text{ev}(T)$ for every k . Consider the projections $P_{b_k}(v) = \alpha_k(v) b_k$, $v \in E$, and take a probability measure as $\mu = \sum_{i=1}^{+\infty} \beta_k \delta_{b_k}$, for $\beta_k \geq 0$. Then, T is diagonalizable under a linear rule conditioned to the basis \mathcal{B} , if we can write by means of the integral formula

$$T(v) = \int_{S_E} \frac{1}{\mu(\{x\})} \cdot T(P_x(v)) \, d\mu(x).$$

Note that $\mu(\{b_k\}) = \beta_k$, and this term does not appear in the integral if $\beta_k = 0$. This expression can be rewritten, for the finite dimensional case, as

$$T(v) = \sum_{k=1}^n T(P_{b_k}(v)) = \sum_{k=1}^n \lambda(P_{b_k}(v)) \cdot P_{b_k}(v) = \sum_{k=1}^n \lambda(\alpha_k(v) b_k) \cdot \alpha_k(v) \cdot b_k.$$

Example 1. Let us see a particular example. Take $E = \mathbb{R}^2$ and the canonical basis $\mathcal{C} = \{b_1, b_2\}$, $b_1 = (1, 0)$, $b_2 = (0, 1)$. Take the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ given by

$$f(x_1, x_2) = (|x_1|x_1, |x_2|x_2), \quad (x_1, x_2) \in \mathbb{R}^2.$$

If we write the eigenvalue equations

$$|x_1|x_1 = \lambda x_1, \quad |x_2|x_2 = \lambda x_2,$$

we directly get that $(x_1, 0)$ and $(0, x_2)$ are eigenvectors for all $x_1, x_2 \in \mathbb{R}$, and the corresponding eigenvalues are $|x_1|$ and $|x_2|$, respectively. That is, all the vectors in the subspaces $\{\tau b_1 : \tau \in \mathbb{R}\}$ and $\{\tau b_2 : \tau \in \mathbb{R}\}$ are eigenvectors of T . Moreover, note that for

$$|x_1| = |x_2| = \lambda,$$

we obtain that the vectors as (x, x) and $(x, -x)$ for $x \in \mathbb{R}$ are also eigenvectors with eigenvalues $|x|$. Figure 1 provides a representation of the eigenvectors of T .

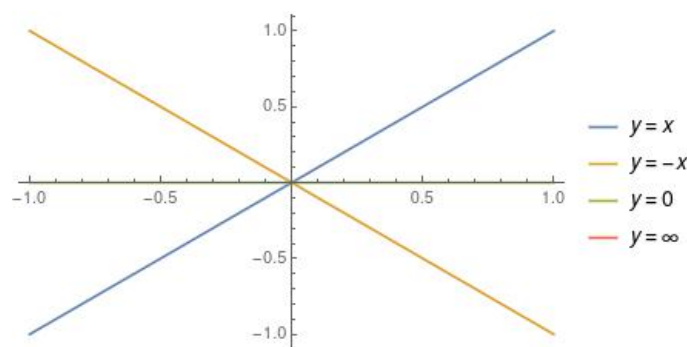


Figure 1. Eigenvectors in Example 1.

Consider the orthonormal projections $P_{b_k}(v) = \langle v, b_k \rangle b_k$, $v \in E$, $k = 1, 2$. Then, T is diagonalizable under the linear rule since

$$T(v) = \sum_{k=1}^2 T(P_{b_k}(v)) = \sum_{k=1}^2 \lambda(P_{b_k}(v)) \cdot P_{b_k}(v) = |x_1| \langle x_1, b_1 \rangle b_1 + |x_2| \langle x_2, b_2 \rangle b_2.$$

According to the given definition, this map is diagonalizable under a linear rule, the one given by this equation. Let us note that there are other eigenvectors that are not used in the representation, with which we could also define a basis of the space.

Example 2. Linearization techniques for nonlinear operators are interesting tools used in many applied contexts. A relevant example is given by the so-called Koopman nodes, which allow one to treat nonlinear systems by means of linear procedures. This method is used to find spectral representations of observables in dynamical systems (see, for example, [28]). In [29], Koopman nodes are used to describe the behavior of complex nonlinear flows, in the context of fluid mechanics.

Let us briefly explain this procedure. Consider a dynamical system given by a function $f : E \rightarrow E, f(x_k) = x_{k+1}$. The Koopman operator is a linear operator U that acts on the space of all scalar functions $\mathcal{F} : X \rightarrow \mathbb{R}$ as $U(\varphi)(x) = \varphi(f(x))$. Since U is linear, we can consider its diagonalization in order to obtain a set of eigenfunctions φ_j associated with eigenvalues $\lambda_j \in \mathbb{C}$ belonging to a certain function space. Consider an observable g , that is, a function $g : E \rightarrow \mathbb{R}^p$ for a given natural number p . If a countable number of eigenfunctions is sufficient to act as a basis of the subspace to which the p coordinate functions of g belong, we can expand g as

$$g(x) = \sum_{j=1}^{\infty} \varphi_j(x) v_j, \quad x \in E,$$

where the p coordinates of the vectors v_j are the corresponding coefficients of the representations on the system $\{\varphi_j\}_{j=1}^{\infty}$ of each coordinate function of g . The functions φ_j are called Koopman eigenfunctions, and the vectors v_j the Koopman modes of the map f .

Our setting is similar in various aspects to the ones of the Koopman method, but we use measures and integral representations to facilitate the further use of integral averages to get approximation formulas. Let us see what happens in the case where the function f is a symmetric linear map $f : E \rightarrow E$ on the Euclidean space E . Take an orthonormal basis of eigenvectors v_1, \dots, v_n for f with eigenvalues $\lambda_1, \dots, \lambda_n$, and the dual basis w_1, \dots, w_n of functions of the dual space E^* . Note that in this case the Koopman operator $U : E^* \rightarrow E^*$ is just the adjoint f^* of f since

$$U(x^*)(x) = \langle f(x), x^* \rangle = \langle x, f^*(x^*) \rangle, \quad x \in E, \quad x^* \in E^*.$$

Take now the identity map $I : E \rightarrow E$ as g : the expansion formula is

$$x = \sum_{j=1}^n \langle x, w_j \rangle v_j,$$

that is, the Koopman nodes are in this case the eigenvectors of f , and we get a representation of f as

$$f(x) = \sum_{j=1}^n \lambda_j \langle x, w_j \rangle v_j.$$

Using our finite measure integration formalism, we can consider the measure $\nu = \frac{1}{n} \sum_{j=1}^n \delta_{w_j}$ and the eigenmeasure $\mu = \frac{1}{n} \sum_{j=1}^n \delta_{v_j}$ to get an integral representation of f as the Bochner integral

$$f(x_0) = n^2 \int_{B_E} \lambda(x) \left(\int_{B_{E^*}} \langle x_0, x^* \rangle \langle x, x^* \rangle d\nu(x^*) \right) x d\mu(x), \quad x_0 \in E,$$

where $\lambda(x) = \lambda_j$ for $x = v_j$ and 0 otherwise. This gives a result equivalent to that obtained by the Koopman operator method. This procedure, which can also be applied in the nonlinear case by following the same steps, shows another example of a linear rule; other examples of application of Koopman nodes to nonlinear problems (and which could be rewritten in our terms) can be found in [29] (see Section 2.2 for application to periodic solutions of dynamical systems).

