ANALYSIS & PDE

Volume 14 No. 3 2021

LEON BUNGERT, MARTIN BURGER,
ANTONIN CHAMBOLLE AND MATTEO NOVAGA

NONLINEAR SPECTRAL DECOMPOSITIONS BY GRADIENT FLOWS OF ONE-HOMOGENEOUS FUNCTIONALS

msp

ANALYSIS AND PDE
Vol. 14, No. 3, 2021
https://doi.org/10.2140/apde.2021.14.823

NONLINEAR SPECTRAL DECOMPOSITIONS BY GRADIENT FLOWS OF ONE-HOMOGENEOUS FUNCTIONALS

LEON BUNGERT, MARTIN BURGER, ANTONIN CHAMBOLLE AND MATTEO NOVAGA

This paper establishes a theory of nonlinear spectral decompositions by considering the eigenvalue problem related to an absolutely one-homogeneous functional in an infinite-dimensional Hilbert space. This approach is motivated by works for the total variation, where interesting results on the eigenvalue problem and the relation to the total variation flow have been proven previously, and by recent results on finite-dimensional polyhedral seminorms, where gradient flows can yield spectral decompositions into eigenvectors.

We provide a geometric characterization of eigenvectors via a dual unit ball and prove that they are subgradients of minimal norm. This establishes the connection to gradient flows, whose time evolution is a decomposition of the initial condition into subgradients of minimal norm. If these are eigenvectors, this implies an interesting orthogonality relation and the equivalence of the gradient flow to a variational regularization method and an inverse scale space flow. Indeed we verify that all scenarios where these equivalences were known before by other arguments — such as one-dimensional total variation, multidimensional generalizations to vector fields, or certain polyhedral seminorms — yield spectral decompositions, and we provide further examples. We also investigate extinction times and extinction profiles, which we characterize as eigenvectors in a very general setting, generalizing several results from literature.

1. Introduction

Spectral properties and spectral decompositions are at the heart of many arguments in mathematics and physics; let us just mention the spectral decomposition of self-adjoint linear operators (see [Reed and Simon 1972; Kato 1966]) and the eigenvalue problems for high-dimensional or nonlinear Schrödinger equations (see [Weinstein 1985]) as two prominent examples. In signal and image processing a variety of successful approaches were based on Fourier transforms and Laplacian eigenfunctions; in image reconstruction and inverse problems the singular-value decomposition is the fundamental tool for linear problems. Over the last two decades variational approaches and other techniques such as sparsity in anisotropic Banach spaces became popular and spectral techniques lost their dominant roles; see [Benning and Burger 2018; Burger and Osher 2013; Caselles et al. 2011; Donoho 2006; Mairal et al. 2014].

Standard examples considered in the nonlinear setting are gradient flows of the form

$$\partial_t u(t) = -p(t), \quad p(t) \in \partial J(u(t)),\ u(0) = f,$$

MSC2010: 35P10, 35P30, 47J10.
Keywords: nonlinear spectral decompositions, gradient flows, nonlinear eigenvalue problems, one-homogeneous functionals, extinction profiles.

in some Hilbert space with ∂J denoting the subdifferential of a seminorm (without Hilbertian structure in general) on a dense subspace, and variational problems of the form

$$\tfrac{1}{2}\|u-f\|^2 + tJ(u) \to \min_u,$$

with the norm in the first term being the one in the Hilbert space. As some recent results demonstrate, the role of eigenvalue problems and even spectral decompositions may be underestimated in such settings. First of all, data f satisfying the nonlinear eigenvector relation

$$\lambda f \in \partial J(f)$$

for a scalar λ give rise to analytically computable solutions for such problems, which was made precise for the TV flow (see [Andreu et al. 2002; Bellettini et al. 2002]) and is hidden in early examples for the variational problem; see [Meyer 2001; Benning and Burger 2013] for a detailed discussion. Secondly, for a general datum f the solution of the gradient flow satisfies

$$f - \bar{f} = \int_0^\infty p(t)\,\mathrm{d}t;$$

i.e., the datum is decomposed into subgradients $p(t)$ of the functional J. Here $\bar{f}$ denotes a component of the datum f which is in the null-space of J and is left-invariant under the flow, for instance the mean value of a function in the case $J = \mathrm{TV}$. In the case that these subgradients are even eigenvectors, this is called a *nonlinear spectral decomposition*.

In [Benning and Burger 2013] some further interesting properties of nonlinear eigenvectors (in a more general setting), such as their use for scale estimates and several relevant examples, have been provided. A rather surprising conjecture was made by Gilboa [2014], suggesting that TV flow and similar schemes can provide a spectral decomposition; i.e., time derivatives of the solution are related to eigenfunctions of the total variation. This was made precise in [Burger et al. 2016] in a certain finite-dimensional setting; furthermore, scenarios where a decomposition into eigenvectors can be computed were investigated. In more detail, functionals of the form $J(u) = \|Au\|_1$ with a matrix A such that AA^T is diagonally dominant lead to such spectral decompositions.

In an infinite-dimensional setting a detailed theory is widely open and will be the subject of this paper. We will consider an absolutely one-homogeneous functional J on a Hilbert space and the corresponding eigenvalue problem $\lambda u \in \partial J(u)$. Effectively this means we look at a seminorm defined on a subspace (being dense in many typical examples) of the Hilbert space and investigate the associated eigenvalue problem and gradient flow. The basic theory does not assume any relation between J and the norm of the Hilbert space, but we shall see that many favorable properties of the gradient flow — such as finite-time extinction, for instance — rely on a Poincaré-type inequality. That is, after factorizing out the null-space of the functional we have a continuous embedding of the Banach space with norm J into the ambient Hilbert space. It is thus natural to think in terms of a Gelfand triple, with subgradients of J existing a priori only in a dual space which is larger than the Hilbert space. The eigenvalue problem and the gradient flow naturally lead to considering only cases with a subgradient in the Hilbert space, which is

an abstract regularity condition known as source condition in inverse problems; see [Burger and Osher 2004; Benning and Burger 2018]. We shall see that a key role is played by the subgradient of minimal norm (known as minimal norm certificate in compressed sensing and related problems; see [Fornasier and Rauhut 2011; Duval and Peyré 2015; Chambolle et al. 2017]. A first key contribution of this paper is a geometric characterization of eigenvectors in such a setting, which is based on a dual characterization of absolutely one-homogeneous convex functionals. Roughly speaking we can interpret all possible subgradients as elements lying on the boundary of a dual unit ball (the subdifferential of J at 0) and single out eigenvectors as those elements which are a normal vector to an orthogonal supporting hyperplane of the ball. Thus, the eigenvalue problem becomes a geometric problem for a Banach space relative to a Hilbert space structure.

We also show that being a subgradient of minimal norm is a necessary condition for eigenvectors. This establishes an interesting connection to gradient flows, which automatically select subgradients of minimal norm as the time derivative of the primal variable. We hence study gradient flows in further detail and conclude that — if the above-noted geometric condition is satisfied — they yield a spectral decomposition, i.e., a representation of the initial data f as an integral of eigenvectors with decreasing frequency (decreasing Hilbert space norm at fixed dual norm). Moreover, we show that if the gradient flow yields a spectral decomposition, this is already sufficient to obtain equivalence to the variational method as well as an inverse scale space method proposed as an alternative to obtain spectral decompositions; see [Burger et al. 2015]. With an appropriate reparametrization from the time in the gradient flow to a spectral dimension we rigorously obtain a spectral decomposition akin to the spectral decomposition of self-adjoint linear operators in Hilbert space as discussed in [Burger et al. 2016]. We apply our theory to several examples: in particular, it matches the finite-dimensional theory for polyhedral regularizations in [Burger et al. 2016], and it can also be used for the one-dimensional total variation flow, a flow of a divergence functional for vector fields, as well as for a combination of divergence and rotation sparsity. Moreover, we visit the simple case of a flow of the L^1-norm, which gives further intuition and limitations in a case where no Poincaré-type estimate between the convex functional and the Hilbert space norm is valid.

