NONLINEAR EIGENVALUE PROBLEMS FOR SEMINORMS AND APPLICATIONS

MARTIN BURGER

Dedicated to the Memory of Joyce McLaughlin and Victor Isakov.

ABSTRACT

The aim of this paper is to discuss recent progress in nonlinear eigenvalue problems for seminorms (absolutely one-homogeneous convex functionals), which find many applications in data science, inverse problems, and image processing. We provide a unified viewpoint on the notion of nonlinear singular vectors and eigenvectors for homogeneous nonlinear operators respectively functionals. We further discuss in particular ground states, i.e., the first eigenvector or eigenfunction. Moreover, we review a recent approach to the analysis of eigenvectors based on duality, which has implications to the possible computation of spectral decompositions, i.e., signal dependent linear expansions in a system of eigenvectors.

Moreover, we discuss some relevant implications such as the refined analysis of variational regularization methods and their bias, as well as the analysis of some iteration methods and time-continuous flows. Finally, we provide more direct applications of the nonlinear eigenvalue problems such as nonlinear spectral clustering on graphs.

MATHEMATICS SUBJECT CLASSIFICATION 2020

Primary 35P30; Secondary 47J10, 68T09, 94A08

KEYWORDS

Nonlinear eigenvalue problems, spectral decomposition, spectral clustering, gradient flows, variational methods

ICM INTERNATIONAL CONGRESS OF MATHEMATICIANS 2022 JULY 6—14

Proc. Int. Cong. Math. 2022, Vol. 7, pp. 5234–5255
DOI 10.4171/ICM2022/140

1. INTRODUCTION

Eigenvalue problems are not just a basic technique in linear algebra (cf. [39, 40]), they also find many applications in several branches of sciences, more recently also in data or image analysis. Prominent examples are the computation of states in quantum mechanics, Fourier decompositions—i.e., expansion in Laplacian eigenvalues—of audio signals or images (cf., e.g., [8]), or spectral clustering based on graph Laplacians (cf. [42]). In most of these applications the eigenvector or eigenfunction is of more importance than the exact eigenvalue, e.g., spectral clustering is based on dividing into the sets where the first nontrivial eigenfunction is negative or positive, respectively. Thus, particular focus is pot on the computation of eigenvectors respectively eigenfunctions.

While eigenvalue problems for linear operators are well understood, nonlinear eigenvalue problems, in particular those being nonlinear in the eigenvector or eigenfunction (cf. [1, 38]), are still a lively topic with many different directions of research. In physics, eigenvalue problems for nonlinear Schrödinger equations are a prominent example (cf. [24, 43]), while eigenvalue problems for p-Laplacian operators (and their graph equivalents) received strong recent attention in partial differential equations and data science (cf., e.g., [12, 21, 33–35]).

In this paper we want to focus on a special type of eigenvalue problems for (positively) zero-homogeneous operators related to the subdifferential of absolutely one-homogeneous functionals, more precisely we look for $\lambda > 0$ and $u \in H$, H a Hilbert space, such that

$$\lambda u \in \partial J(u). \tag{1.1}$$

Here $J : H \to \mathbb{R} \cup \{+\infty\}$ is assumed to be convex and absolutely one-homogeneous, thus it is effectively a seminorm on a subspace of H (cf. [15]). The assumption of one-homogeneity is less restrictive than it seems, since many other homogeneous eigenvalue problems can be reformulated equivalently as one-homogeneous problems, as we shall see in the p-Laplacian case below. Such eigenvalue problems can be rephrased in a variational setting, since we look for stationary points of the Rayleigh-quotient

$$R(u) = \frac{J(u)}{\|u\|}. \tag{1.2}$$

Indeed, (iterative) minimization of the Rayleigh quotient is a key technique for the computation of eigenvectors or eigenfunctions (cf. [13, 25, 29, 31, 32]).

Let us mention a related notion of nonlinear singular values (cf. [3]), given by

$$\lambda K^* K u \in \partial J_0(u), \tag{1.3}$$

where $J_0 : X \to \mathbb{R} \cup \{+\infty\}$ is a convex and absolutely one-homogeneous functional on a Banach space X, and $K : X \to Y$ is a bounded linear operator into the Hilbert space Y. This notion generalizes the linear singular value problem

$$K^* K u = \sigma^2 u, \tag{1.4}$$

with the obvious relation $\sigma = \frac{1}{\sqrt{\lambda}}$ to a nonlinear setting, and finds interesting applications in the regularization theory of inverse problems (cf. [3, 4]). We shall see below that indeed there

is a reformulation of the singular value problem (1.3) as a nonlinear eigenvalue problem of the form (1.1).

Besides some basic issues for eigenvalue problems and their direct applications, we will also discuss the issue of *spectral decompositions* (cf. [15, 18, 19, 27, 28, 30]), i.e., the possibility to develop signals in a systematic way into nonlinear eigenvectors, e.g., as a sum

$$f = \sum_{k=1}^{\infty} c_k u_k,$$

for $f \in H$ and u_k being the eigenvector with eigenvalue λ_k. In a general setting we rather look for a decomposition of the form

$$f = \int_0^{\infty} d\phi_\lambda, \tag{1.5}$$

with a measure ϕ on $\mathbb{R}_+$ valued in the Hilbert space H. Such a decomposition will be called spectral decomposition if the polar composition

$$\phi_\lambda = u_\lambda |\phi_\lambda| \tag{1.6}$$

is such that for each λ in the support of $|\phi_\lambda|$ the unit vector $u_\lambda \in H$ is an eigenvector for the eigenvalue λ. Fundamental questions, only partly answered so far, are the existence of nonlinear spectral decompositions as well as a systematic way to compute such decompositions from data. A particular advantage of a spectral decomposition is the possibility to define filtered versions of f,

$$f_\psi = \int_0^{\infty} \psi(\lambda)\, d\phi_\lambda, \tag{1.7}$$

e.g., with ψ being zero on a certain interval to suppress certain scales related to a range of eigenvalues. Such approaches find applications, e.g., in image or geometry processing (cf. [26, 30]). Moreover, the spectral decompositions of two different data f_1 and f_2 can be mixed, which finds interesting applications, e.g., in image fusion (cf. [5]).

The remainder of this paper is organized as follows: In Section 2 we provide some notations and fundamental properties of eigenvalue problems for seminorms, as well as first examples. We also discuss the motivation for a nonlinear spectral decomposition. Section 3 is devoted to the study of ground states, the eigenvectors for the first nontrivial eigenvalue, which are of particular relevance and also the easiest to compute numerically. Section 4 discusses the relation between eigenvalue problems, on the one hand, and variational methods, iterative schemes, and time-continuous flows, on the other. Here we see that eigenvectors and eigenfunctions yield structured examples of exact solutions for those methods. On the other hand, these methods, in particular gradient flows and time-continuous versions, can be used to compute eigenvectors and possibly even spectral decompositions.