Example 3 (Quasi-linear rules). More sophisticated adaptations of the linear rule could be necessary to cover more complex applications to non-linear operators of interest in physics and other sciences. For example, the moment of inertia is a fundamental magnitude that is used in mechanical systems, and sometimes its time dependence has to be modelled, for example, in the description of the vibratory motion [30]. Let us show a related operator here. Let us consider a 2-dimensional system

and suppose that the time dependence of an moment of inertia is given by the operator $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ (the third coordinate for the variable time t), given by the formula

$$T(v) = \begin{pmatrix} 2t & 1 & 0 \\ 1 & 2t & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ t \end{pmatrix}, \quad v = (x, y, t),$$

where x and y are the usual Cartesian coordinates of the points in \mathbb{R}^2 , and the third coordinate represents a weighted counter of the time t . After a diagonalization process of T , the analyst can get (time-dependent) information about the principal axis of inertia of the physical systems that are represented by this operator. As in the linear case, a direct computation involving the characteristic polynomial of the matrix above provides eigenvalues of T . Indeed, the (time-dependent) solutions of the equation

$$\begin{vmatrix} 2t - \lambda & 1 & 0 \\ 1 & 2t - \lambda & 0 \\ 0 & 0 & 3 - \lambda \end{vmatrix} = 0, \quad \text{for } \lambda \in \mathbb{R}^2,$$

are $\lambda_1(t) = 2t - 1, \lambda_2(t) = 2t + 1$, and $\lambda_3(t) = 3$. Note that the system is not linear, so no information about the general structure of the sets of eigenvectors is available. However, the vectors $(1, -1, 0), (1, 1, 0)$, and $(0, 0, 1)$ provide a basis for \mathbb{R}^3 of eigenvectors of T associated with λ_1, λ_2 , and λ_3 , respectively. If we write x', y' , and t' for the coordinates of the vectors in \mathbb{R}^3 with respect to this new basis, we get the change of variables are given by the equations $x' = (x - y)/2, y' = (x + y)/2$ and $t = t'$. A straightforward calculation shows that it provides an “integral representation” of T as

$$\begin{aligned} T(v) &= (2t - 1) \begin{pmatrix} (x - y)/2 \\ (y - x)/2 \\ 0 \end{pmatrix} + (2t + 1) \begin{pmatrix} (x + y)/2 \\ (x + y)/2 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 0 \\ t \end{pmatrix} \\ &= \lambda_1(t) x' e_1 + \lambda_2(t) y' e_2 + \lambda_3(t) t' e_3. \end{aligned}$$

Thus, a quasi-linear “integral representation” as

$$T(v) = \sum_{e_i \in S \subset ev(T)} \lambda_i(v) P_{e_i}(v) e_i$$

has been obtained, where $\lambda_i(v)$ is the eigenvalue associated with e_i depending on v , and where $P_{e_i}(v)$ gives the coordinate of v in the new basis.

3.2. Topological Rules

Although the approximation of a linear operator is essentially a topological question, the usual methods and techniques are strongly influenced by linearity. Point approximation can also be performed using sequences of locally constant functions, and, in fact, these functions are primarily used for approximation in the context of vector-valued integrable functions. These functions can usually be written as pointwise limits of simple functions, provided they belong to any Köthe–Bochner function space under some weak requirements of density or order continuity.

Let us recall the definition of (strong) eigenmeasure μ for an operator, applied to the context of functions $f : E \rightarrow E$, that will be useful in the present section. Although the definition provided in [26] is rather abstract, for the aim of this paper we can use the following version: let $\mathcal{B}(E_0)$ be the sigma-algebra of Borel sets of a multiple E_0 of the unit ball of the Banach space E associated with the norm topology. A measure $\mu : \mathcal{B}(E_0) \rightarrow \mathbb{R}^+$ is an eigenmeasure for a Bochner μ -integrable map $f : E_0 \rightarrow E$ (that is, the restriction of $f : E \rightarrow E$ to E_0) if it is a regular Borel probability measure such that

- $ev(f) \in \mathcal{B}(E_0)$, and
- for every $A \in \mathcal{B}(E_0), \mu(A) = 0$ if $\mu(A \cap ev(f)) = 0$.

Note that, in this case, the Bochner integral of $f(x)$ can be used to reconstruct the diagonal part of the function f by the integral equations

$$\int_A f(x) d\mu(x) = \int_A \lambda(x) x d\mu(x), \quad A \in \mathcal{B}(E_0).$$

In the case that the set $ev(f)$ is finite, μ can always be defined as a proper convex combination of Dirac’s deltas $\delta_{x_i}, x_i \in ev(f)$. Let us use these integral equations for the next construction.

Consider a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n, n \in \mathbb{N}$, which has a finite set of eigenvectors $ev(f) = \{x_1, \dots, x_m\}$. Define a probability measure $\mu : \mathcal{B}(E_0) \rightarrow \mathbb{R}^+$ by

$$\mu(A) = \sum_{i=1}^m \alpha_i \delta_{x_i}(A), \quad \sum_{i=1}^m \alpha_i = 1, \quad \alpha_i > 0 \text{ for every } i, \text{ and } A \in \mathcal{B}(E_0),$$

where $E_0 = \max_{i=1, \dots, m} \|x_i\| \cdot B_E$.

Let us define the *metric rule* for the extension of such a function $f_0 : ev(f) \rightarrow \mathbb{R}^+$ as follows: for every point $x \in E$, we compute the (measurable) set

$$M_x := \left\{ x_i \in ev(f) : \min_{k=1, \dots, m} \|x_k - x\| = \|x_i - x\| \right\}.$$

Then, we define the extension \hat{f} by

$$\hat{f}(x) := \frac{1}{\mu(M_x)} \int_{M_x} \lambda(y) y d\mu(y).$$

Therefore, we say that a function $f : E \rightarrow E$ with a finite number of eigenvectors is *diagonalizable under a metric rule* if there is an eigenmeasure μ such that

$$f(x) = \frac{1}{\mu(M_x)} \int_{M_x} \lambda(y) y d\mu(y) \quad \text{for all } x \in E.$$

Example 4. Consider the function $\phi : (\mathbb{R}^2, \|\cdot\|_2) \rightarrow (\mathbb{R}^2, \|\cdot\|_2)$ given by

$$\phi((x_1, x_2)) = \begin{cases} (0, 1) & \text{if } x_2 > \max\{x_1, -x_1\} \\ (0, -1) & \text{if } x_2 < \min\{x_1, -x_1\} \\ (-1, 0) & \text{if } x_1 < \min\{x_2, -x_2\} \\ (1, 0) & \text{if } x_1 > \max\{x_2, -x_2\} \end{cases}$$

and having in the diagonal lines the averages of the values of the adjacent sectors. Thus, it is a constant function inside any of the regions defined by the diagonal lines in Figure 2, for which the minimal distance to any of the reference points $(1, 0), (-1, 0), (0, 1),$ and $(0, -1)$ provides the value of the constant image vector.

The description of the function ϕ is easy: if we divide \mathbb{R}^2 into four sectors by the diagonal lines that intersect at $(0, 0)$ that can be seen in Figure 2, we have that the function is constant in each of them, and it is also constant in the diagonal lines themselves.