Finally, we also discuss the extinction times and extinction profiles of gradient flows, a problem that was studied for TV flow in detail before; see [Andreu et al. 2002; Giga and Kohn 2011; Bonforte and Figalli 2012]. Our theory is general enough to allow for a direct generalization of the results to flows of absolutely one-homogeneous convex functionals and simplifies many proofs. In particular, we can show that under generic conditions the gradient flows have finite extinction time and there is an extinction profile, i.e., a left limit of the time derivative at the extinction, which is an eigenvector. Furthermore, we give sharp upper and lower bounds of the extinction time. For flows that yield a spectral decomposition, we obtain a simple relation between the initial datum, the extinction time, and the extinction profile. In the case of the one-dimensional total variation flow we get the results in [Bonforte and Figalli 2012] as special cases.

The remainder of the paper is organized as follows: In Section 2 we discuss some basic properties of absolutely one-homogeneous functionals and the related nonlinear eigenvalue problem. Section 3 is devoted to obtain further geometric characterizations of eigenvectors and to work out connections to

subgradients of minimal norm. In Section 4 we discuss the potential to obtain spectral decompositions by gradient flows. For this sake we give an overview of the classical theory by Brézis, which shows that gradient flows generate subgradients of minimal norm, and we also provide equivalence results to other methods in the case of spectral decompositions, for which we give a geometric condition. Moreover, we discuss the appropriate scaling of the spectrum from time to eigenvalues in order to obtain a more suitable decomposition. In Section 5 we show that the geometric condition for obtaining a spectral decomposition is indeed satisfied for relevant examples such as total variation flow in one dimension and multidimensional flows of vector fields with functionals based on divergence and rotation. Finally, in Section 6 we investigate the extinction profiles of the gradient flows, which we show to be eigenvectors even if the flow itself does not necessarily produce a spectral decomposition.

2. Absolutely homogeneous convex functionals and eigenvalue problems

In the following we collect some results about the basic class of convex functionals we consider in this paper; moreover we provide basic definitions and results about the nonlinear eigenvalue problem related to such functionals.

2A. ***Notation.*** Let $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ be a real Hilbert space with induced norm $\|\cdot\| := \sqrt{\langle\cdot,\cdot\rangle}$ and J be a functional in the class $\mathcal{C}$, defined as follows:

Definition 2.1. The class $\mathcal{C}$ consists of all maps $J : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ such that J is convex, lower semicontinuous with respect to the strong topology on $\mathcal{H}$, and absolutely one-homogeneous, i.e.,

$$J(cu) = |c| J(u) \quad \text{for all } u \in \mathcal{H},\ c \in \mathbb{R} \setminus \{0\}, \tag{2-1}$$

$$J(0) = 0. \tag{2-2}$$

The effective domain and null-space of $J \in \mathcal{C}$ are given by

$$\operatorname{dom}(J) := \{u \in \mathcal{H} : J(u) < \infty\}, \tag{2-3}$$

$$\mathcal{N}(J) := \{u \in \mathcal{H} : J(u) = 0\}. \tag{2-4}$$

Note that $\operatorname{dom}(J)$ and $\mathcal{N}(J)$ are not empty due to (2-2) and that $\mathcal{N}(J)$ is a (strongly and weakly) closed linear space [Bungert and Burger 2019] whose orthogonal complement we denote by $\mathcal{N}(J)^\perp$. The effective domain $\operatorname{dom}(J)$ is also a linear space but not closed, in general. Given any $f \in \mathcal{H}$, the orthogonal projection $\bar{f}$ of f onto $\mathcal{N}(J)$ is

$$\bar{f} := \arg\min_{u \in \mathcal{N}(J)} \|u - f\|, \tag{2-5}$$

and the "dual norm" of $f \in \mathcal{H}$ with respect to J is defined as

$$\|f\|_* := \sup_{u \in \mathcal{N}(J)^\perp} \frac{\langle f, u\rangle}{J(u)}. \tag{2-6}$$

Considering J as a norm on the Banach space

$$\mathcal{V} = \operatorname{dom}(J) \cap \mathcal{N}(J)^\perp, \tag{2-7}$$

the above dual norm is indeed defined on the dual space of $\mathcal{V}$ (respectively a predual if it exists). Hence, we naturally obtain a Gelfand triple structure $\mathcal{V} \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{V}^*$ and the eigenvalue problem can also be understood as a relation between the geometries of the Hilbert space $\mathcal{H}$ and the Banach spaces $\mathcal{V}$ and $\mathcal{V}^*$.

By $\partial J(u)$ we denote the subdifferential of J in $u \in \operatorname{dom}(J)$, given by

$$\partial J(u) := \{p \in \mathcal{H} : \langle p, v\rangle \le J(v) \text{ for all } v \in \mathcal{H},\ \langle p, u\rangle = J(u)\}, \tag{2-8}$$

and define $\operatorname{dom}(\partial J) := \{u \in \mathcal{H} : \partial J(u) \neq \varnothing\}$. Of particular importance will be the subdifferential in zero

$$\partial J(0) = \{p \in \mathcal{H} : \langle p, v\rangle \le J(v) \text{ for all } v \in \mathcal{H}\}, \tag{2-9}$$

which uniquely determines J as we will see. Using the definition of the dual norm (2-6), it can be easily seen that

$$\partial J(0) = \{p \in \mathcal{N}(J)^\perp : \|p\|_* \le 1\}; \tag{2-10}$$

i.e., roughly speaking the subdifferential of J in 0 coincides with the dual unit ball in the space $\mathcal{V}^*$.

Lastly, we recall the definitions of the *Fenchel conjugate* of a general function $\Phi : \mathcal{H} \to [-\infty, +\infty]$ as $\Phi^*(u) := \sup_{v\in\mathcal{H}} \langle v, u\rangle - \Phi(v)$ and of the *indicator function* of a subset $K \subset \mathcal{H}$ as

$$\chi_K(u) := \begin{cases} 0, & u \in K, \\ +\infty, & u \notin K. \end{cases}$$

We refer to [Bauschke and Combettes 2011] for fundamental properties.

2B. ***Absolutely one-homogeneous functionals.*** In the following, we collect some elementary properties of functionals in the class $\mathcal{C}$ defined in Definition 2.1 and their subdifferentials whose proofs are either trivial or can be found in [Burger et al. 2016].

Proposition 2.2. *Let $J \in \mathcal{C}$, $u, v \in \mathcal{H}$, and $c > 0$. It holds*:

(1) $J(u) \ge 0$.

(2) $J(u+v) \le J(u) + J(v)$ *and* $J(u) - J(v) \le J(u-v)$.

(3) $J(u+w) = J(u)$ *for all* $w \in \mathcal{N}(J)$.

(4) $J^*(u) = \chi_{\partial J(0)}(u)$.

(5) $\partial J(u)$ *is convex and weakly closed and it holds* $\partial J(u) \subset \partial J(0)$.

(6) $\partial J(cu) = \partial J(u)$.

(7) $\partial J(u) \perp \mathcal{N}(J)$; *i.e.*, $\langle p, w\rangle = 0$ *for all* $p \in \partial J(u)$, $w \in \mathcal{N}(J)$.

As already indicated, the knowledge of the set $\partial J(0)$ suffices to uniquely determine a functional $J \in \mathcal{C}$. Furthermore, any such functional has a canonical dual representation, similar to the concept of dual norms. This is no surprise since the elements of $\mathcal{C}$ are seminorms on subspaces of $\mathcal{H}$ and norms if and only if they have trivial null-space.

Theorem 2.3. *A functional $J : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ belongs to $\mathcal{C}$ if and only if*

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

for a set $K \subset \mathcal{H}$ that satisfies $K = -K$. In this case, it holds

$$\partial J(0) = \overline{\mathrm{conv}}(K),$$

where $\overline{\mathrm{conv}}$ denotes the closed convex hull of a set.