2. BASIC PROPERTIES AND FORMULATIONS

In the following we fix some notation, discuss some basic properties of nonlinear eigenvalue problems such as (1.1), and unify the formulations of eigenvalues and singular values.

2.1. Seminorms, duality, and subdifferentials

Throughout the whole paper we assume that $J : H \to \mathbb{R} \cup \{+\infty\}$ is convex and absolutely one-homogeneous, i.e.,

$$J(tu) = |t|J(u), \quad \forall t \in \mathbb{R}. \tag{2.1}$$

This implies that J satisfies a triangle inequality, since

$$J(u_1 + u_2) = 2J\left(\frac{1}{2}u_1 + \frac{1}{2}u_2\right) \leq 2\left(\frac{1}{2}J(u_1) + \frac{1}{2}J(u_2)\right) = J(u_1) + J(u_2).$$

Moreover, the set

$$H_0 = \{u \in H \mid J(u) < \infty\}$$

is a subspace of H on which J is a seminorm, hence our nomenclature as eigenvalue problems for seminorms.

For completeness, let us recall the definition of the subdifferential of a convex functional J,

$$\partial J(u) = \{p \in H^* \mid \langle p, v - u\rangle \leq J(v) - J(u), \forall v \in H\}, \tag{2.2}$$

and the polar function (or convex dual),

$$J^*(p) = \sup_{u \in H} \langle p, u\rangle - J(u). \tag{2.3}$$

Note that for $p \in \partial J(u)$ we have

$$\langle p, u\rangle = J(u)$$

and

$$\langle p, v\rangle \leq J(v)$$

for each $v \in H$, and these properties are actually an equivalent characterization of subgradients under our assumptions (cf. [15]). Since J is a norm on a subspace, we can define a dual norm

$$\|p\|_* = \sup_{u \in H, J(u) \leq 1} \langle p, u\rangle, \tag{2.4}$$

which is interesting for the analysis of subgradients. Indeed, it can be shown that

$$\partial J(u) \subset \partial J(0) = \{p \in H^* \mid \|p\|_* \leq 1\},$$

for each $u \in H$, i.e., subdifferentials are contained in the dual unit ball.

The eigenvalue problem (1.1) can be interpreted in a dual way, by noticing that each eigenvector u is also a multiple of a subgradient p, respectively as $p \in \lambda \partial J^*(p)$. A key observation made in [15] is that these subgradients arising in the eigenvalue problems are of minimal norm.

Proposition 2.1. *Let u be an eigenvector of J satisfying $\lambda u = p \in \partial J(u)$. Then p is a subgradient of minimal norm, i.e.,*

$$\|p\| \leq \|q\|, \quad \forall q \in \partial J(u).$$

In [15] a further geometric characterization of eigenvectors has been derived, which relates to the minimal norm property.

Proposition 2.2. *An element $p \in \partial J(0)$ defines an eigenvector $u = \frac{1}{\lambda} p$ for some $\lambda > 0$ if and only if p satisfies the extremal property,*

$$\langle p, p - q \rangle \geq 0, \quad \forall q \in \partial J(0). \tag{2.5}$$

At least from a theoretical point of view, this yields an option to obtain all eigenvectors of functional as follows: first of all, compute for each $u \in H$ with $\|u\| = 1$ the subgradient of minimal norm, i.e.,

$$p = \arg\min\{\|q\| \mid q \in \partial J(u)\},$$

and subsequently check condition (2.5). In case of satisfaction, u is an eigenvector.

Example 2.3. Consider the simple example $J(u) = \sqrt{\langle u, Au \rangle}$ for a positive semidefinite operator A. In this case

$$\partial J(u) = \frac{1}{J(u)} Au$$

for $u \neq 0$, and it is easy to see that

$$\partial J(0) = \{p = Aw \mid w \in H \langle w, Aw \rangle \leq 1\}.$$

Let u be a linear eigenvector with eigenvalue $\lambda \neq 0$, i.e., $\lambda u = Au$, then (2.5) with $p = \frac{1}{J(u)} Au = \frac{\lambda}{J(u)} u$ becomes

$$\frac{\lambda}{J(u)} \left\langle u, \frac{1}{J(u)} Au - Aw \right\rangle \geq 0.$$

This is satisfied, since it is equivalent to

$$\sqrt{\langle u, Au \rangle} \geq \langle u, Aw \rangle,$$

and this inequality holds due to the Cauchy–Schwarz inequality in the scalar product induced by A.

Example 2.4. Consider a polyhedral functional, i.e.,

$$J(u) = \chi_C^* = \sup_{p \in C} \langle p, u \rangle,$$

with the symmetric polyhedral set

$$C = \operatorname{conv}(\{p_1, \ldots, p_m, -p_1, \ldots, -p_m\}).$$

Then p_j satisfies (2.5) if the plane orthogonal to p_j only intersects C in p_j.

Let us make this more concrete in $\mathbb{R}^2$ in polyhedra with $m = 2$. We start with the example $p_1 = (1, 1)$ and $p_2 = (-1, 1)$, i.e., C is the unit ball in ℓ^∞. The lines orthogonal to $\pm p_j$ only intersect C in p_j, thus all p_j are eigenvectors. As a specific case, we explicitely compute (2.5) for p_1 and $q = (r, s) \in C$,

$$\langle p_1, p_1 - q \rangle = 2 - r - s \geq 0,$$

since $r, s \in [-1, 1]$.

As a second example consider $p_1 = (1, 1)$ and $p_2 = (\varepsilon, 1)$ with $\varepsilon \in (0, 1)$. With the reasoning as above, we see that p_1 is an eigenvector. However, p_2 is not as we see with $q = (1, 1) \in C$,

$$\langle p_2, p_2 - q\rangle = \varepsilon^2 + 1 - \varepsilon - 1 = \varepsilon(\varepsilon - 1) < 0.$$

2.2. Singular values and eigenvalues

In the following we discuss the reformulation of the nonlinear singular value problem (1.3) as a nonlinear eigenvalue problem. Throughout this section we assume that J_0 is a seminorm on a subspace of X extended by $+\infty$, and $K : X \to Y$ is a bounded linear operator. In order to unify the formulation with (1.1), we want to define a functional J on a subspace of Y, respectively on values $v = Ku$. Hence, it is natural to define the space H as the closure of the range of K in Y. If K has a nontrivial nullspace, the definition of $J_0(v)$ as $J(u)$ with $Ku = v$ is not unique, however. We thus first provide a property of eigenvectors that will enable a unique definition.