It can also be equivalently defined, as explained in the explanation of the metric rule above. Consider the points $e_1 = (1, 0), e_2 = (0, 1), e_3 = (-1, 0)$ and $e_4 = (0, -1)$. It is clear that they are eigenvectors for ϕ with eigenvalues equal to 1. Consider the eigenmeasure $\mu(\cdot) = \frac{1}{4} \sum_{k=1}^4 \delta_{e_k}(\cdot)$. For each point x that is not in the diagonal lines, there is a point e_k of these ones such that the set M_x is given by $M_x = \{e_k\}$ since the distance to these points actually defines in which sector the point x is (see Figure 2). Therefore, for every element x inside of a sector, we apply the linear rule

$$\phi(x) = \frac{1}{\mu(M_x)} \int_{M_x} \phi(x) d\mu = \frac{1}{1/4} \phi(e_k) \cdot \frac{1}{4} = e_k.$$

Case x belongs to the diagonal lines, but it is not $(0, 0)$; we have that there are two eigenvectors e_k and e_j that are at the same distance to x , getting in this case $M_x = \{e_k, e_j\}$, and so

$$\phi(x) = \frac{1}{\mu(M_x)} \int_{M_x} \phi(x) d\mu = \frac{1}{1/2} \left(\phi(e_k) \cdot \frac{1}{4} + \phi(e_j) \cdot \frac{1}{4} \right) = \frac{1}{2}(e_k + e_j).$$

Finally, $M_{(0,0)} = \{e_1, e_2, e_3, e_4\}$ and then

$$\phi((0,0)) = \frac{1}{\mu(M_{(0,0)})} \int_{M_{(0,0)}} \phi(x) d\mu = \left(\sum_{k=1}^4 \phi(e_k) \cdot \frac{1}{4} \right) = (0,0).$$

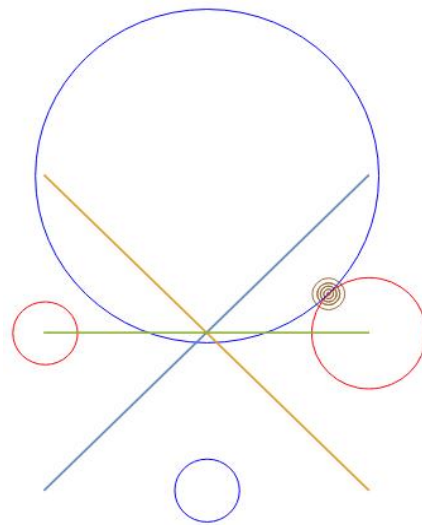


Figure 2. Representation of the distances to the nearest points of the point $(3/4, 1/4)$: according to the definition, in this case the value of $\phi((3/4, 1/4))$ is $(1, 0)$.

Example 5. This easy procedure can be adapted to other situations to approximate non-linear operators using the minimal distance criterion. The idea is to use the suitable inequalities appearing in the description of the operator to control the error committed. For instance, in [31] the authors analyze a nonlinear multiparameter eigenvalue problem arising in the theory of nonlinear multifrequency electromagnetic wave propagation. Under certain conditions, the problem can be written as a system of n nonlinear single parameter eigenvalue problems, with linear and non-linear components. It is proved that the corresponding linear problems have a finite number of positive eigenvalues; however, the nonlinear ones have infinite positive eigenvalues, which provide eigentuples of the original multiparameter problem not related to the solutions of the linear one-parameter problems. After some transformation, an eigenvector equation appearing in the formalism is defined as

$$f'' + \alpha f^3 + \epsilon f = \lambda f$$

with f belonging to the space of functions with continuous second derivatives (see Equation (6) in [31]; see also Equation (15) in [32]). We can then search for the eigenvectors of the operator $T(f) = f'' + \alpha f^3 + \epsilon f$. Once we get a set $ev(T)_0$ of these eigenfunctions, we can find a naive approximation to $T(f)$ the value $T(u)$ for the function $u \in ev(T)_0$ that minimizes a certain “error function”, using the inequality

$$\|T(f) - T(u)\|_2 \leq \|f'' - u''\|_2 + |\alpha| \|f^3 - u^3\|_2 + \epsilon \|f - u\|_2,$$

where we consider for example the norm $\|\cdot\|_2$ of function space L^2 . Suppose that we can assume that map $g \mapsto g''$ is a Lipschitz map when considered from $ev(T)_0 \cup \{f\}$ to L^2 and that all these

functions are pointwise uniformly bounded in their domains by a constant Q . Then, we can get a Lipschitz type estimate for $\|T(f) - T(u)\|_2$ as

$$\|T(f) - T(u)\|_2 \leq K\|f - u\|_2 + |\alpha|\|(f - u)(f^2 + fu + u^2)\|_2 + \epsilon\|f - u\|_2 \leq R\|f - u\|_2$$

for $R = (K + |\alpha|3Q^2 + \epsilon)$. The best approximation by the metric rule would be given in this case by $f \approx \lambda_u u$, where u is the function in $ev(T)_0$ for which the minimal L^2 distance to f is attained. A more precise calculation would lead to a domination by the norm of the Sobolev space $W^{2,2}$.

There are a lot of nonlinear problems directly connected to differential operators arising in physics, biology, economics, and other sciences. Just to mention some recent developments, the papers [33,34] study the principal eigenvalue of p -biharmonic and (p, q) -biharmonic operators, which is applied to various differential equation problems appearing in physics. More specific results could be obtained by setting some requirements on the functions involved beforehand. Deeper assumptions of the functional setting would provide more possibilities to find more concrete applications, both in terms of how to obtain the eigenvectors of the operators and the general metric structure of the function spaces under consideration. For example, fixed-point results are often used in the search for eigenvalues of nonlinear functions, and more general notions of distance and function spaces can be used, including order-based tools [35–38].

3.3. The Lipschitz Case

Let us show now how a diagonalization procedure can be implemented in the context of Lipschitz maps. Obviously, the essential requirement in this case is that the operators to be diagonalized have to be Lipschitz; let us note the example of application of a linear rule proposed in Section 3.1 (Example 1 is not Lipschitz, so the method presented here is not an extension of a linear-type rule). In fact, metric topology is used instead as it corresponds to the use of Lipschitz operators, but an additional tool is available in this case, so this situation is not an extension of the metric rule shown in Section 3.2: note that Example 4 is not given by a Lipschitz map.

Let us show a non-trivial case of diagonalizable Lipschitz map.

Example 6. Let E be a Banach space, and consider the function $T : E \rightarrow E$ given by

$$T(x) := \min\{1, \|x\|\}x, \quad x \in E.$$

Let us show that T is a Lipschitz map. First, suppose that $\|y\| \leq 1$. Then, we have that

$$\begin{aligned} \|T(x) - T(y)\| &= \|\min\{1, \|x\|\}x - \|y\|y\| \\ &= \|\min\{1, \|x\|\}x - \min\{1, \|x\|\}y + \min\{1, \|x\|\}y - \|y\|y\| \\ &\leq \min\{1, \|x\|\} \|x - y\| + \left| \min\{1, \|x\|\} - \|y\| \right| \|y\| \\ &\leq \|x - y\| + \left| \|x\| - \|y\| \right| \|y\| \leq 2 \|x - y\|. \end{aligned}$$

Of course, the same computations give the same result in case $\|x\| \leq 1$. Finally, if $\|x\| > 1$ and $\|y\| > 1$, we obtain

$$\|T(x) - T(y)\| = \|x - y\| \leq 2\|x - y\|,$$

which gives that T is Lipschitz with constant $K \leq 2$.