Proof. It is well known that the Fenchel conjugate of an absolutely one-homogeneous functional J is given by $J^*(u) = \chi_{\partial J(0)}$. Hence, for the first implication we observe that by choosing $K := \partial J(0)$ and using that — being lower semicontinuous and convex — J equals its double Fenchel conjugate, it holds $J(u) = J^{**}(u) = \chi_K^*(u) = \sup_{p \in K} \langle p, u \rangle$. Furthermore, K satisfies $K = -K$, which can be seen from $\langle -p, u \rangle = \langle p, -u \rangle \le J(-u) = J(u)$ for $p \in K$.

Conversely, any J given by $J(u) = \sup_{p \in K} \langle p, u \rangle = \chi_K^*(u)$, where $K = -K$, trivially belongs to $\mathcal{C}$, and it holds $J^* = \chi_K^{**} = \chi_{\overline{\mathrm{conv}}(K)}$. Hence, by standard subdifferential calculus one concludes $\partial J(0) = \{p \in \mathcal{H} : J^*(p) = 0\} = \overline{\mathrm{conv}}(K)$. □

Remark 2.4. Using the convexity and homogeneity of J, Jensen's inequality immediately implies that the generalized triangle inequality

$$J\left(\int_a^b u(t)\,\mathrm{d}t\right) \le \int_a^b J(u(t))\,\mathrm{d}t \tag{2-11}$$

for a function $u : [a, b] \to \mathcal{H}$ holds, whenever these expressions make sense.

Due to Theorem 2.3, we will from now on assume that

$$J(u) = \sup_{p \in K} \langle p, u \rangle, \quad u \in \mathcal{H},\ K := \partial J(0), \tag{2-12}$$

after possibly replacing K by the closure of its convex hull.

2C. *Subgradients and eigenvectors.* In this section, we will define the nonlinear eigenvalue problem under consideration and provide first insights into the geometric connection of eigenvectors and the dual unit ball K. We start with general subgradients of a functional $J \in \mathcal{C}$ before we turn to the special case of subgradients which are eigenvectors.

Definition 2.5 (subgradients). Let $J \in \mathcal{C}$ and $u \in \mathrm{dom}(\partial J)$. Then the elements of the set $\partial J(u)$ are called subgradients of J in u.

Proposition 2.6. *Let p be a subgradient of J in $u \neq 0$ and let K be as in (2-12). Then p lies in the relative boundary $\partial_{\mathrm{rel}} K$ of K, where $\partial_{\mathrm{rel}} K = K \setminus \mathrm{relint}(K)$ and $\mathrm{relint}(K) = \{p \in K :$ there exists $c > 1$ such that $cp \in K\}$ denotes the relative interior of K.*

Proof. We observe that for all $c \in [0, 1]$ and $v \in \mathcal{H}$ it holds

$$\langle cp, v \rangle = c\langle p, v \rangle \le cJ(v) \le J(v);$$

i.e., $cp \in K$. Let us assume that there is $c > 1$ such that $cp \in K$. Then, using that $p \in \partial J(u)$ yields

$$c\langle p, u\rangle = \langle cp, u\rangle \leq J(u) = \langle p, u\rangle,$$

which clearly contradicts $c > 1$. Hence, we have established the claim. □

Interestingly, due to the fact that subdifferentials $\partial J(u)$ are convex sets and lie in the (relative) boundary of the convex set K, they are either singletons or lie in a "flat" region of the boundary of K.

Definition 2.7 (eigenvectors). Let $J \in \mathcal{C}$. We say that $u \in \mathcal{H}$ is an eigenvector of ∂J with eigenvalue $\lambda \in \mathbb{R}$ if

$$\lambda u \in \partial J(u). \tag{2-13}$$

Remark 2.8. Due to the positive zero-homogeneity of $\partial J(u)$, any multiple cu of u with $c > 0$ is also an eigenvector of ∂J with eigenvalue λ/c. To avoid this ambiguity one sometimes additionally demands $\|u\| = 1$ from an eigenvector u. In this work, however, we do not adopt this normalization since it does not match the flows that we are considering. As a consequence, all occurring eigenvalues should be multiplied by the norm of the eigenvector to become interpretable. For example, let $p \in \partial J(p)$. Then $q := p/\|p\|$ has unit norm and is an eigenvector of ∂J with eigenvalue $\lambda := \|p\|$ since $\lambda q = p \in \partial J(p) = \partial J(q)$. The last equality follows from Proposition 2.2(6).

Now we collect some basic properties of eigenvectors and, in particular, we show that eigenvectors are the elements of minimal norm in their respective subdifferential. Hence, one can restrict the search for eigenvectors to the subgradients of minimal norm; this shows a first connection to gradient flows which select the subgradients of minimal norm, as already indicated.

Proposition 2.9 (properties of eigenvectors). *Let $u \in \mathcal{H}$ be an eigenvector of ∂J with eigenvalue $\lambda \in \mathbb{R}$. Then it holds*:

(1) *$-u$ is eigenvector with eigenvalue λ.*

(2) *$\lambda \geq 0$ and $\lambda = 0$ if and only if $u = 0$.*

(3) $\lambda u = \arg\min\{\|p\| : p \in \partial J(u)\}$.

Proof. We only prove the third item. It holds for all $p \in \partial J(u)$ that

$$\lambda\|u\|^2 = \langle \lambda u, u\rangle = J(u) = \langle p, u\rangle \leq \|p\|\|u\|$$

and, since $\lambda \geq 0$, we obtain $\|\lambda u\| \leq \|p\|$. The convexity of $\partial J(u)$ shows that λu is the unique element of minimal norm. □

Remark 2.10. In the following, we will simply talk about eigenvectors whilst suppressing the dependency upon ∂J, for brevity.

It is trivial that all elements in the null-space of J are eigenvectors with eigenvalue 0. However, these eigenvectors are only of minor interest, as the example of total variation shows, where the null-space consists of all constant functions. Hence, in [Benning and Burger 2013] so-called ground states were considered. These are eigenvectors in the orthogonal complement of the null-space with the lowest

possible positive eigenvalue and, hence, the second-largest eigenvalue of the operator ∂J. In our setting, these ground states correspond to vectors with minimal norm in the relative boundary $\partial_{\mathrm{rel}} K$ of K, which was defined in Proposition 2.6.

Proposition 2.11. *Let u_0 be a ground state of J, defined as*

$$u_0 \in \arg\min_{\substack{u \in \mathcal{N}(J)^\perp \\ \|u\|=1}} J(u), \tag{2-14}$$

and let $\lambda_0 := J(u_0)$. Then $p_0 := \lambda_0 u_0$ is an element of minimal norm in $\partial_{\mathrm{rel}} K$ and an eigenvector.

Proof. It has been shown in [Benning and Burger 2013] that ground states u_0 in (2-14) exist under relatively weak assumptions[1] and are eigenvectors with eigenvalue λ_0. Furthermore, λ_0 is by definition the smallest eigenvalue that a normalized eigenvector in $\mathcal{N}(J)^\perp$ can have. Hence, $p_0 := \lambda_0 u_0 \in \partial J(u_0) = \partial J(p_0)$, which shows that p_0 is an eigenvector. Let us assume there is $q \in \partial_{\mathrm{rel}} K$ such that $\|q\| < \|p_0\| = \lambda_0$. This implies, by the Cauchy–Schwarz inequality and the definition of λ_0,

$$\langle q, u\rangle \le \|q\|\|u\| < \lambda_0 \|u\| = J(u) \quad \text{for all } u \in \mathcal{N}(J)^\perp.$$

Since this inequality is strict, we can conclude that $q \notin \partial_{\mathrm{rel}} K$. Therefore, $\|p_0\|$ is minimal, as claimed. □

3. Geometric characterization of eigenvectors

In this section, we give a novel geometric characterizing of eigenvectors. Simply speaking, eigenvectors p (with eigenvalue 1) are exactly those vectors on the relative boundary of K for which there exists a supporting hyperplane of K through p that is orthogonal to p. In other words, there is a multiple of the Hilbert unit ball which is tangential to $\partial_{\mathrm{rel}} K$ at p. All other eigenvectors are multiples of these. In particular, the eigenvalue problem can be viewed as studying the *relative geometry* of the unit balls in $\mathcal{H}$ and $\mathcal{V}^*$, respectively, where $\mathcal{V}$ is given by (2-7). Since this geometric interpretation is not very handy in the case of infinite-dimensional Hilbert spaces (e.g., function spaces), we will only work with an algebraic characterization in the following. We start with a lemma that allows us to limit ourselves to the study of eigenvectors with eigenvalue 1 without loss of generality.