Lemma 2.5. *Let u be a nonlinear singular vector according to (1.3). Then $J_0(u) \le J_0(w)$ for all w such that $Kw = Ku$.*

Proof. We take a duality product of λK^*Ku with $u - w$ to obtain

$$\lambda\big\langle K^*Ku, w - u\big\rangle = \lambda\langle Ku, Ku - Kw\rangle = 0.$$

On the other hand, from the singular value equation (1.3) we find

$$\lambda\big\langle K^*Ku, w - u\big\rangle = \langle p, w - u\rangle = \langle p, w\rangle - J_0(u) \le J_0(w) - J_0(u).$$

Hence, $J_0(u) \le J_0(w)$. ∎

From this result we see that we need to define J via the minimal value of J_0, more precisely

$$J(v) := \inf_{u, Ku = v} J_0(u). \tag{2.6}$$

It is straightforward to check that $J : H \to \mathbb{R} \cup \{+\infty\}$ is an absolutely one-homogeneous convex functional. Moreover, there is a direct relation between subgradients: we find $p \in \partial J(Ku)$ if and only if $K^*p \in \partial J_0(u)$. Thus, we find the equivalence between

$$\lambda v = \lambda Ku \in \partial J(v) \quad \text{and} \quad \lambda K^*Ku = \lambda K^*v \in \partial J_0(u).$$

2.3. Spectral decomposition

An interesting question is the possible existence of a spectral decomposition in the nonlinear case. Let us recall the well-known spectral decomposition of a positive semi-definite linear operator A on a Hilbert space H: there exists an operator-valued spectral measure E supported on the spectrum of A such that

$$A = \int_0^\infty \lambda \, dE_\lambda.$$

This allows extending functions $f : \mathbb{R}^+ \to \mathbb{R}$ to the operator A as

$$f(A) = \int_0^\infty f(\lambda)\, dE_\lambda.$$

In the case of a compact operator, the spectral measure is concentrated on a countable set and takes the form

$$E = \sum_{k=1}^{\infty} u_k \otimes u_k \delta_{\lambda_k},$$

where u_k is an eigenvector for the eigenvalue λ_k and $\otimes$ denotes the outer product. A positive semidefinite linear operator is the canonical choice in our setting, since it defines a convex absolutely one-homogeneous functional

$$J(u) = \sqrt{\langle u, Au \rangle}.$$

In general, we cannot expect to obtain some kind of spectral decomposition from a convex functional J, respectively its subdifferential ∂J, but we can hope to have a pointwise decomposition, corresponding in the linear case to

$$Au = \int_0^\infty \lambda\, d(E_\lambda u) = \int_0^\infty \lambda\, d\phi_\lambda,$$

with a spectral measure ϕ valued in the Hilbert space H. In particular, this allows for the reconstruction of u from the spectral measure via

$$u = \int_0^\infty d\phi_\lambda,$$

as well as some spectral filtering by integrating some function of λ, e.g., a characteristic function in some region.

In general, there is no unique way to construct a unique spectral decomposition of this kind. For example, for total variation regularization in one dimension (with appropriate definition of the variation on the boundary), it was shown in [3] that the Haar wavelet basis is an orthogonal basis of nonlinear eigenfunctions, hence there exists an atomic spectral decomposition in this basis. However, it also has been shown that there is a continuum of further eigenfunctions, necessarily linearly dependent, hence further spectral decompositions can be obtained by exchanging parts of the Haar wavelet basis. An interesting question is to define a generic spectral decomposition by a natural technique.

3. GROUND STATES

In the following we investigate the first nontrivial eigenvalue and its corresponding eigenvector or eigenfunction, which we call ground state. More precisely, let

$$\mathcal{N}(J) = \{u \in H \mid J(u) = 0\}$$

be the nullspace of J. Due to the properties of J, the nullspace is indeed a linear subspace (cf. [3,15]), and we can define its orthogonal complement $H_0 := \mathcal{N}(J)^\perp$ in H. It can further be shown that for each $u \in H$ and $u_0 \in \mathcal{N}(J)$ the identity

$$J(u + u_0) = J(u) + J(u_0)$$

holds, i.e., there is an analogue of the orthogonal decomposition at the level of J. Finally, let $u \in H$ be an eigenvector for the eigenvalue $\lambda \neq 0$. Then we find for $u_0 \in \mathcal{N}(J)$,

$$\lambda \langle u, u_0 \rangle = \langle p, u_0 \rangle \leq J(u + u_0) - J(u) = 0.$$

Since $-u$ is also an eigenvector, we obtain the opposite inequality and hence the orthogonality of u and u_0. Thus, we get rid of the trivial eigenvalues and the corresponding eigenvectors by restriction to H_0, which leads in particular to the definition of a ground state according to [3].

Definition 3.1. Let $J : H \to \mathbb{R} \cup \{+\infty\}$ be an absolutely one-homogeneous convex functional and H_0 be the orthogonal of its nullspace as above. Then we call $u \in H$ a *ground state* of J if

$$u \in \arg\min_{u \in H_0} \frac{J(u)}{\|u\|}. \tag{3.1}$$

Let us mention that we can rescale u in the above definition and consider equivalently a ground state as a minimizer of J on the unit sphere $\{u \in H_0 \mid \|u\| = 1\}$. The latter is useful for proving the existence of ground states. If J is lower semicontinuous and the sublevel sets of J are precompact, existence follows from a standard argument (cf. [3]). It is apparent for normalized eigenvectors u that $\lambda = J(u)$, thus

$$\lambda_0 := \min_{u \in H_0} \frac{J(u)}{\|u\|} \leq \lambda$$

for each nontrivial eigenvalue. On the other hand, λ_0 is indeed an eigenvalue for each eigenvector u_0 minimizing the Rayleigh-quotient. For this, define $p_0 = \lambda_0 u_0$. Then we have

$$\langle p_0, u_0 \rangle = \lambda_0 \langle u_0, u_0 \rangle = \lambda_0 = J(u_0),$$

and for arbitrary $u \in H \setminus \{0\}$,

$$\langle p_0, u \rangle = \lambda_0 \langle u_0, u \rangle \leq \lambda_0 \|u\| \leq \frac{J(u)}{\|u\|} \|u\| = J(u).$$

Thus $p_0 = \lambda_0 u_0 \in \partial J(u_0)$.

We finally recall the relation to the case of nonlinear singular values. The ground state in this case can be equivalently computed from minimizing

$$u \in \arg\min_{u \in X} \frac{J_0(u)}{\|Ku\|}, \tag{3.2}$$

which is often more accessible.

3.1. p-Laplacian eigenvalues

Ground states of the p-Laplacian are a well-studied problem in partial differential equations, as well as on graphs. In the standard setting, one would look for the first eigenvalue in the problem

$$-\nabla \cdot \left(|\nabla u|^{p-2} \nabla u \right) = \lambda_1 u |u|^{p-2},$$

in a domain Ω with homogeneous Neumann or Dirichlet boundary values. This is, however, related to the eigenvalue of the p-Laplacian energy

$$E_p(u) = \int_\Omega \left|\nabla u(x)\right|^p dx$$

in $L^p(\Omega)$, while our Hilbert space setting corresponds to solving

$$-\nabla \cdot \left(|\nabla u|^{p-2}\nabla u\right) = \lambda_1 u \|u\|_{L^2}^{p-2}.$$

Since $u \mapsto E_p(u)^{1/p}$ is an absolutely one-homogeneous convex functional, the ground state can be computed by minimizing the corresponding Rayleigh quotient

$$R(u) = \frac{E_p(u)^{1/p}}{\|u\|_{L^2}},$$

which corresponds to our setup in this paper. For $p = 2$, all formulations simply yield the standard linear eigenvalue problem for the Laplacian and, indeed, the formulation with the Rayleigh quotient is related to the fact that the first eigenvalue is the best constant in the Poincaré-inequality. On graphs, the corresponding problem for the graph Laplacian is fundamental for spectral clustering techniques (cf. [42]).