The main property of the function T is that it is “absolutely” diagonalizable: indeed, all the vectors in the space are eigenvectors, and for each of them the eigenvalue is given by $\min\{1, \|x\|\}$. Similar examples could be obtained for special operators defined as $T(x) = f(x) \cdot x$, with f being a real bounded Lipschitz function.

4. Approximation of Lipschitz Maps by the Proximity Rule

In case we have no more information than that provided by the fact that T is Lipschitz, we have to apply direct algorithms for obtaining an approximation to T . Let us show two methods to do it.

4.1. The Minimal Distance Extension of $T|_{ev(T)}$

Take a strictly decreasing sequence of positive numbers $(\epsilon_i)_{i=1}^{+\infty}$ such that $\lim_i \epsilon_i = 0$, and consider an eigenmeasure μ for T . We can get an approximation to T as follows. We have to assume that norm-bounded subsets of $ev(T)$ are relatively norm-compact, which is obviously true if E is finite dimensional.

- Fix an element $x \in E$ and define $d_0 = \inf\{\|x - y\| : y \in ev(T)\}$. Consider the set $M_{\epsilon_1}(x) = \{y \in S : \|x - y\| \leq d_0 + \epsilon_1\}$.
- Choose an element $y_1 \in M_{\epsilon_1}(x)$.
- We consider the value of the first extension of T as $\lambda_{y_1} \cdot y_1$, where λ_{y_1} is the eigenvalue associated with y_1 .
- Repeat the procedure to get a sequence $(y_i)_{i=1}^n$ in $ev(T)$ with associated eigenvalues $(\lambda_{y_i})_{i=1}^{+\infty}$. By the assumption on compactness, and taking into account that the sequence of vectors is bounded, there is a subsequence that converges to an element y . Let us write again $(y_i)_{i=1}^{+\infty}$ for the subsequence. Note that for each $i \in \mathbb{N}$, we have

$$\|\lambda_{y_i} \cdot y_i - T(y)\| = \|T(y_i) - T(y)\| \leq Lip(T) \cdot \|y_i - y\|,$$

and $\lim_i \|y_i - y\| = 0$.

This allows one to define the extension \hat{T} as $\hat{T}(x) = T(y) = \lim_i \lambda_{y_i} \cdot y_i$.

- Note that, due to the Lipschitz inequality for T , we can write the following bound for the error committed.

$$\begin{aligned} \|T(x) - \hat{T}(x)\| &= \|T(x) - T(y)\| = \|T(x) - \lim_i T(y_i)\| \leq Lip(T) \cdot \lim_i \|x - y_i\| \\ &= Lip(T) \cdot \inf\{\|x - y\| : y \in ev(T)\} = Lip(T) \cdot d_0. \end{aligned}$$

This gives a basic extension procedure, for which the only requirements are the existence of a convenient set of eigenvectors and the Lipschitz inequality for T . Indeed, in case we have no information about the map T other than it is Lipschitz; we have to choose the point y in $ev(T)$ that makes the bound $\inf_{x_i \in ev(T)} \|x_i - x\|$ to be (at least almost) attained. That is, if a point $x_{i_0} \in ev(T)$ attains this infimum, we approximate $T(x)$ by $T(x_{i_0})$.

However, this is not the best way of obtaining a good approximation for $T(x)$ if we can assume certain convexity properties for the map T . Essentially, we have to assume that the map T is nearly convex, that is, we can approximate it by a convex combination of eigenvectors as

$$T\left(\sum_{x_i \in ev(T)} \alpha_i x_i\right), \quad \text{by} \quad \sum_{x_i \in ev(T)} \alpha_i T(x_i).$$

In case if we assume such property, we can find better error bounds, which suggests other approximation methods instead of the one provided by the nearest point in $ev(T)$. Recall that for this point to exist and to be unique for the next definition makes sense, we need some strict convexity conditions on X ; we are assuming that the space X is uniformly convex. This is explained in Section 5. Before doing it, let us present a most elaborated version of the minimal distance extension just explained.

4.2. The Average Value Extension of $T|_{ev(T)}$

As in the previous case, take a strictly decreasing sequence of positive numbers $(\epsilon_i)_{i=1}^{+\infty}$ such that $\lim_i \epsilon_i = +\infty$ and consider an eigenmeasure μ for T . We can get an approximation to T as follows.

Fix an element $x \in E$, define $d_0 = \inf\{\|x - y\| : y \in \text{ev}(T)\}$, and consider the sets $M_{\epsilon_i}(x) = \{y \in S : \|x - y\| \leq d_0 + \epsilon_i\}$, $i \in \mathbb{N}$.

Let us assume first that

$$\lim_i \mu(M_{\epsilon_i}(x)) > 0;$$

we will see after about what to do in the case $\lim_i \mu(M_{\epsilon_i}(x)) = 0$.

Write $M_0(x)$ for the measurable set $\bigcap_{i=1}^{+\infty} M_{\epsilon_i}(x)$. Note that, by the monotone convergence theorem,

$$\mu(M_0(x)) = \mu\left(\bigcap_{i=1}^{+\infty} M_{\epsilon_i}(x)\right) = \lim_i \mu(M_{\epsilon_i}(x)) > 0$$

and so $M_0(x)$ is non-null.

- Write $\widehat{T}^{\epsilon_1}(x) = \frac{1}{\mu(M_{\epsilon_1}(x))} \int_{M_{\epsilon_1}(x)} \lambda_y \cdot y \cdot \chi_{\{y\}} d\mu(y)$, and define $\widehat{T}^{\epsilon_i}(x)$ in the same way, as

$$\widehat{T}^{\epsilon_i}(x) = \frac{1}{\mu(M_{\epsilon_i}(x))} \int_{M_{\epsilon_i}(x)} \lambda_y \cdot y \cdot \chi_{\{y\}} d\mu(y),$$

for all $i \in \mathbb{N}$.

- The set $\{\lambda_y \cdot y : y \in M_{\epsilon_1}(x)\}$ is norm bounded: indeed, we have that for every $y \in M_{\epsilon_1}(x)$,

$$\begin{aligned} \|\lambda_y \cdot y\| &\leq \|T(x)\| + \|T(x) - T(y)\| \\ &\leq \|T(x)\| + \|T(x) - \lambda_y \cdot y\| \leq \|T(x)\| + d_0 + \epsilon_1. \end{aligned}$$