Lemma 3.1. *We have $u \in \mathcal{H}$ is an eigenvector with eigenvalue $\lambda > 0$ if and only if $p := \lambda u$ is an eigenvector with eigenvalue* 1.

Proof. The statement follows directly from Proposition 2.2(6). □

A key geometric characterization is provided by the following result:

Proposition 3.2. *We have $p \in K$ is an eigenvector with eigenvalue* 1 *if and only if*

$$\langle p, p - q\rangle \ge 0 \quad \textit{for all } q \in K. \tag{3-1}$$

[1] e.g., if J satisfies (6-3)

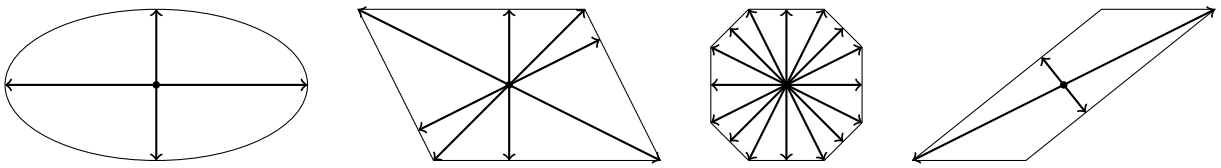

Figure 1. Four different dual unit balls K with all eigenvectors with eigenvalue 1.

Proof. It holds that p is an eigenvector with eigenvalue 1 if and only if

$$\langle p, p\rangle = J(p) = \sup_{q\in K}\langle p, q\rangle.$$

This is equivalent to (3-1), which concludes the proof. □

The statement of Proposition 3.2 is illustrated in Figure 1, which shows four different dual unit balls $K \subset \mathbb{R}^2$ together with all eigenvectors with eigenvalue 1. It is obvious from the picture that any other vector on the boundary does not meet the boundary orthogonally.

Proposition 3.2 can be used to obtain the following results:

Corollary 3.3. *Let p be a point of maximal norm in K; i.e., $\|p\| \geq \|q\|$ for all $q \in K$. Then every positive multiple of p is an eigenvector.*

Example 3.4. Consider the linear eigenvalue problem for a symmetric and positively semidefinite matrix $A \in \mathbb{R}^{n\times n}$. The corresponding functional is given by $J(u) = \sqrt{\langle u, Au\rangle}$ and K is an ellipsoid. Along the main axes of the ellipsoid the hypersurface $p + \operatorname{span}\{p^\perp\}$ is tangential; hence the main axes define the eigenvectors.

Remark 3.5 (existence of nonlinear spectral decompositions). Unlike in the above-noted linear case where there are exactly n different eigendirections, nonlinear eigenvectors in our setting may constitute an overcomplete generating set of the ambient space, as can be seen in Figure 1. The leftmost set corresponds to the linear case and has two different eigendirections. In contrast, in the nonlinear cases one can have significantly more different eigendirections, which makes the set of eigenfunctions an overcomplete system in these examples. Whether this is true in general remains an open question. However, in Section 5 we give several relevant examples where there are sufficiently many nonlinear eigenvectors to represent any datum as linear combination of such vectors. Note that from the second and fourth sets in Figure 1, it also becomes clear that in the nonlinear case one cannot expect to have orthogonal eigenvectors, as is the case for compact self-adjoint linear operators.

Next we investigate for which elements $q \in K$ the characterizing inequality (3-1) for eigenvectors is actually an equality.

Proposition 3.6. *Let $u \in \operatorname{dom}(\partial J)$ and $p := \arg\min\{\|q\| : q \in \partial J(u)\}$ be an eigenvector with eigenvalue 1. Then it holds*

$$\langle p, p - q\rangle = 0 \quad \textit{for all } q \in \partial J(u), \tag{3-2}$$

which can be reformulated as $\partial J(u) \subset \partial J(p)$.

Proof. The nonnegativity of the left-hand side follows directly from the assumption that p is an eigenvector and thus fulfills (3-1). For the other inequality, we let $q \in \partial J(u)$ arbitrary and consider $r := \lambda q + (1-\lambda)p$ for $\lambda \in (0, 1)$, which is in $\partial J(u)$, as well, due to convexity. Using (3-1) and the minimality of $\|p\|$ yields

$$\lambda\langle p, q\rangle + (1-\lambda)\|p\|^2 = \langle p, r\rangle \le \|p\|^2 \le \|r\|^2 = \lambda^2\|q\|^2 + (1-\lambda)^2\|p\|^2 + 2\lambda(1-\lambda)\langle p, q\rangle.$$

Dividing this by $\lambda(1-\lambda)$ one finds

$$\frac{1}{1-\lambda}\langle p, q\rangle + \frac{1}{\lambda}\|p\|^2 \le \frac{\lambda}{1-\lambda}\|q\|^2 + \frac{1-\lambda}{\lambda}\|p\|^2 + 2\langle p, q\rangle,$$

which can be simplified to

$$\frac{1}{1-\lambda}\langle p, q\rangle \le \frac{\lambda}{1-\lambda}\|q^2\| - \|p\|^2 + 2\langle p, q\rangle.$$

Letting λ tend to zero and reordering shows $\langle p, p-q\rangle \le 0$; hence equality holds. □

The converse statement of Proposition 3.6 is false in general. This can be seen by choosing $K \subset \mathbb{R}^n$ to be an ellipsoid. In this case all subdifferentials $\partial J(u)$ are singletons since the boundary of an ellipsoid does not contain convex sets consisting of two or more points. Hence, (3-2) is always met but not every boundary point has an orthogonal tangent hyperplane. However, the converse is true in finite dimensions if $K \subset \mathbb{R}^n$ is a polyhedron. In [Burger et al. 2016] the authors introduced condition (3-2), which they termed (MINSUB), to study the case of polyhedral functionals J, meaning that the set K is a polyhedron. Using (MINSUB) together with another, relatively strong, condition they were able to prove the converse of Proposition 3.6, namely that all subgradients of minimal norm are eigenvectors. Below we show that in fact the other condition is superfluous.

Proposition 3.7. *Let $K \subset \mathbb{R}^n$ be a convex polyhedron such that for all $u \in \mathbb{R}^n$ the element $p :=$ $\arg\min\{\|q\| : q \in \partial J(u)\}$ satisfies condition*

$$\langle p, p-q\rangle = 0 \quad \textit{for all } q \in \partial J(u) \tag{MINSUB}$$

from [Burger et al. 2016]. *Then p is an eigenvector.*

Proof. Let us fix $u \in \mathbb{R}^n$ and let p be the element of minimal norm in $\partial J(u)$. By the definition of the subdifferential and by Proposition 2.6 we infer that $\partial J(u)$ — being the intersection of K and the hypersurface $\{q \in \mathbb{R}^n : \langle q, u\rangle = J(u)\}$ — must coincide with a facet F of the polyhedron. Due to (MINSUB), the set $\mathcal{S} := \{q \in \mathbb{R}^n : \langle q, p\rangle = \|p\|^2\}$ defines a hypersurface through p such that $F \subset \mathcal{S}$ and $\mathcal{S}$ is orthogonal to p. Since K is convex, all other points in K lie on one side of $\mathcal{S}$ which implies that $\mathcal{S}$ is supporting K, and hence $\langle p, p-q\rangle \ge 0$ for all $q \in K$. With Proposition 3.2 we conclude that p is an eigenvector with eigenvalue 1. □

Remark 3.8. Notably, the minimality of $\|p\|$ does not play a role in the proof of Proposition 3.7. However, from the Cauchy–Schwarz inequality it follows that only subgradients of minimal norm can satisfy (MINSUB).