Particularly interesting cases are, of course, the limiting ones $p = 1$ and $p = \infty$. For $p = 1$, the ground state is the first eigenfunction of total variation, and, due to area and coarea formula, the L^2-norm and total variation can be related to the volume, respectively perimeter, of level sets (cf. [22]). In this way and similar to the classical Cheeger problems (cf. [37]), it can be shown that, indeed, ground states only take two-function values and the interface between solves an isoperimetric problem. On a graph the ground states of total variation can be related in a similar way to a graph cut, the so-called Cheeger cut (cf. [41]). In one dimension, for a modified version of total variation that takes into account also the variation across the boundary (assuming extension by zero outside), the ground state can be computed as a piecewise constant function with single discontinuity in the midpoint of the interval. For this approach, also scaling of the eigenfunction is possible. For simplicity, consider $\Omega = (0, 1)$ and let u_1 be the ground state. Then indeed, for $s < 1$, the function

$$u_1^s(x) = \begin{cases} \frac{1}{\sqrt{s}} u_1(\frac{x}{s}) & \text{if } x < s, \\ 0 & \text{if } x > s \end{cases}$$

is another eigenfunction for a larger eigenvalue. Moreover, the dilation

$$u_1^{s,t}(x) = \begin{cases} 0 & \text{if } x < t, \\ \frac{1}{\sqrt{s}} u_1(\frac{x-t}{s}) & \text{if } x < s + t, \\ 0 & \text{if } x > s + t \end{cases}$$

is another eigenfunction if $t \leq 1 - s$. Indeed, such results can be generalized to anisotropic total variation in multiple dimension by scaling and dilation along the coordinate axes.

In the case $p = \infty$, the setup in a Hilbert space is not the one usually referred to as ∞-Laplacian, which rather corresponds to the treatment of the ∞-Laplacian energy

$$J(u) = \|\nabla u\|_\infty$$

in $L^\infty(\Omega)$, while we use the Hilbert space $H = L^2(\Omega)$. In this case an interesting problem is to consider an extended version of J, defined as $+\infty$ if u does not have homogeneous boundary values, i.e., we effectively solve the homogeneous Dirichlet problem. Here the ground state can be computed more easily with a different scaling $J(u) = 1$, i.e., we effectively maximize $\|u\|_2$ subject to $|\nabla u(x)| \leq 1$ almost everywhere. For a domain Ω, this is indeed the case if u is the distance function to the boundary, thus computing the ground state yields an alternative way to compute distance functions, respectively solve the eikonal equation (cf. [17] for further details). On a graph, the computation of the ground state with the above normalization yields a way to define and compute a distance function.

3.2. Ground states and sparsity

An interesting class of functionals J in many applications in signal and image processing are ℓ^1-norms or their continuum counterpart, the total variation of a measure. Such functionals are used to enforce sparsity of their minimizers, i.e., a minimal size of the support.

Let us start in the finite-dimensional case $X = \mathbb{R}^n$, $K : \mathbb{R}^n \to \mathbb{R}^m$ (the latter equipped with the Euclidean norm), and

$$J_0(u) = \|u\|_1 = \sum_{i=1}^{n} |u_i|.$$

Here we use the singular value formulation (3.2), i.e., we want to compute

$$u \in \arg\min_{u \in X} \frac{\|u\|_1}{\|Ku\|_2}.$$

Indeed, the sparsity is present also in the ground state. To see this, let e_i be the ith unit vector and $\tilde{e}_i = \operatorname{sign}(u_i) e_i$. Then, for arbitrary u, we have

$$\|Ku\|_2 = \|u\|_1 \left\| \sum_{i=1}^{n} \sigma_i K \tilde{e}_i \right\|_2,$$

with $\sigma_i = \frac{|u_i|}{\|u\|_1}$. By convexity, we find further

$$\|Ku\|_2 \leq \|u\|_1 \sum_{i=1}^{n} \sigma_i \|K\tilde{e}_i\|_2,$$

and equality holds if u has a single nonzero entry. In particular, we find

$$R(u) = \frac{\|u\|_1}{\|Ku\|_2} \geq \frac{1}{\max_i \|Ke_i\|_2},$$

and thus $u = e_j$ with j such that

$$\|Ke_j\|_2 \geq \|Ke_i\|_2, \quad \forall i \in \{1, \dots, n\}$$

is a ground state. Moreover, the proof shows that there are no other ground states, i.e., all of them have maximal sparsity.

In the infinite-dimensional case, we have $X = \mathcal{M}(\Omega)$ for some domain Ω and K mapping to some Hilbert space, typically with the assumption that K is the adjoint of an

operator L mapping from the Hilbert space to the predual space $C_b(\Omega)$ (cf. [9]). Using the latter issues with the dual space of $\mathcal{M}(\Omega)$ can be avoided and the analysis can be carried out in the predual space. The functional J_0 is the total variation norm

$$J_0(u) = \sup_{\varphi \in C_b(\Omega), \|\varphi\|_\infty \leq 1} \int_\Omega \varphi \, du.$$

With analogous reasoning as above, interpreting a general measure after division by its total variation norm as a convex combination of signed concentrated measures, we see that the ground states are of the form $u = \delta_z$ for $z \in \Omega$ such that

$$\|K\delta_z\| \geq \|K\delta_x\|, \quad \forall x \in \Omega.$$

Example 3.2. Let us consider $\Omega \subset \mathbb{R}^d$ and let $k : \mathbb{R}^d \to \mathbb{R}$ be a continuous and integrable kernel. We consider the convolution operator

$$K : \mathcal{M}(\Omega) \to L^2(\mathbb{R}^d), \quad u \mapsto \int_\Omega k(\cdot - y) du(y).$$

We see that $K\delta_z = k(\cdot - z)$ and thus

$$\|K\delta_z\|_{L^2}^2 = \int_{\mathbb{R}^d} k(x - z)^2 \, dx = \int_{\mathbb{R}^d} k(y)^2 \, dy.$$

Hence, the maximum of $\|K\delta_z\|$ is attained for any $z \in \Omega$, which implies that the concentrated measure δ_z is a ground state for any $z \in \Omega$.