- Note that, if $j < i$,

$$\begin{aligned} &\|\widehat{T}^{\epsilon_i}(x) - \widehat{T}^{\epsilon_j}(x)\| \\ &\leq \left\| \left(\frac{\mu(M_{\epsilon_j}(x)) - \mu(M_{\epsilon_i}(x))}{\mu(M_{\epsilon_i}(x)) \cdot \mu(M_{\epsilon_j}(x))} \right) \int_{M_{\epsilon_i}(x)} \lambda_y \cdot y \cdot \chi_{\{y\}} d\mu(y) \right\| \\ &\quad + \left\| \frac{1}{\mu(M_{\epsilon_j}(x))} \left(\int_{M_{\epsilon_j}(x)} \lambda_y \cdot y \cdot \chi_{\{y\}} d\mu(y) - \int_{M_{\epsilon_j}(x)} \lambda_y \cdot y \cdot \chi_{\{y\}} d\mu(y) \right) \right\| \\ &\leq \left\| \left(\frac{\mu(M_{\epsilon_j}(x) \setminus M_{\epsilon_i}(x))}{\mu(M_{\epsilon_i}(x)) \cdot \mu(M_{\epsilon_j}(x))} \right) \int_{M_{\epsilon_i}(x)} \lambda_y \cdot y \cdot \chi_{\{y\}} d\mu(y) \right\| \\ &\quad + \left\| \frac{1}{\mu(M_{\epsilon_j}(x))} \left(\int_{M_{\epsilon_j}(x) \setminus M_{\epsilon_i}(x)} \lambda_y \cdot y \cdot \chi_{\{y\}} d\mu(y) \right) \right\| \\ &\leq \left| \frac{\mu(M_{\epsilon_j}(x) \setminus M_{\epsilon_i}(x))}{\mu(M_{\epsilon_i}(x)) \cdot \mu(M_{\epsilon_j}(x))} \right| \cdot (\|T(x)\| + d_0 + \epsilon_1) \cdot \mu(M_{\epsilon_i}(x)) \\ &\quad + \frac{1}{\mu(M_{\epsilon_j}(x))} \cdot (\|T(x)\| + d_0 + \epsilon_1) \cdot \mu(M_{\epsilon_j}(x) \setminus M_{\epsilon_i}(x)) \\ &\leq 2 \frac{\mu(M_{\epsilon_j}(x) \setminus M_{\epsilon_i}(x))}{\mu(M_{\epsilon_j}(x))} \cdot (\|T(x)\| + d_0 + \epsilon_1) \\ &\leq 2 \frac{\mu(M_{\epsilon_j}(x)) - \mu(M_{\epsilon_i}(x))}{\mu(M_0(x))} \cdot (\|T(x)\| + d_0 + \epsilon_1) \\ &\leq 2 \frac{|\mu(M_{\epsilon_j}(x)) - \mu(M_0(x))| + |\mu(M_0(x)) - \mu(M_{\epsilon_i}(x))|}{\mu(M_0(x))} \cdot (\|T(x)\| + d_0 + \epsilon_1), \end{aligned}$$

that converges with i, j to 0. Consequently, $\|\widehat{T}^{\epsilon_i}(x) - \widehat{T}^{\epsilon_j}(x)\| \rightarrow_{i,j} 0$, and so we can define the approximation at x by

$$\widehat{T}^0(x) := \lim_i \widehat{T}^{\epsilon_i}(x) = \lim_i \frac{1}{\mu(M_{\epsilon_i}(x))} \int_{M_{\epsilon_i}(x)} \lambda_y \cdot y \cdot \chi_{\{y\}} d\mu(y).$$

Finally, in case $\mu(M_0(x)) = 0$, we fix a small $\delta > 0$ and change d_0 in the definition of the sets $M_{\epsilon_i}(x)$ by $d_0 = \inf\{\|x - y\| : y \in ev(T)\} + \delta$ to get a non- μ -null set $M_0(x)$.

5. Approximation of Lipschitz Maps by the Convexity Rule

In this section, we show how a Lipschitz map T can be approximated using the evaluation of the function just in the set of eigenvectors, under some convexity assumptions on T . Although the explanation we develop here is of theoretical nature, we are thinking in effective algorithms to approximate functions by this method, so minimization of the errors explained in Section 2 of the paper for getting eigenvectors or “almost” eigenvectors to provide a general method for extending the idea of average reconstruction of a Lipschitz map by means of its eigenvectors. We assume here that the involved maps are “approximately convex”, that is, the values of the maps at all the points are near to a convex combination of the eigenvectors that are near them. This is the main difference with respect to the methods explained in the previous section.

In order to obtain reasonable error bounds, we need the functions to satisfy some requirements. The main idea is that we can perform approximations of the value of the function at a given point using its metric properties, which are essentially determined by the Lipschitz condition of the function if nothing else is assumed for the function. So, we have to argue in topological terms, in the line of what explained in Section 3.2. We propose to define the pointwise approximation to the Lipschitz map by convex combination of eigenvector evaluations. Consequently, the formalism of eigenmeasure, as well as the computation of nearest points to a given set in a Banach space endowed with a uniformly convex norm, and the inequalities associated with the Lipschitz character of the function, will be required.

Given a Lipschitz function $T : X \rightarrow X$, we will reconstruct the function T by convex combinations of eigenvectors. We have to assume that $ev(T)$ is non-empty. In order to do it, we need to find what is the convex hull of finite sets of eigenvectors that is as near as possible to the point x . In case a given point $x \in X$ belongs to any convex combination of elements x_1, \dots, x_n of $ev(T)$ with eigenvalues $\lambda_1, \dots, \lambda_n$, we can approximate $T(x)$ as the convex combination

$$\sum_{i=1}^n \alpha_i T(x_i) = \alpha_1 \lambda_1 x_1 + \dots + \alpha_n \lambda_n x_n.$$

Since several convex combinations involving different elements of $ev(T)$ can occur, we have to choose one of them: the Lipschitz inequality for T gives the key of how to do it. The idea is to divide the error in two terms, with the first one corresponding to the distance of x to the convex hull of x_1, \dots, x_n and the other one measuring the error committed by the approximation of T as a convex function.

If we write $co(S)$ for the convex hull of any subset S of the Banach space X and $\overline{co(S)}$ for its norm closure, we should write $d(x, S) = \inf_{z \in S} \|x - z\|$ for the distance from S to x . Recall the following fundamental result on optimal approximations in Banach spaces: if X is a uniformly convex Banach space, and $x \in X$ and $C \subset X$ are closed and convex, then there is a unique $y \in C$ such that $\|x - y\| = \inf_{z \in C} \|x - z\| = d(x, C)$. We say as usual that y is the nearest point to x in C . A direct consequence of this general fact is the following.

Proposition 4. *Let X be a uniformly convex Banach space, and let $T : E \rightarrow E$ be a Lipschitz map of norm $Lip(T)$. Suppose that $S = \{x_1, \dots, x_n\} \in ev(T)$. Then, there exists a convex combination $z = \sum_{i=1}^n \alpha_i^0 x_i \in co(S)$, $\alpha_i^0 \geq 0$, such that*

- (i) The minimal distance $\inf \left\{ \left\| \sum_{i=1}^n \alpha_i \lambda_i x_i - T(x) \right\| : \sum_{i=1}^n \alpha_i = 1, \alpha_i \geq 0 \right\}$ is attained at z , and
- (ii) there is a bound for this minimal distance as

$$\left\| \sum_{i=1}^n \alpha_i^0 \lambda_i x_i - T(x) \right\| \leq Lip(T) \cdot \min_{i=1, \dots, n} \|x_i - x\|.$$

Proof.