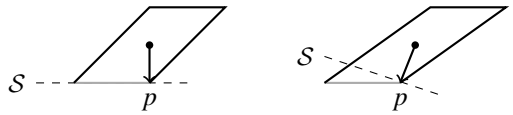

Figure 2. Left: (MINSUB) is met since $\mathcal{S}$ is supporting. Right: Here (MINSUB) is violated; i.e., $\mathcal{S}$ is not supporting and p is no eigenvector.

Figure 2 shows two polyhedrons together with a subgradient of minimal norm in the subdifferential marked in gray. The left polyhedron satisfies (MINSUB) but the right one does not, because it is too distorted. Consequently, for the left polyhedron the subgradient of minimal norm is an eigenvector whereas this is not true for the right one.

4. Spectral decompositions by gradient flows

The fact that eigenvectors are subgradients of minimal norms motivates us to further study processes that generate such subgradients. Indeed, the theory of maximal monotone evolution equations shows that gradient flows have this desirable property.

4A. ***Gradient flow theory from maximal monotone evolution equations.*** In this section we give a concise recap of the theory of nonlinear evolution equations due to [Brézis 1973]; see also [Kōmura 1967] for an earlier existence result due to Kōmura. The theory deals with the solution of the differential inclusion

$$\begin{cases} \partial_t u(t) + Au(t) \ni 0, \\ u(0) = f, \end{cases} \tag{4-1}$$

for times $t > 0$. Here A denotes a potentially *nonlinear* and *multivalued* operator defined on a subset $\operatorname{dom}(A) := \{u \in \mathcal{H} : Au \neq \varnothing\}$ and is assumed to be *maximal monotone*. That is,

$$\langle p - q, u - v\rangle \geq 0 \quad \text{for all } p \in Au,\ q \in Av, \tag{4-2}$$

and A cannot be extended to a monotone operator with larger domain; see [Brézis 1973] for a precise definition. Furthermore, one defines

$$A^0 u := \arg\min\{\|p\| : p \in Au\}, \quad u \in \operatorname{dom}(A), \tag{4-3}$$

which is the norm-minimal element in the convex set Au.

For these class of operators one has the following:

Theorem 4.1 [Brézis 1973]. *For all $f \in \operatorname{dom}(A)$ there exists a unique function $u : [0, \infty) \to \mathcal{H}$ such that*:

(1) *$u(t) \in \operatorname{dom}(A)$ for all $t > 0$.*

(2) *u is Lipschitz continuous on $[0, \infty)$; i.e., $\partial_t u \in L^\infty(0, \infty; \mathcal{H})$ (as distributional derivative) and it holds*

$$\|\partial_t u\|_{L^\infty(0,\infty;\mathcal{H})} \leq \|A^0 f\|. \tag{4-4}$$

(3) *(4-1) holds for almost every $t \in (0, \infty)$.*

(4) *u is right-differentiable for* all $t \in (0, \infty)$ *and it holds*

$$\partial_t^+ u(t) + A^0 u(t) = 0 \quad \textit{for all } t \in (0, \infty). \tag{4-5}$$

(5) *The function $t \mapsto A^0 u(t)$ is right-continuous and the function $t \mapsto \|A^0 u(t)\|$ is nonincreasing.*

Proof. For the proof see [Brézis 1973, Theorem 3.1]. □

An important instance of maximally monotone operators are subdifferentials ∂J of lower semicontinuous convex functionals $J : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$. For these one can relax the assumption $f \in \operatorname{dom}(\partial J)$ to $f \in \overline{\operatorname{dom}(\partial J)}$ and obtains:

Theorem 4.2 [Brézis 1973]. *Let $A = \partial J$, where $J : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous, convex, and proper, and let $f \in \overline{\operatorname{dom}(A)}$. Then there exists a unique continuous function $u : [0, \infty) \to \mathcal{H}$ with $u(0) = f$ such that:*

(1) *$u(t) \in \operatorname{dom}(A)$ for all $t > 0$.*

(2) *u is Lipschitz continuous on $[\delta, \infty)$ for all $\delta > 0$ and it holds*

$$\|\partial_t u\|_{L^\infty(\delta,\infty;\mathcal{H})} \le \|A^0 v\| + \frac{1}{\delta}\|f - v\| \quad \textit{for all } v \in \operatorname{dom}(A), \textit{ for all } \delta > 0. \tag{4-6}$$

(3) *u is right-differentiable for* all $t \in (0, \infty)$ *and it holds*

$$\partial_t^+ u(t) + A^0 u(t) = 0 \quad \textit{for all } t \in (0, \infty). \tag{4-7}$$

(4) *The function $t \mapsto A^0 u(t)$ is right-continuous for all $t > 0$ and the function $t \mapsto \|A^0 u(t)\|$ is nonincreasing.*

(5) *The function $t \mapsto J(u(t))$ is convex, nonincreasing, Lipschitz continuous on $[\delta, \infty)$ for all $\delta > 0$ and it holds*

$$\frac{\mathrm{d}^+}{\mathrm{d}t} J(u(t)) = -\|\partial_t^+ u(t)\|^2 \quad \textit{for all } t > 0. \tag{4-8}$$

Proof. For the proof see [Brézis 1973, Theorem 3.2], where it should be noted that right-differentiability of the map $t \mapsto J(u(t))$ follows since it is Lipschitz continuous and nonincreasing. □

Applying Theorem 4.2 to the so-called *gradient flow* of the functional J

$$\begin{cases} \partial_t u(t) = -p(t), \\ p \in \partial J(u(t)), \\ u(0) = f, \end{cases} \tag{GF}$$

yields the existence of a unique solution $u : [0, \infty) \to \mathcal{H}$ with associated subgradients $p(t) := -\partial_t^+ u(t) \in \partial J(u(t))$, which have minimal norm in $\partial J(u(t))$ for all $t > 0$.

Remark 4.3. From now on, we will denote all occurring right derivatives with the usual derivative symbols ∂_t and $\mathrm{d}/\mathrm{d}t$ to simplify our notation.

Having the geometric characterization of eigenvectors and the existence theory of gradient flows at hand, we are now interested in the scenario that the gradient flow yields a sequence of subgradients $p(t)$ which are eigenvectors, i.e., $p(t) \in \partial J(p(t))$. Reformulating this using the eigenvector characterization from Proposition 3.2 we obtain:

Theorem 4.4. *The gradient flow (GF) yields a sequence of eigenvectors $p(t)$ for $t > 0$ if and only if $\langle p(t), p(t) - q\rangle \geq 0$ for all $q \in K$, for all $t > 0$.*

Before giving examples of functionals $J \in \mathcal{C}$ that guarantee this to happen, we investigate the consequences of such a behavior of the flow. We prove that, in this case, the gradient flow is equivalent to a variational regularization method and an inverse scale space flow. Furthermore, disjoint increments $p(t) - p(s)$ of eigenvectors will turn out to be mutually orthogonal. Finally, we will use the subgradients $p(t)$ of an arbitrary gradient flow to define a measure that acts as generalization of the spectral measure corresponding to a self-adjoint/compact linear operator in the case that $p(t)$ are eigenvectors.

4B. *Equivalence of gradient flow, variational method, and inverse scale space flow.* First we show that if the gradient flow (GF) generates eigenvectors, this implies the equivalence with a variational regularization problem (VP), and the inverse scale space flow (ISS), given by

$$\begin{cases} v(t) = \arg\min_{v\in\mathcal{H}} E_t(v), \\ E_t(v) = \frac{1}{2}\|v - f\|^2 + tJ(v), \end{cases} \tag{VP}$$

$$\begin{cases} \partial_\tau r(\tau) = f - w(\tau), \\ r(\tau) \in \partial J(w(\tau)), \\ r(0) = 0, \quad w(0) = \bar{f}. \end{cases} \tag{ISS}$$

Here $t > 0$ denotes a time/regularization parameter, whereas $\tau > 0$ will turn out to correspond to the "inverse time" $1/t$. Furthermore, the initial condition $w(0) = \bar{f}$ of the inverse scale space flow denotes the orthogonal projection of f onto the null-space of J defined in (2-5); see [Bungert and Burger 2019] for more details. Note that all time derivatives ought to be understood in the weak sense, existent for almost all times, or in the sense of right derivatives that exist for all times.