Example 3.3. We return to the case of a polyhedral regularization $J = \chi_C^*$, but actually the argument holds for general convex sets C. Since we know that the ground state is an eigenvector and thus a subgradient in $\partial J(0) = C$, it suffices to minimize the Rayleigh quotient over C. Moreover, the extremal property (2.5) can be satisfied only for $p \in \partial C$, hence we further restrict the possible minimization. Let p be the solution of

$$p = \arg\min_{q \in \partial C} \|q\|,$$

i.e., the element of minimal norm in ∂C. Then

$$R(p) = \frac{J(p)}{\|p\|} = \frac{1}{\|p\|} \sup_{q \in C} \langle q, p \rangle = \frac{1}{\|p\|} \langle p, p \rangle = \|p\|.$$

By analogous reasoning, we can show

$$R(q) \geq \|q\| \geq \|p\| = R(p).$$

Thus, p is indeed a ground state of J.

4. VARIATIONAL PROBLEMS, ITERATIONS, AND FLOWS

A first motivation of the definition of nonlinear singular values was to obtain exact solutions of variational problems of the form

$$u \in \arg\min_{u \in X} \frac{1}{2} \|Ku - f\|^2 + \alpha J_0(u),$$

which frequently arise in the regularization of inverse problems, image processing, and data analysis (cf. [4] and the references therein). It was further extended to some iterative methods and time-continuous flows, which in turn can be used as methods to compute eigenvectors or singular vectors. We shall discuss these developments in the following. For the sake of a simple notation, we denote by u_λ an eigenvector with eigenvalue λ, i.e.,

$$\lambda u_\lambda = p_\lambda \in \partial J(u_\lambda). \tag{4.1}$$

Moreover, throughout the whole section we will use data $f \in H_0$, since for arbitrary $f \in H$ we can factor out the component in $\mathcal{N}(J)$. The latter is just technical and beyond our interest of highlighting the main points of the analysis in this paper.

4.1. Variational regularization methods

We start with a discussion of variational regularization methods, which we rephrase as in Section 2 in a Hilbert space setting, i.e.,

$$u \in \arg\min_{u \in H} \frac{1}{2}\|u - f\|^2 + \alpha J(u). \tag{4.2}$$

We consider f being a multiple of the eigenvector u_λ, i.e., $f = cu_\lambda$, $c > 0$, and look for a solution of the form $u = C(\alpha, \lambda)u_\lambda$. The basis for this investigation is the optimality condition

$$u - f + \alpha p = 0, \quad p \in \partial J(u)$$

satisfied by a solution u of (4.2). Making the ansatz $p = p_\lambda = \lambda u_\lambda$, which is in the subdifferential of $C(\alpha, \lambda)u$ due to the zero-homogeneity of ∂J, we arrive at the scalar relation

$$C(\alpha, \lambda) - c + \alpha\lambda = 0,$$

which yields a positive solution if $c > \alpha\lambda$. If $c \leq \alpha\lambda$, we obtain a solution by choosing $C(\alpha, \lambda) = 0$, since $\frac{c}{\alpha\lambda} \in \partial J(0)$. Thus, we find

$$C(\alpha, \lambda) = (c - \alpha\lambda)_+, \tag{4.3}$$

i.e., the solution is a multiple of an eigenvector with the magnitude obtained by a shrinkage formula. We see that obviously the shrinkage is stronger for larger α, but also for larger λ. Hence, there is less change in smaller eigenvalues (low frequencies) than in larger eigenvalues (high frequencies).

The solutions of this kind can be investigated with respect to their robustness with respect to noise (errors in f) and bias (errors due to positive values of α), see [3]. Let us detail some aspects of bias in the following, a particularly interesting property is that the ground state yields the minimal bias (cf. [3]).

Theorem 4.1. *Let $\alpha > 0$ and $u \notin \mathcal{N}(J)$ be a solution of (4.2). Then*

$$\|u - f\| \geq \alpha\lambda_0,$$

where λ_0 is the minimal eigenvalue of J.

Proof. We employ the optimality condition $p = \frac{1}{\alpha}(f - u)$ with $p \in \partial J(u)$ to obtain

$$J(u) = \langle p, u \rangle = \frac{1}{\alpha}\langle f - u, u \rangle \leq \frac{1}{\alpha}\|u\|\|u - f\|.$$

Moreover, due to our assumption on f and since $p \in H_0$ for every subgradient, we also conclude $u \in H_0$. Due to the definition of the minimal eigenvalue via the ground state, we conclude

$$\|u\| \leq \frac{1}{\lambda_0} J(u).$$

Inserting this relation into the above inequality and canceling $J(u)$, which is possible due to $u \notin \mathcal{N}(J)$, yields the assertion. ∎

In order to get rid of bias effects for low frequencies, several two-step approaches have been proposed in literature for examples of J. A structured approach has been derived in [11], which computes a solution v via minimizing

$$\|v - f\| \to \min_v \quad \text{subject to } J(v) - J(u) - \langle p, v - u \rangle = 0,$$

where u is the solution of (4.2) and p the corresponding subgradient arising in the optimality condition. We can elucidate this scheme in the case of $f = cu_\lambda$. If $c > \alpha\lambda$, we have a nontrivial solution $u = C(\alpha, \lambda)u_\lambda$, thus $v = cu_\lambda$ satisfies $J(v) - J(u) - \langle p, v - u \rangle = 0$ and clearly minimizes $\|v - f\|$. This means that low frequencies are exactly reconstructed by this two-step procedure. An alternative approach to reduce bias is iterative regularization, in particular the Bregman iteration, which can be interpreted as an inexact penalization of the above constraint. We will further investigate the behavior of singular vectors in this iteration in the next part.

4.2. Bregman iterations and inverse scale space flows

The Bregman iteration is obtained by subsequently computing

$$u^{k+1} \in \arg\min_u \frac{1}{2}\|u - f\|^2 + \alpha\big(J(u) - J\big(u^k\big) - \big\langle p^k, u - u^k \big\rangle\big),$$

the penalty being the Bregman distance between u and the last iterate u^k, and p^k the subgradient from the optimality condition for u^k (cf. [36]). The optimality condition directly yields an update formula for the subgradients in the form

$$p^{k+1} = p^k + \frac{1}{\alpha}(f - u^{k+1}).$$

Let us mention that for consistency with the variational method the choice $p^0 = 0$ and $u^0 \in \mathcal{N}(J)$ is usually assumed, without loss of generality we can choose $u^0 = 0$. In this case the variational method (4.2) is just the first step of the Bregman iteration. In order to obtain a suitable result, α has to be chosen large in the Bregman iteration, however.