- (i) Consider the convex hull of the points $\lambda_1 x_1, \dots, \lambda_n x_n$. By the quoted result, there is an optimal point y in the convex hull in which the minimal distance to $T(x)$ is attained that is also an element of X ; let us write it as the convex addition $\sum_{i=1}^n \alpha_i^0 \lambda_i x_i$.
- (ii) Then, since x_i are eigenvectors, if we consider any other convex combination $\sum_{i=1}^n \alpha_i \lambda_i x_i$ we obtain

$$\begin{aligned} \left\| \sum_{i=1}^n \alpha_i^0 \lambda_i x_i - T(x) \right\| &\leq \left\| \sum_{i=1}^n \alpha_i \lambda_i x_i - T(x) \right\| = \left\| \sum_{i=1}^n \alpha_i \lambda_i x_i - \sum_{i=1}^n \alpha_i T(x) \right\| \\ &= \left\| \sum_{i=1}^n \alpha_i (T(x_i) - T(x)) \right\| \leq Lip(T) \cdot \left(\sum_{i=1}^n \alpha_i \|x_i - x\| \right). \end{aligned}$$

Thus, computing the infimum on the right hand side, we get

$$\left\| \sum_{i=1}^n \alpha_i^0 \lambda_i x_i - T(x) \right\| \leq Lip(T) \cdot \inf_{\sum \alpha_i = 1} \left(\sum_{i=1}^n \alpha_i \|x_i - x\| \right) = Lip(T) \cdot \min_{i=1, \dots, n} \|x_i - x\|,$$

as we wanted to prove.

□

Remark 3. Proposition 4 provides an approximation tool for Lipschitz maps, which works best whenever T behaves more or less like a convex function. Note that the nearest point provides another way to produce an approximation of a Lipschitz map in terms of eigenmeasures, again as an integral average. Let us show this in the simple case of a finite set of eigenvalues; the general procedure is explained below, although we do not use the integral formalism in the subsequent explanation for the sake of clarity. Suppose that we have a set $S = \{x_1, \dots, x_n\}$ of eigenvectors of a Lipschitz map T . Write $N_S(x)$ for the nearest point of $\text{co}(S)$ to x , and $P_{\{y\}}(N_S(x))$ (where $y \in S$), for the coefficient α_i^0 associated with $y = x_i$ of the convex combination that gives $N_S(x)$. Note that, by the requirements on E , these coefficients are unique so the expression is well-defined. Take an eigenmeasure μ associated with S , $\mu(\cdot) = \sum_{i=1}^n \mu(\{x_i\}) \cdot \delta_{x_i}(\cdot)$. Then, we have that the best approximation \hat{T} is given by the spectral integral

$$\hat{T}(x) := \int_E \frac{1}{\mu(\{y\})} \lambda(y) P_{\{y\}}(N_S(x)) y d\mu(y).$$

Fixing T , and assuming that convexity is a good method for approximating it, we can define two different errors that can be controlled using complementary arguments. We have to point out that getting the exact value of $T(x)$ just knowing the eigenvectors of T is not possible in general: stronger requirements (such as linearity, for example) are needed. However, we provide an approximate value of $T(x)$ and fix small bounds for the errors committed.

Definition 3. Let X be uniformly convex and consider a Lipschitz map $T : X \rightarrow X$. For each set, $S = \{x_1, \dots, x_n\} \subset \text{ev}(T)$ and $x \in X$; if $\sum_{i=1}^n \alpha_i^0 \lambda_i x_i$ is the nearest point of S to x , we define two approximation errors as follows .

(1) The convexity error ε_c , which measures how far the map T is from being convex, is given by

$$\varepsilon_c(T, S, x) := \left\| T\left(\sum_{i=1}^n \alpha_i^0 x_i\right) - \sum_{i=1}^n \alpha_i^0 \lambda_i x_i \right\|.$$

(2) The distance error $\varepsilon_d(T, S, x)$, which gives the distance of $T(x)$ to the value of the map to the nearest point in S , is given by

$$\varepsilon_d(T, S, x) := \left\| T(x) - T\left(\sum_{i=1}^n \alpha_i^0 x_i\right) \right\|.$$

Next result gives the main tool for the approximation of T ; the way the error is decomposed allows one to evaluate how far the image of the nearest point of S to x by T could be a good approximation for $T(x)$. On one hand, we have that the possibility of writing x as a convex combination of eigenvectors makes the distance error equal to 0, and so we only have to consider the convexity error, which measures how far the function T fails to be convex. Indeed, if it is convex, we have that $T(\sum_{i=1}^n \alpha_i^0 x_i) = \sum_{i=1}^n \alpha_i^0 T(x_i)$, and then $\varepsilon_c(T, S, x) = 0$.

On the other hand, in case T is convex, for getting an exact formula for T in terms of its eigenvectors, we only need the set $ev(T)$ of these eigenvectors to be big enough to cover all the space as convex combinations of the elements of $ev(T)$.

Let us write this explicitly in the next result, which can be considered as a basic form of an approximated Spectral Theorem for Lipschitz operators.

Theorem 2. Let X be a uniformly convex Banach space. Let $T : X \rightarrow X$ be a Lipschitz map. Fix a set $S = \{x_1, \dots, x_n\} \subset ev(T)$ and $x \in X$. Then, there is a convex combination $\sum_{i=1}^n \alpha_i^0 x_i \in co(S)$ such that

$$\left\| T(x) - \sum_{i=1}^n \alpha_i^0 \lambda_i x_i \right\| \leq \varepsilon_c(T, S, x) + \varepsilon_d(T, S, x)$$

and both errors can be bounded as

$$\varepsilon_c(T, S, x) \leq Lip(T) \cdot \left(\sum_{i=1}^n \alpha_i^0 \|x_i - \sum_{j=1}^n \alpha_j^0 x_j\| \right) \quad \text{and} \quad \varepsilon_d(T, S, x) \leq Lip(T) \cdot d(x, S).$$

Proof. Again, the requirements on X and T allow one to ensure that there is a unique point that attains the distance from $co(S)$ to x (note that the convex hull of S is also closed, since S is finite; otherwise, we would need to write a “limit version” of the present result, which we prefer to show in this version for the aim of clarity). So, we fix the point $\sum_{i=1}^n \alpha_i^0 \lambda_i x_i$ by computing the nearest point $\sum_{i=1}^n \alpha_i^0 x_i$ in S to x . Then,

$$\begin{aligned} \left\| \sum_{i=1}^n \alpha_i^0 \lambda_i x_i - T(x) \right\| &= \left\| \sum_{i=1}^n \alpha_i^0 \lambda_i x_i - T\left(\sum_{i=1}^n \alpha_i^0 x_i\right) + T\left(\sum_{i=1}^n \alpha_i^0 x_i\right) - T(x) \right\| \\ &\leq \left\| \sum_{i=1}^n \alpha_i^0 \lambda_i x_i - T\left(\sum_{i=1}^n \alpha_i^0 x_i\right) \right\| + \left\| T\left(\sum_{i=1}^n \alpha_i^0 x_i\right) - T(x) \right\| \\ &= \varepsilon_c(T, S, x) + \varepsilon_d(T, S, x). \end{aligned}$$

On the other hand,

$$\begin{aligned} \varepsilon_c(T, S, x) &= \left\| \sum_{i=1}^n \alpha_i^0 T(x_i) - T\left(\sum_{i=1}^n \alpha_i^0 x_i\right) \right\| \leq \left\| \sum_{i=1}^n \alpha_i^0 \cdot T\left(\sum_{i=1}^n \alpha_i^0 x_i\right) - \sum_{i=1}^n \alpha_i^0 T(x_i) \right\| \\ &\leq Lip(T) \cdot \left(\sum_{i=1}^n \alpha_i^0 \|x_i - \sum_{j=1}^n \alpha_j^0 x_j\| \right), \end{aligned}$$

and

$$\varepsilon_d(T, S, x) \leq \left\| T(x) - T\left(\sum_{i=1}^n \alpha_i^0 x_i\right) \right\| \leq Lip(T) \cdot \left\| x - \sum_{i=1}^n \alpha_i^0 x_i \right\| = Lip(T) \cdot d(x, S). \quad \square$$

Note that the bound obtained above for the convexity error $\varepsilon_c(T, S, x)$ does not allow one to prove that it equals 0 in case T is convex: this has to be done directly by the assumption of certain convexity properties for T , which have to be applied in a second step. If we do not have such information, it could be better to approximate $T(\sum_{i=1}^n \alpha_i^0 x_i)$ by $\lambda_{i_0} x_{i_0}$, where x_{i_0} is the nearest point in S to $\sum_{i=1}^n \alpha_i^0 x_i$, in the line of the minimal distance approximation explained in Section 4.1.