Theorem 4.5 (equivalence of GF and VP). *Let (u, p) be a solution of the gradient flow (GF) and assume that for all $t > 0$ it holds $p(t) \in \partial J(p(t))$. Then for $s \leq t$ it holds $p(s) \in \partial J(u(t))$. Moreover, $u(t) = v(t)$, where $v(t)$ solves (VP).*

Proof. From (3-1) we see that in particular $\langle p(t), p(t) - p(s)\rangle \geq 0$ holds for all $0 < s \leq t$ and hence, using $p(t) = -\partial_t u(t)$ together with Theorem 4.2(5),

$$0 \geq \langle -p(t), p(t) - p(s)\rangle = \frac{\mathrm{d}}{\mathrm{d}t} J(u(t)) - \langle \partial_t u(t), p(s)\rangle.$$

Integrating from s to t yields

$$J(u(t)) - J(u(s)) - \langle u(t), p(s)\rangle + J(u(s)) \leq 0,$$

which is equivalent to $J(u(t)) \leq \langle u(t), p(s)\rangle$ and hence $p(s) \in \partial J(u(t))$. That $v(t) := u(t)$ solves (VP) follows by observing that the Fejér mean $q(t) := (1/t)\int_0^t p(s)\,\mathrm{d}s$ is the appropriate subgradient for the

optimality condition of $v(t)$ being a minimizer of E_t; i.e.,

$$v(t) - f + tq(t) = 0, \quad q(t) \in \partial J(v(t)).$$

The fact that (GF) and (VP) possess unique solutions concludes the proof. □

Now we prove the equivalence of the variational problems and the inverse scale space flow. To avoid confusion due to the time reparametrization connecting t and τ, we denote the derivatives of v and q with respect to the regularization parameter t in (VP) by v' and q', respectively. For instance, $v'(1/\tau)$ simply means $(\partial_t v(t))|_{t=1/\tau}$ and this expression exists since by the previous theorem $v = u$ and u is right-differentiable for all $t > 0$.

Theorem 4.6 (equivalence of VP and ISS). *Let the gradient flow (GF) generate eigenvectors $p(t) \in \partial J(p(t))$. Let, furthermore, $\{v(t) : t > 0\}$ be the solutions of the variational problem (VP) with subgradients $\{q(t) : t > 0\}$. Then for $\tau := 1/t$ the pair (w, r), given by*

$$\begin{aligned} w(\tau) &:= v\Big(\frac{1}{\tau}\Big) - \frac{1}{\tau} v'\Big(\frac{1}{\tau}\Big), \\ r(\tau) &:= q\Big(\frac{1}{\tau}\Big), \end{aligned} \tag{4-9}$$

is a solution of the inverse scale space flow (ISS).

Proof. From the optimality conditions of (VP) for $t > 0$ we deduce $q(t) = [f - v(t)]/t$. By the quotient rule we obtain $q'(t) = [v(t) - f - tv'(t)]/t^2$. Inserting $t = 1/\tau$ yields $q'(1/\tau) = \tau^2[v(1/\tau) - f - (1/\tau)v'(1/\tau)]$. Using this we find with the chain rule

$$\partial_\tau^+ r(\tau) = -\frac{1}{\tau^2} q'\Big(\frac{1}{\tau}\Big) = f - v\Big(\frac{1}{\tau}\Big) + \frac{1}{\tau} v'\Big(\frac{1}{\tau}\Big) = f - w(\tau);$$

hence, (w, r) fulfills the inverse scale space equality. It remains to check that $r(\tau) \in \partial J(w(\tau))$. We use the fact that according to Theorem 4.5 the solutions of the gradient flow and the variational problem coincide, i.e, $v = u$ and $v' = u' = -p$, to obtain

$$J(w(\tau)) = J\Big(v\Big(\frac{1}{\tau}\Big) - \frac{1}{\tau} v'\Big(\frac{1}{\tau}\Big)\Big) = J\Big(v\Big(\frac{1}{\tau}\Big) + \frac{1}{\tau} p\Big(\frac{1}{\tau}\Big)\Big) \le J\Big(v\Big(\frac{1}{\tau}\Big)\Big) + \frac{1}{\tau} J\Big(p\Big(\frac{1}{\tau}\Big)\Big).$$

Using that the subgradients $p(1/\tau)$ are eigenvectors with minimal norm in $\partial J(u(1/\tau))$ and invoking Proposition 3.6 we infer $J(p(1/\tau)) = \|p(1/\tau)\|^2 = \langle q(1/\tau), p(1/\tau)\rangle$. By inserting this and using that $q(1/\tau) \in \partial J(v(1/\tau))$ we obtain

$$J(w(\tau)) \le \Big\langle q\Big(\frac{1}{\tau}\Big), v\Big(\frac{1}{\tau}\Big)\Big\rangle + \frac{1}{\tau}\Big\langle q\Big(\frac{1}{\tau}\Big), p\Big(\frac{1}{\tau}\Big)\Big\rangle = \langle r(\tau), w(\tau)\rangle.$$

The fact that $r(\tau) = q(1/\tau) \in K$ finally shows that $r(\tau) \in \partial J(w(\tau))$. Once again, the uniqueness of solutions of (VP) and (ISS) concludes the proof. □

4C. ***Orthogonality of the decomposition.*** Simple examples of flows with subgradients which are piecewise constant in time show that it is false that two subgradients of a gradient flow corresponding to different time points are orthogonal. However, we are able to show that the differences of subgradients are orthogonal if the subgradients themselves are eigenvectors.

Theorem 4.7. *If the gradient flow (GF) generates eigenvectors, it holds for* $0 < r \le s \le t$ *that*

$$\langle p(t), p(s) - p(r) \rangle = 0. \tag{4-10}$$

Proof. As stated in Theorem 4.5, it holds $p(r), p(s), p(t) \in \partial J(u(t))$ with $\|p(t)\|$ minimal in $\partial J(u(t))$. Hence, the assertion follows directly by employing Proposition 3.6 with $u := u(t)$, $\hat{p} := p(t)$, and $p \in \{p(s), p(r)\}$ to obtain $\langle p(t), p(s) \rangle = \|p(t)\|^2 = \langle p(t), p(r) \rangle$. □

We already know from Theorem 4.5 that if the gradient flow generates eigenvectors it holds $p(s) \in \partial J(u(t))$ for all $s \le t$. Using Theorem 4.7 above one can even show:

Corollary 4.8. *If the gradient flow (GF) generates eigenvectors, it holds for all* $0 < s \le t$ *that*

$$p(s) \in \partial J(p(t)). \tag{4-11}$$

Proof. Equation (4-10) implies $\|p(t)\|^2 = \langle p(s), p(t) \rangle$ for all $0 \le s \le t$. This yields

$$J(p(t)) = \|p(t)\|^2 = \langle p(s), p(t) \rangle \le J(p(t)),$$

and hence $p(s) \in \partial J(p(t))$. □

At this point the most important consequence of Theorem 4.7 is the following corollary, which states that differences of eigenvectors generated by the gradient flow are orthogonal.

Corollary 4.9. *If one defines the* ***spectral increments***

$$\phi(s,t) = p(t) - p(s), \quad s, t > 0,$$

then Theorem 4.7 implies

$$\langle \phi(s_1, t_1), \phi(s_2, t_2) \rangle = 0$$

for all $0 < s_1 \le t_1 \le s_2 \le t_2$.