It is instructive to investigate again the case $f = cu_\lambda$ in the Bregman iteration, looking for a solution of the form $u^k = C^k(\alpha, \lambda)u_\lambda$. If α is large, we may expect to have $c < \alpha\lambda$ and hence $u^1 = 0$, which yields $p^1 = \frac{c}{\alpha}u_\lambda$. Indeed, we obtain

$$p^k = \frac{ck}{\alpha}u_\lambda \quad \text{for } k \leq \frac{\alpha\lambda}{c}.$$

The first iteration step with a nonzero solution u^k is given by $k = K(\alpha, \lambda)$ with

$$K(\alpha, \lambda) = \min\left\{k \in \mathbb{N} \mid k > \frac{\alpha\lambda}{c}\right\}.$$

Here we can easily compute $p^k = \lambda u_\lambda$ and thus

$$u^k = cu_\lambda + \alpha\left(\frac{c(k-1)}{\alpha} - \lambda\right)u_\lambda,$$

i.e.,

$$C^k(\alpha, \lambda) = c - \big(\alpha\lambda - c(k-1)\big) = ck - \alpha\lambda.$$

For $k = K(\alpha, \lambda) + 1$, we obtain again u^k, being a nontrivial multiple of u_λ, and the corresponding subgradient is $p^k = \lambda u_\lambda = p^{k-1}$, which implies $u^k = f$. For further iterations, the result clearly does not change anymore. Hence, the number of iterations needed to obtain the exact solution behaves like $\frac{\alpha\lambda}{c}$, and we see again that eigenvectors for smaller eigenvalues (low frequency) are reconstructed faster, while eigenvectors for larger eigenvalues will appear only very late in the iteration.

The computations are a bit more precise in the limit $\alpha \to \infty$, which yields (after appropriate rescaling of step sizes) a time-continuous flow, the so-called inverse scale space method (cf. [20])

$$\partial_t p(t) = f - u(t), \quad p(t) \in \partial J\big(u(t)\big).$$

By analogous reasoning, we can compute the solution for $u(0) = p(0) = 0$ and $f = cu_\lambda$ as

$$u(t) = \begin{cases} 0, & t < \frac{\lambda}{c}, \\ cu_\lambda, & t > \frac{\lambda}{c}. \end{cases}$$

Thus, the reconstruction becomes exact at a time proportional to the eigenvalue.

4.3. Gradient flows

Another iterative scheme obtained from the variational method (4.2) is to start with $u = f$ and solve for

$$u^{k+1} \in \arg\min_u \frac{1}{2}\|u - u^k\|^2 + \alpha J(u).$$

Again, the first step is consistent with (4.2), but the dynamics is very different from the Bregman iteration, in particular for small α, which is the relevant case here. Choosing $f = cu_\lambda$, the optimality condition

$$u^{k+1} - u^k + \alpha p^{k+1} = 0, \quad p^{k+1} \in \partial J(u^{k+1})$$

yields $u^1 = (c - \alpha\lambda)u_\lambda$, and by analogous reasoning

$$u^k = (c - k\alpha\lambda)u_\lambda,$$

as long as $k < \frac{c}{\alpha\lambda}$. For $k > \frac{c}{\alpha\lambda}$, we can indeed verify that $u^k = 0$ solves the problem. Here we see that eigenvectors related to larger eigenvalues (high frequencies) shrink faster to zero, whereas the ground state is the last to disappear.

Again, this can be made more precise for the time-continuous variant, this time obtained as $\alpha \to 0$. The above iteration scheme is also known as minimizing movement scheme (cf. [23]) and the limit for appropriately scaled time is the gradient flow

$$\partial_t u(t) = -p(t), \quad p(t) \in \partial J\big(u(t)\big), \tag{4.4}$$

with initial value $u(0) = f$. Again, the solution of the gradient flow has a simple form if $f = cu_\lambda$, namely

$$u(t) = \begin{cases} (c - \lambda t)u_\lambda, & t < \frac{c}{\lambda}, \\ 0, & t > \frac{c}{\lambda}. \end{cases}$$

This means that solutions shrink to zero linearly in time, and the extinction time $\frac{c}{\lambda}$ again changes with the eigenvalue. Due to the inverse relation with λ, low frequencies get extinct later than high ones.

The behavior on eigenvectors motivates studying the gradient flow also for arbitrary initial values f. First of all, we can generalize the finite time extinction. For initial value $f \in H_0$, it is easy to show that $u(t) \in H_0$ for all $t > 0$, since for $v \in \mathcal{N}(J)$ we have

$$\big\langle u(t), v\big\rangle = \langle f, v\rangle - \int_0^t \big\langle p(s), v\big\rangle\, ds = 0.$$

Now we can use we use the standard dissipation relation

$$\big\|u(t)\big\|^2 + 2\int_0^t J\big(u(s)\big)\, ds \le \|f\|^2,$$

and $J(u(s)) \ge \lambda_0 \|u(s)\|$, resulting in

$$\big\|u(t)\big\|^2 + 2\lambda_0 \int_0^t \big\|u(s)\big\|\, ds \le \|f\|^2.$$

Similar to the proof of the Gronwall inequality, this allows deducing

$$\big\|u(t)\big\| \le \|f\| - \lambda_0 t, \quad \text{for } t < \frac{\|f\|}{\lambda_0}.$$

Thus $u(t) = 0$ for $t = \frac{\|f\|}{\lambda_0}$, and it is easy to show that for $t > \frac{\|f\|}{\lambda}_0$ the unique solution of the gradient flow is given by $u(t) = 0$ and $p(t) = 0$. Thus, the gradient flow exhibits a finite extinction phenomenon, the solution vanishes after finite time. We define the extinction time as

$$t_*(f) = \inf\big\{t > 0 \mid u(t) = 0\big\}. \tag{4.5}$$

Our analysis above yields the following upper bound on the extinction time:

Theorem 4.2. *Let $f \in H_0$ and $u \in C(0, T; H)$ be a solution of the gradient flow (4.4). Then the extinction time defined by (4.5) satisfies*

$$t_*(f) \le \frac{\|f\|}{\lambda_0}, \tag{4.6}$$

where λ_0 is the minimal nontrivial eigenvalue of J.

From the special case of f being a multiple of a ground state, we see that (4.6) is sharp for suitable initial values. In order to gain further understanding, it is instructive to investigate scalar products of the solution u with eigenvectors. This leads to

$$\langle u(t), u_\lambda \rangle = \langle f, u_\lambda \rangle - \int_0^t \langle p(s), u_\lambda \rangle \, ds \geq \langle f, u_\lambda \rangle - \lambda t.$$

Thus we obtain lower bounds on the extinction time of the form

$$t_*(f) \geq \frac{1}{\lambda} \big| \langle f, u_\lambda \rangle \big|$$

and see that the extinction time will be larger the more the initial value is correlated with low frequencies.

The extinction time is not the only relevant quantity, but also the so-called *extinction profile* is of high relevance. The extinction profile v_f is defined as

$$v_f = \lim_{\tau \downarrow 0} \frac{1}{\tau} u\big(t_*(f) - \tau\big),$$

i.e., it is the left-sided derivative of the gradient flow at the extinction time. Surprisingly, it can be shown that v_f is an eigenvector of J, under suitable conditions even that it is the ground state. This was shown first for the total variation flow (cf., e.g., [2]) and later also other zero-homogeneous evolution equations such as the fast diffusion equation (cf. [6,7]). In [15,16] this has been reconsidered in the abstract setting of eigenvectors of seminorms and general results on the extinction profile could be obtained. Let us just motivate formally why it can be expected that the extinction profile is an eigenvector. From the optimality condition in the minimizing movement scheme with τ chosen appropriately, we obtain

$$\frac{1}{\tau} u\big(t_*(f) - \tau\big) = \frac{1}{\tau} \big(u\big(t_*(f) - \tau\big) - u\big(t_*(f)\big) \big) = p\big(t_*(f)\big).$$

Hence, if $p(t_*(f))$ is not vanishing, the limit $\tau \downarrow 0$ yields $v_f = p(t_*(f))$ and, due to the homogeneity of the subdifferential, we also obtain $p(t_*(f)) \in \partial J(v_f)$ in the limit. Thus, v_f, respectively its rescaled version, is an eigenvector of J.