Consequences of this result under reasonable assumptions can be directly obtained. We can control both errors separately since each of them has different meaning.

To obtain the definition of the approximation to T , we try to find a compromise between the approximate convexity expected for T and the error control by using the distance between points, taking into account that T is Lipschitz. We describe the construction of through several steps; as we explained before, we are searching for an extension $\widehat{T} : E \rightarrow E$ for the function T restricted to a subset $S \subseteq ev(T)$, T_S . For helping in the construction process, we first fix an eigenmeasure μ for the set S .

Let us define our general technique under certain restrictions. For the aim of simplicity, we choose a countable set $S \subseteq ev(T)$ to start with, which has the property that for every $R > 0$, $S \cap R \cdot B_E$ is finite. We will use metric tools through the entire process to obtain a pointwise definition of the function \widehat{T} . Take also a positive strictly increasing sequence $(\varepsilon_i)_{i=1}^{+\infty}$ such that $\lim_i \varepsilon_i = +\infty$. The approximation obtained will depend both on the elections of S and $(\varepsilon_i)_{i=1}^{+\infty}$. Fix also $r > 1$.

- Fix a point $x \in X$ and define $d_0 = \inf_{y \in S} \|y - x\|$. Our aim is to work with a *finite* subset of S that is “around” the point x . To take it, consider the ratio $r\|x\|$ to get the set $S \cap (d_0 + r\|x\|) \cdot B_1(x)$. As we said, this ratio r is chosen at the beginning and has to be fixed in order the function \widehat{T} to be univocally defined.
- Take $0 < \varepsilon_1 < r\|x\|$, be sure that r is big enough for this to hold, and define the set $M_{\varepsilon_1}(x) = \{y \in S : \|x - y\| \leq d_0 + \varepsilon_1\}$. Let us write n_{ε_1} for the finite cardinal of this set and write $x_1, \dots, x_{n_{\varepsilon_1}}$ for the elements of $M_{\varepsilon_1}(x)$. Compute the best approximation in $co(M_{\varepsilon_1}(x))$ to x , which is written as $\sum_{i=1}^{n_{\varepsilon_1}} \alpha_i^{\varepsilon_1} x_i$.
- We define the approximation $\widehat{T}^{\varepsilon_1}$ at the point x as $\widehat{T}^{\varepsilon_1}(x) := \sum_{i=1}^{n_{\varepsilon_1}} \alpha_i^{\varepsilon_1} \lambda_i x_i$. We also compute the associated bounds for the errors $\varepsilon_c(T, M_{\varepsilon_1}, x)$ and $\varepsilon_d(T, M_{\varepsilon_1}, x)$, given by the product of $Lip(T)$ and

$$v_c(\varepsilon_1) := \sum_{i=1}^{n_{\varepsilon_1}} \alpha_i^{\varepsilon_1} \|x_i - \sum_{j=1}^{n_{\varepsilon_1}} \alpha_j^{\varepsilon_1} x_j\|, \text{ and } v_d(\varepsilon_1) := d(x, S) = \|x - \sum_{j=1}^{n_{\varepsilon_1}} \alpha_j^{\varepsilon_1} x_j\|$$

respectively.

- We have to balance these errors to choose the best value for the approximation. In general, both bounds are independent, and since we want to obtain a decision rule, we choose a parameter $0 \leq \tau \leq 1$ that will depend on the election of the decision maker. So we will use the value of

$$V(\varepsilon_i) := \tau_1 \cdot v_c(\varepsilon_i) + (1 - \tau) \cdot v_d(\varepsilon_i)$$

for further comparison.

- Consider now ε_2 and repeat the construction of the approximation obtained above for ε_2 , which gives

$$\widehat{T}^{\varepsilon_2}(x) := \sum_{i=1}^{n_{\varepsilon_2}} \alpha_i^{\varepsilon_2} \lambda_i x_i.$$

We compute also the associated bounds and the values of $V(\epsilon_1)$ and $V(\epsilon_2)$.

- After comparing $V(\epsilon_1)$ and $V(\epsilon_2)$, we consider the lower value as the better approximation.
- We follow the same procedure until the last approximation is computed; note that by the construction, the number of steps needed to finish it is finite.

6. Some Hints for Practical Application: An Algorithm for Convex Approximation of Lipschitz Maps in Euclidean Space

We have already fixed some general procedures for approximating a given map by means of the properties that it satisfies. We have focused our attention firstly on Lipschitz maps and secondly on those operators for which some convexity assumption can be assumed.

We assume that we have already fixed a method for computing a significant subset of eigenvectors; we have not developed this question in this paper, but we have suggested that Monte Carlo integration together with our error formulas would help for the determination of such a subset. The scientific literature provides a large source of procedures for doing this (see, for example, [10,14]; see also [17,18,20] for concrete applications in quantum physics).

- For the aim of simplicity, fix E to be a (finite dimensional) Euclidean space; we center the attention in this case because it is the standard case for applications. Once a Lipschitz map $T : E \rightarrow E$ is fixed, the first step consists on computing an approximation to the set $ev(T)$. In order to do it, we use any optimization method.
- The most basic one is a Monte Carlo calculation: just fix a certain $\epsilon > 0$ and define a set $ev_\epsilon(T)$ by taking vectors around the region we want, and satisfying that $\cos(x, T(x)) = |\langle \frac{x}{\|x\|}, \frac{T(x)}{\|T(x)\|} \rangle| > 1 - \epsilon$. The idea is to minimize the diagonalization error provided in the first section using the norm, $\|T(x) - \lambda \cdot x\|$.
- The corresponding eigenvalues can be fixed by using the approximate equation $\|T(x)\| = |\lambda| \|x\|$ and taking the sign of the scalar product for these points.
- Use the set $ev_\epsilon(T)$ and any of the methods explained in Sections 4 and 5 to get an approximation of the errors. The error bounds explained there can be adapted for this case by including the error associated with the previously explained approximation to the set $ev(T)$, which is given by the numbers $\|T(x) - \lambda \cdot x\|, x \in ev_\epsilon(x)$.