Ideally, one would like to obtain this orthogonality relation for the time derivative of p which, however, only exists in a distributional sense. Formally, one defines

$$\phi(t) = -t\, \partial_t p(t)$$

to obtain an orthogonal spectral representation of the data f in the sense of [Burger et al. 2016]; i.e.,

$$f - \bar{f} = \int_0^\infty \phi(t)\, \mathrm{d}t, \quad \langle \phi(s), \phi(t) \rangle = 0, \ s \ne t.$$

Here, ϕ formally acts as spectral measure. Since, however, this approach fails due to the lacking regularity of $t \mapsto p(t)$, we will present a rigorous definition of a spectral measure in the next section.

Remark 4.10. It remains to be mentioned that all results in Sections 4B and 4C have already been proved in a finite-dimensional setting [Burger et al. 2016] and assuming that $K \subset \mathbb{R}^n$ is a polyhedron that satisfies (MINSUB) or the stronger condition that $J(u) = \|Au\|_1$ with a matrix A such that AA^T is diagonally dominant. However, these results are not stronger than ours because using the results in Section 3 it is trivial to show that K satisfying (MINSUB) implies that the gradient flow yields eigenvectors (see also Theorem 5.1). Hence, our result are a proper generalization to infinite dimensions and, furthermore, do not require a special structure of the functional J but only that the gradient flow produces eigenvectors.

4D. ***Large-time behavior and the spectral measure.*** In this section we aim at defining a measure that corresponds to the spectral measure of linear operator theory in the case that the gradient flow (GF) generates eigenvectors. However, all statements in this section are true without this assumption. As already mentioned in the Introduction, the gradient flow (GF) formally gives rise to a decomposition of an arbitrary datum into subgradients which takes the form

$$f - \bar{f} = \int_0^\infty p(t)\,\mathrm{d}t, \tag{4-12}$$

where $\bar{f}$ is the null-space component of $f \in \mathrm{dom}(J)$ given by (2-5). This decomposition is the basis for constructing the spectral measure; however, up to now it is formal due to the improper integral.

In order to make (4-12) rigorous, we have to investigate the large-time behavior of the gradient flow first. First we will show that the gradient flow has a unique strong limit as time tends to infinity which is given by $\bar{f}$. This will allow us to compute

$$\int_0^\infty p(t)\,\mathrm{d}t = -\int_0^\infty \partial_t u(t)\,\mathrm{d}t = u(0) - \lim_{t\to\infty} u(t) = f - \bar{f},$$

which gives the decomposition. From there on we will construct the spectral measure.

Theorem 4.11 (large-time behavior). *Let u solve (GF). Then $u(t)$ strongly converges to $\bar{f}$ as $t \to \infty$.*

Proof. The proof for strong convergence of $u(t)$ to some $u_\infty \in \mathcal{N}(J)$ as $t \to \infty$ is given in [Bruck 1975, Theorem 5] for even and hence, in particular, for absolutely one-homogeneous J. To see that $u_\infty = \bar{f}$ we observe

$$\langle u(t) - f, v\rangle = \int_0^t \langle \partial_t u(s), v\rangle\,\mathrm{d}s = -\int_0^t \langle p(s), v\rangle\,\mathrm{d}s = 0 \quad \text{for all } v \in \mathcal{N}(J),\ t > 0,$$

by using Proposition 2.2(7). Hence $u(t) - f \in \mathcal{N}(J)^\perp$ and by using the strong closedness of the orthogonal complement in Hilbert spaces we infer $u_\infty - f \in \mathcal{N}(J)^\perp$ and $u_\infty \in \mathcal{N}(J)$. This is equivalent to $u_\infty = \bar{f}$. □

Corollary 4.12. *Let u solve (GF), with $f \in \mathrm{dom}(J)$. Then it holds*

$$\lim_{T\to\infty} \left\| \int_0^T p(t)\,\mathrm{d}t - (f - \bar{f}) \right\| = 0, \tag{4-13}$$

which implies (4-12).

Finally, the following lemma implies that one can without loss of generality assume $\bar{f} = 0$.

Lemma 4.13. *The solution of $\partial_t u = -p,\ p \in \partial J(u)$, with $u(0) = f$, is given by $u = v + \bar{f}$, where v solves $\partial_t v = -p,\ p \in \partial J(v)$, with $v(0) = f - \bar{f}$.*

Proof. The proof is trivial by using Proposition 2.2(3) and (7). □

Now we consider (u, p) solving $\partial_t u = -p$ with $u(0) = f$ and, without loss of generality, we can assume $\bar{f} = 0$. This can always be achieved by replacing f with $f - \bar{f}$ and using the previous lemma. Note that $\bar{f} = 0$ if and only if $f \in \mathcal{N}(J)^\perp$. Then by Corollary 4.12 we infer that

$$f = \int_0^\infty p(t)\,\mathrm{d}t. \tag{4-14}$$

Let us compare this to the statement of the spectral theorem for a self-adjoint linear operator $T : \mathcal{H} \to \mathcal{H}$. For these one has

$$\Phi(T) f = \int_{\sigma(T)} \Phi(\lambda)\, \mathrm{d}E_\lambda f,$$

where $\sigma(T)$ denotes the spectrum of T, E is the spectral measure, $\Phi : \mathbb{R} \to \mathbb{R}$ is a function, and $f \in \mathcal{H}$. By choosing $\Phi = \mathrm{id}$ one obtains the spectral decomposition of the operator T itself and by choosing $\Phi = 1$ one obtains the decomposition of the identity instead:

$$f = \int_{\sigma(T)} \mathrm{d}E_\lambda f. \tag{4-15}$$

If T is even compact, the spectral measure becomes atomic and is given by

$$E_\lambda = \sum_{k=1}^{\infty} \delta_{\lambda_k}(\lambda) \langle \cdot\,, e_k \rangle e_k, \tag{4-16}$$

where $(e_k)_{k\in\mathbb{N}}$ denotes a set of orthonormal eigenvectors of T with a null-sequence of eigenvalues $(\lambda_k)_{k\in\mathbb{N}}$. Plugging this in, one can express any $f \in \mathcal{H}$ by a linear combination of eigenvectors:

$$f = \sum_{k=1}^{\infty} \langle f, e_k \rangle e_k. \tag{4-17}$$

Consequently, our aim is to manipulate the measure $p(t)\,\mathrm{d}t$ in (4-14) in such a way that it becomes a nonlinear generalization of (4-15)–(4-17) for the case of the maximal monotone operator ∂J. In particular, it should become atomic if $t \mapsto p(t)$ is a sequence of countably/finitely many distinct eigenvectors with eigenvalue 1. To achieve this we condense all $t > 0$ with the same value of $\|p(t)\|$ into one atomic point $\lambda(t) = \|p(t)\|$, which can be considered as the corresponding eigenvalue of $e(t) := p(t)/\|p(t)\|$. Since the function $t \mapsto \lambda(t)$ is nonincreasing and converges to zero according to Theorem 4.2 and [Brézis 1975, Theorem 7], this yields a perfect analogy to the linear situation where only orthogonality has to be replaced by orthogonality of differences of eigenvectors.

Definition 4.14 (spectral measure). Let (u, p) solve $\partial_t u = -p$, $p \in \partial J(u)$, with $u(0) = f \in \mathcal{N}(J)^\perp \cap \mathrm{dom}(J)$, and let $\tilde{\mu}$ be the measure

$$\tilde{\mu}(A) = \int_A p(t)\,\mathrm{d}t \tag{4-18}$$

for Borel sets $A \subset (0, \infty)$. If we set $\lambda : (0, \infty) \to [0, \infty)$, $t \mapsto \lambda(t) := \|p(t)\|$, then the *spectral measure* μ of f with respect to J is defined as the pushforward of $\tilde{\mu}$ through λ; i.e.,

$$\mu(B) := \tilde{\mu}(\lambda^{-1} B) \tag{4-19}$$

for Borel sets $B \subset [0, \infty)$. The *spectrum* of μ is given by

$$\sigma(\mu) := \{\lambda(t) : t > 0\}. \tag{4-20}$$

Remark 4.15. Note that $t \mapsto \lambda(t)$ is indeed a measurable map since it is nonincreasing according to Theorem 4.2(4), which makes μ well-defined.