We finally mention that an extension of the results on the extinction profile has been carried out in [14], which analyzes the fine asymptotics for gradient flows of p-homogeneous functionals. In the case $p < 2$, there is still an extinction profile with similar properties, for $p \geq 2$ there is only decay as $t \to \infty$, however. Appropriately rescaled versions of the asymptotics of the solution are again eigenvectors of the underlying functionals.

4.4. Gradient flows and spectral decompositions

Gradient flows are particularly interesting for computing eigenvectors and even spectral decompositions, since the classical theory by Brezis (cf. [10]) implies that indeed the solution selects subgradients of minimal norm, i.e.,

$$\partial_t u(t) = -p^0(t), \quad p^0(t) = \arg\min \big\{ \|p\| \mid p \in \partial J\big(u(t)\big) \big\}.$$

Thus, we obtain a spectral decomposition into eigenvectors $p^0(t)$ with the Lebesgue measure on $\mathbb{R}_+$ if all subgradients of minimal norm are indeed eigenvectors. This means that (2.5)

needs to be satisfied for the subgradients of minimal norm in $\partial J(v)$ for all $v \in H$. Then we have indeed

$$f = \int_0^\infty p(t)\, dt,$$

but this is not yet a spectral decomposition in the above sense, since the integration is not with respect to a measure of the eigenvalue. However, as seen in [15], a change of measure from t to $\lambda(t)$ indeed yields a spectral decomposition. Note, however, that since $p(t)$ and thus $\lambda(t)$ can be piecewise constant, the arising measure in the spectral domain is not absolutely continuous with respect to the Lebesgue measure in typical cases.

In [16] an alternative way to obtain a spectral decomposition in separable Hilbert spaces was derived via extinction profiles. This countable spectral decomposition is obtain by first computing the extinction profile of the gradient flow with starting value f and then projecting f onto the space orthogonal to the first extinction profile. This projection is used again as starting value of the gradient flow and then again f is projected onto the space orthogonal to the new extinction profile. Iterating this procedure the projections converge to zero and the sum of the orthogonal components yields an atomic spectral decomposition.

There are several examples of flows that yield a spectral decomposition, the most prominent one being the one-dimensional total variation (cf. [15]). Other examples are polyhedral regularizations with sufficiently regular convex sets C (cf. [18, 19]) and one-homogeneous functionals vector fields using divergence and rotation (cf. [15]).

5. APPLICATIONS

In order to illustrate the use of nonlinear eigenvalue problems in data science, we discuss two toy examples representing wider classes of applications in this section.

5.1. 1-Laplacian graph clustering

We start with a common technique for data clustering, namely the computation of the first eigenfunction of the 1-Laplacian on graphs. For this, we acquire data on a surface by a laser scanner with random sampling, as illustrated in Figure 1. This resembles the classical two-moons data set frequently used for the evaluation of clustering methods. Based on those data points we build a nearest neighbor graph as illustrated in the right image of Figure 1.

On the arising graph we compute the first nontrivial eigenfunction of the classical graph Laplacian, which is shown in the left part of Figure 2. This serves mainly for comparison with the eigenfunction of the 1-Laplacian on the graph (the ground state of the graph total variation), which is shown in the right part. The ground state can be computed as an extinction profile of the gradient flow with the graph Laplacian eigenfunction as a starting value (cf. [16]). It is apparent that the eigenfunction of the 1-Laplacian has a much sharper transition between positive and negative values, which corresponds closely to the geometric structure in the data. This leads to improved spectral clustering as shown in Figure 3, namely the sub- and superlevel sets at zero (in red, respectively blue). One observes a rather linear

FIGURE 1
Image of traditional Austrian Christmas cookies (Vanillekipferl) and random sampling of points on the surface (left, respectively middle) and neighborhood graph built out of the sample points.

FIGURE 2
First nontrivial eigenfunction of the graph (2-)Laplacian (left) and the graph 1-Laplacian (right).

FIGURE 3
Spectral clustering based on the graph (2-)Laplacian (left) and the graph 1-Laplacian (right).

structure in the clustering with the graph Laplacian, while the clustering with the 1-Laplacian perfectly adapts to the structure in the data set.

5.2. Distance functions from ∞-Laplacians

In the following we illustrate the computation of distance functions by minimizing the ∞-Laplacian energy

$$J(u) = \operatorname{ess\,sup}_x \left|\nabla u(x)\right|. \tag{5.1}$$

We use the graph Laplacian energy on a grid graph built on the map of the United Kingdom with a large stencil. In this case we compute the (nonnegative) ground state over the set of functions on the graph vanishing on a predefined boundary (corresponding to the geographical boundary). We then normalize it such that $J(u) = 1$, which implies that u becomes the

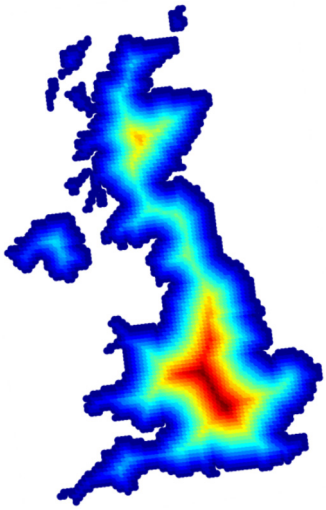

FIGURE 4
Distance function on a grid graph for the geometry of the United Kingdom.

distance function to the boundary. The result is shown in Figure 4 and generalizes results obtained by solving the eikonal equation in the continuum setting.

ACKNOWLEDGMENTS

The author thanks Daniel Tenbrinck (FAU Erlangen-Nürnberg) for measurements and computations in Section 5.1 and Leon Bungert (Bonn University) for computations in Section 5.2.

FUNDING

This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 777826 (NoMADS) and by the ERC via Grant EU FP7—ERC ConsolidatorGrant 615216 LifeInverse.

REFERENCES

[1] H. Amann, Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces. *SIAM Rev.* **18** (1976), 620–709.

[2] G. Bellettini, V. Caselles, and M. Novaga, The total variation flow in $\mathbb{R}^n$. *J. Differential Equations* **184** (2002), 475–525.

[3] M. Benning and M. Burger, Ground states and singular vectors of convex variational regularization methods. *Methods Appl. Anal.* **20** (2013), 295–334.

[4] M. Benning and M. Burger, Modern regularization methods for inverse problems. *Acta Numer.* **27** (2018), 1–111.