Author Contributions: Conceptualization, E.E. and E.A.S.P.; methodology, E.E. and E.A.S.P.; formal analysis, E.E. and E.A.S.P.; investigation, E.E.; writing—original draft preparation, E.A.S.P.; writing—review and editing, E.E. and E.A.S.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by the Grant PID2020-112759GB-I00 funded by MCIN/AEI/10.13039/501100011033 and by “ERDF A way of making Europe”.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Appell, J.; de Pascale, E.; Vignoli, A. *Nonlinear Spectral Theory*; Walter de Gruyter: Berlin, Germany, 2004; Volume 10.
2. Fučík, S.; Nečas, J.; Souček, V.; Souček, J. *Spectral Analysis of Nonlinear Operators*; Springer: Berlin, Germany, 1973.
3. da Silva, E.B.; Fernandez, D.L.; de Andrade Neves, M.V. A spectral theorem for bilinear compact operators in hilbert spaces. *Banach J. Math. Anal.* **2021**, *15*, 28. [CrossRef]
4. da Silva, E.B.; Fernandez, D.L.; de Andrade Neves, M.V. Schmidt representation of bilinear operators on hilbert spaces. *Indag. Math.* **2021**, in press. [CrossRef]
5. Phillips, J.R. Eigenfunction Expansions for Self-Adjoint Bilinear Operators in Hilbert Space. Ph.D. Thesis, Oregon State University, Corvallis, OR, USA, 1966.
6. Krasnoselskij, M.A. On a topological method in the problem of eigenfunctions for nonlinear operators. *Dokl. Akad. Nauk. SSSR* **1950**, *74*, 5–7. (In Russian)
7. Nemytskij, V.V. Structure of the spectrum of completely continuous nonlinear operators. *Mat. Sb.* **1953**, *33*, 545–558. (In Russian)
8. Kachurovskij, R.I. Regular points, spectrum and eigenfunctions of nonlinear operators. *Sov. Math. Dokl.* **1969**, *10*, 1101–1105.
9. Calamai, A.; Furi, M.; Vignoli, A. An overview on spectral theory for nonlinear operators. *Commun. Appl. Anal.* **2009**, *13*, 509–534.
10. Dörfner, M. A numerical range for nonlinear operators in smooth banach spaces. *Z. Anal. Anwend.* **1996**, *15*, 445–456. [CrossRef]

11. Choi, G.; Jung, M.; Tag, H.-J. On the lipschitz numerical index of banach spaces. *arXiv* **2021**, arXiv:2110.13821.
12. Dong, X.; Wu, D. The kachurovskij spectrum of lipschitz continuous nonlinear block operator matrices. *Adv. Oper. Theory* **2021**, *6*, 1–18. [[CrossRef](#)]
13. Albu-Schaeffer, A.; Della Santina, C. A review on nonlinear modes in conservative mechanical systems. *Annu. Rev. Control* **2020**, *50*, 49–71. [[CrossRef](#)]
14. de Raedt, H.; Frick, M. Stochastic diagonalization. *Phys. Rep.* **1993**, *231*, 107–149. [[CrossRef](#)]
15. Mendes, R.V. Resonant hole configurations and hole pairing. *J. Phys. Condens. Matter* **1991**, *3*, 6757–6761. [[CrossRef](#)]
16. Lee, D.; Salwen, N.; Lee, D. The diagonalization of quantum field hamiltonians. *Phys. Lett. B* **2001**, *503*, 223–235. [[CrossRef](#)]
17. Mendes, R.V. Stochastic techniques in condensed matter physics. In *Stochastic Analysis and Applications in Physics*; Springer: Dordrecht, The Netherlands, 1994; pp. 239–281.
18. Subhajt, N.; Chaudhury, P.; Bhattacharyya, S.P. Stochastic diagonalization of hamiltonian: A genetic algorithm-based approach. *Int. J. Quantum Chem.* **2002**, *90*, 188–194.
19. Lee, D.; Salwen, N.; Windolowski, M. Introduction to stochastic error correction methods. *Phys. Lett. B* **2001**, *502*, 329–337. [[CrossRef](#)]
20. Husslein, T.; Fettes, W.; Morgenstern, I. Comparison of calculations for the hubbard model obtained with quantum-monte-carlo, exact, and stochastic diagonalization. *Int. J. Mod. Phys. C* **1997**, *8*, 397–415. [[CrossRef](#)]
21. Spencer, J.S.; Blunt, N.S.; Choi, S.; Etrych, J.; Filip, M.A.; Foulkes, W.M.C.; Franklin, R.S.; Handley, W.J.; Malone, F.D.; Neufeld, V.A.; et al. The hande-qmc project: Open-source stochastic quantum chemistry from the ground state up. *J. Chem. Theory Comput.* **2019**, *15*, 1728–1742. [[CrossRef](#)]
22. Diestel, J.; Jarchow, H.; Tonge, A. *Absolutely Summing Operators*; Cambridge University Press: Cambridge, UK, 1995; Volume 43.
23. Diestel, J.; Uhl, J.J. *Vector Measures*; American Mathematical Society: Providence, RI, USA, 1977.
24. Okada, S.; Ricker, W.; Sánchez Pérez, E.A. *Optimal Domain and Integral Extension of Operators*; Operator Theory: Advances and Applications; Springer: Berlin, Germany, 2008.
25. Cobzaş, Ş.; Miculescu, R.; Nicolae, A. *Lipschitz Functions*; Lecture Notes in Mathematics; Springer: Cham, Switzerland, 2019; Volume 2241.
26. Erdoğan, E.; Sánchez Pérez, E.A. Eigenmeasures and stochastic diagonalization of bilinear maps. *Math. Meth. Appl. Sci.* **2021**, *44*, 5021–5039. [[CrossRef](#)]
27. Aliprantis, C.D.; Border, K. *Infinite Dimensional Analysis*; Springer: Berlin, Germany, 1999.
28. Mezić, I. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dyn.* **2005**, *41*, 309–325. [[CrossRef](#)]
29. Rowley, C.W.; Mezić, I.; Bagheri, S.; Schlatter, P.; Henningson, D.S. Spectral analysis of nonlinear flows. *J. Fluid Mech.* **2009**, *641*, 115–127. [[CrossRef](#)]
30. Krylov, S.; Lurie, K.; Yakobovitz, A. Compliant structures with time-varying moment of inertia and non-zero averaged momentum and their application in angular rate microsensors. *J. Sound Vib.* **2011**, *330*, 4875–4895. [[CrossRef](#)]
31. Tikhov, S.V.; Valovik, D.V. Perturbation of nonlinear operators in the theory of nonlinear multifrequency electromagnetic wave propagation. *Commun. Nonlinear Sci. Numer. Simul.* **2019**, *75*, 76–93. [[CrossRef](#)]
32. Valovik, D.V. Nonlinear multi-frequency electromagnetic wave propagation phenomena. *J. Opt.* **2017**, *19*, 115502. [[CrossRef](#)]
33. Jiří, B.; Drábek, P. Estimates of the principal eigenvalue of the p-biharmonic operator. *Nonlinear Anal. Theory Methods Appl.* **2012**, *75*, 5374–5379.
34. Kong, L.; Nichols, R. On principal eigenvalues of biharmonic systems. *Commun. Pure Appl. Math.* **2021**, *20*, 1–15. [[CrossRef](#)]
35. Amann, H. Fixed point equations and nonlinear eigenvalue problems in ordered banach spaces. *SIAM Rev.* **1976**, *18*, 620–709. [[CrossRef](#)]
36. Dukić, D.; Paunović, L.; Radenović, S. Convergence of iterates with errors of uniformly quasi-lipschitzian mappings in cone metric spaces. *Kragujev. J. Math.* **2011**, *35*, 399–410.
37. Henderson, J.; Wang, H. Nonlinear eigenvalue problems for quasilinear systems. *Comput. Math. Appl.* **2005**, *49*, 1941–1949. [[CrossRef](#)]
38. Todorčević, V. *Harmonic Quasiconformal Mappings and Hyperbolic Type Metrics*; Springer: Cham, Switzerland, 2019.