[5] M. Benning, M. Möller, R. Z. Nossek, M. Burger, D. Cremers, G. Gilboa, and C. Schönlieb, Nonlinear spectral image fusion. In *International conference on scale space and variational methods in computer vision*, pp. 41–53, Springer, 2017.

[6] M. Bonforte and A. Figalli, Total variation flow and sign fast diffusion in one dimension. *J. Differential Equations* **252** (2012), 4455–4480.

[7] M. Bonforte, G. Grillo, and J. Vazquez, Behaviour near extinction for the Fast Diffusion Equation on bounded domains. *J. Math. Pures Appl.* **97** (2012), 1–38.

[8] R. Bracewell, *Fourier analysis and imaging*. Springer, 2004.

[9] K. Bredies and H. Pikkarainen, Inverse problems in spaces of measures. *ESAIM Control Optim. Calc. Var.* **19** (2013), 190–218.

[10] H. Brezis, *Operateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert*. Elsevier, 1973.

[11] E. M. Brinkmann, M. Burger, J. Rasch, and C. Sutour, Bias reduction in variational regularization. *J. Math. Imaging Vision* **59** (2017), 534–566.

[12] T. Bühler and M. Hein, Spectral clustering based on the graph p-Laplacian. In *Proceedings of the 26th annual international conference on machine learning*, pp. 81–88, ACM, 2009.

[13] L. Bungert and M. Burger, Gradient flows, nonlinear power methods, and computation of nonlinear Eigenfunctions. In *Handbook of numerical analysis*, to appear.

[14] L. Bungert and M. Burger, Asymptotic profiles of nonlinear homogeneous evolution equations of gradient flow type. *J. Evol. Equ.* **20**, 1061–1092.

[15] L. Bungert, M. Burger, A. Chambolle, and M. Novaga, Nonlinear spectral decompositions by gradient flows of one-homogeneous functionals. *Anal. PDE* **14** (2021), 823–860.

[16] L. Bungert, M. Burger, and D. Tenbrinck, Computing nonlinear eigenfunctions via gradient flow extinction. In *International conference on scale space and variational methods in computer vision*, pp. 485–497, Springer, 2019.

[17] L. Bungert, Y. Korolev, and M. Burger, Structural analysis of an L^∞ variational problem and relations to distance functions. *Pure Appl. Anal.* **2** (2020), 703–738.

[18] M. Burger, L. Eckardt, G. Gilboa, and M. Moeller, Spectral representations of one-homogeneous functionals. In *International conference on scale space and variational methods in computer vision*, pp. 16–27, Springer, 2015.

[19] M. Burger, G. Gilboa, M. Moeller, L. Eckardt, and D. Cremers, Spectral decompositions using one-homogeneous functionals. *SIAM J. Imaging Sci.* **9** (2016), 1374–1408.

[20] M. Burger, G. Gilboa, S. Osher, J. Xu, and W. Yin, Nonlinear inverse scale space methods. *Commun. Math. Sci.* **4** (2006), 179–212.

[21] J. Calder, The game theoretic p-Laplacian and semi-supervised learning with few labels. *Nonlinearity* **32** (2018), 301.

[22] A. Chambolle, V. Caselles, D. Cremers, M. Novaga, and T. Pock, An introduction to total variation for image analysis. In *Theoretical foundations and numerical methods for sparse recovery*, pp. 1455–1499, De Gruyter, Berlin, 2010.

[23] E. De Giorgi, New problems on minimizing movements. In *Ennio de Giorgi: selected papers*, pp. 699–713, Springer, 1993.

[24] M. Esteban, M. Lewin, and E. Sere, Variational methods in relativistic quantum mechanics. *Bull. Amer. Math. Soc.* **45** (2008), 535–593.

[25] T. Feld, J. Aujol, G. Gilboa, and N. Papadakis, Rayleigh quotient minimization for absolutely one-homogeneous functionals. *Inverse Probl.* **35** (2019), 06400.

[26] M. Fumero, M. Moeller, and E. Rodola, Nonlinear spectral geometry processing via the TV transform. *ACM Trans. Graph.* **39** (2020), 1–16.

[27] G. Gilboa, A spectral approach to total variation. In *International conference on scale space and variational methods in computer vision*, pp. 36–47, Springer, 2013.

[28] G. Gilboa, A total variation spectral framework for scale and texture analysis. *SIAM J. Imaging Sci.* **7** (2014), 1937–1961.

[29] G. Gilboa, Iterative methods for computing Eigenvectors of nonlinear operators. In *Handbook of mathematical models and algorithms in computer vision and imaging: mathematical imaging and vision*, pp. 1–28, Springer, 2021.

[30] G. Gilboa, M. Moeller, and M. Burger, Nonlinear spectral analysis via one-homogeneous functionals: Overview and future prospects. *J. Math. Imaging Vision* **56** (2016), 300–319.

[31] M. Hein and T. Bühler, An inverse power method for nonlinear eigenproblems with applications in 1-spectral clustering and sparse PCA. In *Advances in neural information processing systems 23*, edited by J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, pp. 847–855, Curran Associates, 2010.

[32] R. Hynd and E. Lindgren, Approximation of the least Rayleigh quotient for degree-p homogeneous functionals. *J. Funct. Anal.* **272** (2012), 4873–4918.

[33] J. Jost, R. Mulas, and D. Zhang, p-Laplace operators for oriented hypergraphs. *Vietnam J. Math.* (2021), 1–36.

[34] B. Kawohl and P. Lindqvist, Positive eigenfunctions for the p-Laplace operator revisited. *Analysis* **26** (2006), 545–550.

[35] A. Le, Eigenvalue problems for the p-Laplacian. *Nonlinear Anal.* **64** (2006), 1057–1099.

[36] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, An iterative regularization method for total variation-based image restoration. *Multiscale Model. Simul.* **4** (2005), 460–489.

[37] E. Parini, An introduction to the Cheeger problem. *Surv. Math. Appl.* **6**, 9–21.

[38] P. H. Rabinowitz, Some global results for nonlinear eigenvalue problems. *J. Funct. Anal.* **7** (1971), 487–513.

[39] Y. Saad, *Numerical methods for large eigenvalue problems, revised edn*. Society for Industrial and Applied Mathematics, Philadelphia, 2011.

[40] G. Strang, *Introduction to linear algebra. 5th edn*. Wellesley-Cambridge Press, Wellesley, 1993.

[41] A. Szlam and X. Bresson, Total variation and Cheeger cuts. In *ICML'10: Proceedings of the 27th International Conference on International Conference on Machine Learning*, pp. 1039–1046, Ominpress, Haifa, 2010.

[42] U. von Luxburg, A tutorial on spectral clustering. *Stat. Comput.* **17** (2007), 395–416.

[43] M. I. Weinstein, Nonlinear Schrödinger equations and sharp interpolation estimates. *Comm. Math. Phys.* **87** (1982), 567–576.

MARTIN BURGER

Department Mathematik, Friedrich-Alexander Universität Erlangen-Nürnberg, Cauerstr. 11, 91058 Erlangen, Germany, martin.burger@fau.de