By definition of μ it holds

$$\int_{\sigma(\mu)} \mathrm{d}\mu = \int_{(0,\infty)} \mathrm{d}\tilde{\mu} = f;$$

i.e., μ has a reconstruction property like (4-15). Furthermore, if $t \mapsto p(t)$ is a collection of countably many distinct eigenvectors, the map $t \mapsto \lambda(t)$ only has a countable range. Consequently, the measure μ — which is supported on the range of λ — becomes concentrated in countably many points and, hence, atomic. This is the analogy to the linear case (4-16).

4E. ***A necessary and sufficient condition for spectral decompositions.*** Before moving to examples of specific functionals whose gradient flows yield spectral decompositions of *any data*, we first give a necessary and sufficient condition on the data such that the gradient flow of *any functional* computes a spectral decomposition of the data. More precisely, we show that the necessary condition (4-11) derived in Corollary 4.8 is also sufficient. This condition generalizes the (SUB0) + orthogonality condition defined in [Schmidt et al. 2018], which appears to be the only sufficient condition in that line.

Theorem 4.16. *Let $T > 0$ and assume that*

$$f = \int_0^\infty p(s)\,\mathrm{d}s, \quad p(s) \in \partial J(p(t)) \quad \textit{for all } 0 < s \le t. \tag{4-21}$$

Then $u(t) = \int_t^\infty p(s)\,\mathrm{d}s$ solves the gradient flow (GF); i.e., $\partial_t u(t) = -p(t) \in \partial J(u(t))$ for $t > 0$.

Proof. Obviously, it holds $\partial_t u(t) = -p(t)$ and thus it remains to be checked that $p(t) \in \partial J(u(t))$ for $t > 0$. We compute

$$J(u(t)) = J\left(\int_t^\infty p(s)\,\mathrm{d}s\right) \le \int_t^\infty J(p(s))\,\mathrm{d}t = \int_t^T \langle p(t), p(s)\rangle\,\mathrm{d}t = \langle p(t), u(t)\rangle.$$

Together with $p(t) \in K$ for all $t > 0$, this concludes the proof. □

Definition 4.17 ((SUB0) + orthogonality [Schmidt et al. 2018]). We say that $f \in \mathcal{H}$ satisfies the (SUB0) + orthogonality condition if there are $N \in \mathbb{N}$ and positive numbers γ_i, λ_i for $i = 1, \dots, N$ such that

$$f = \sum_{i=1}^N \gamma_i u_i,$$

where $p_i := \lambda_i u_i$ satisfy

- $p_i \in \partial J(p_i)$ for $i = 1, \dots, N$,
- $\sum_{i=j}^N p_i \in \partial J(0)$ for all $j = 1, \dots, N$,
- $\langle p_i, p_j\rangle = 0$ for $i, j = 1, \dots, N$ with $i \neq j$.

In [Schmidt et al. 2018] the authors proved that the inverse scale space flow (ISS) yields a decomposition into eigenvectors if the datum satisfies the conditions above. In the following we prove that their condition is a special case of our necessary and sufficient condition (4-21).

Theorem 4.18. *Any datum f fulfilling the conditions from Definition 4.17 also satisfies (4-21).*

Proof. Any finite representation of the data as

$$f = \sum_{i=1}^{N} \gamma_i u_i,$$

with numbers $\gamma_i \geq 0$ and eigenvectors u_i satisfying $\lambda_i u_i \in \partial J(u_i)$, can be rewritten as

$$f = t_1 \sum_{i=1}^{N} \lambda_i u_i + (t_2 - t_1) \sum_{i=2}^{N} \lambda_i u_i + \cdots + (t_N - t_{N-1}) \sum_{i=N}^{N} \lambda_i u_i,$$

where $t_i := \gamma_i / \lambda_i$, and we can assume the ordering $t_i < t_j$ for $i < j$. Consequently, we can define

$$p(s) := \sum_{i=j}^{N} \lambda_i u_i, \quad s \in [t_{j-1}, t_j],$$

for $i = 1, \ldots, N$ and $t_0 := 0$. On one hand, this yields $f = \int_0^T p(s)\,\mathrm{d}s$, where $T := t_N$. On the other hand, using [Schmidt et al. 2018, Proposition 3.4] one can easily calculate that $p(s) \in \partial J(p(t))$ holds for $s \leq t$, which is condition (4-21). □

4F. *Short-time behavior.* Finally, we investigate the consequences of a gradient flow which generates eigenvectors for the short-time behavior of the map $t \mapsto u(t)$. From the temporal continuity it is obvious that $u(t) \to u(0) = f$ in $\mathcal{H}$ as $t \searrow 0$, which holds in general. However, in the case that $f \in \mathrm{dom}(J)$ and the gradient flow yields a spectral decomposition into eigenvectors, more can be shown. In fact, $u(t)$ converges to f also in the seminorm J, which generalizes known results for the one-dimensional ROF problem [Overgaard 2018].

Theorem 4.19. *Let $f \in \mathrm{dom}(J)$ and assume that the gradient flow (GF) generates eigenvectors. Then it holds $J(f - u(t)) \to 0$ as $t \searrow 0$.*

Proof. Using the definition of the gradient flow, we write $f - u(t) = \int_0^t p(s)\,\mathrm{d}s$. Since all the subgradients $p(s)$ for $s > 0$ are eigenvectors, this implies, together with the triangle inequality from Remark 2.4,

$$0 \leq J(f - u(t)) \leq \int_0^t J(p(s))\,\mathrm{d}s = \int_0^t \|p(s)\|^2\,\mathrm{d}s.$$

According to [Brézis 1971] the map $s \mapsto \|p(s)\|$ is in $L^2(0, \delta)$ for all $\delta > 0$ if and only if $f \in \mathrm{dom}(J)$. Hence, the dominated convergence theorem yields $\lim_{t \searrow 0} J(f - u(t)) = 0$, as desired. □

5. Examples of flows that yield a spectral decomposition

In the following we discuss some relevant examples of functionals J and corresponding flows that yield a spectral decomposition, meaning that the decomposition

$$f - \bar{f} = \int_0^\infty p(t)\,\mathrm{d}t$$

induced by the gradient flow is a decomposition into eigenvectors $p(t) \in \partial J(p(t))$ for all $t > 0$.

5A. *Polyhedral flow with* (MINSUB). As already indicated, we can prove that condition (MINSUB) introduced in Proposition 3.7 is already sufficient for the gradient flow yielding eigenvectors, which is an improvement of the results in [Burger et al. 2016].

Theorem 5.1 (polyhedral gradient flow). *Let $K = \partial J(0)$ be a polyhedron satisfying (MINSUB) from Proposition 3.7. Then the gradient flow $\partial_t u = -p$, $p \in \partial J(u)$ with $u(0) = f$, yields a sequence of eigenvectors $p(t)$ for $t > 0$. Furthermore, $t \mapsto p(t)$ takes only finitely many different values and there is $T > 0$ such that $p(t) = 0$ for all $t \geq T$.*

Proof. The gradient flow selects subgradients of minimal norm (see Theorem 4.2); all of them are eigenvectors due to Proposition 3.7. According to [Burger et al. 2016] we also know that p is piecewise constant in t and that the flow becomes extinct in finite time. □

As mentioned in Remark 4.10, in finite dimensions functionals of the type $J(u) = \|Au\|_1$ with AA^T diagonally dominant satisfy (MINSUB) and thus suffice for the gradient flow to yield eigenvectors. Therefore, we will now study functionals of the type $J(u) = \|Au\|_1$ on $\mathcal{H} = L^2$, where A is a suitable operator such that formally AA' is diagonally dominant. Here A' denotes the adjoint of A with respect to the inner product on L^2. Apart from the trivial choice $A = \mathrm{id}$, which will be shortly dealt with in the following section, a natural choice is $A = \mathrm{div}$ since then formally $AA' = \mathrm{div}\,\nabla = -\Delta$, which is diagonally dominant. Indeed, this does the trick as we will see. Note that in one space dimension, A reduces to the usual derivation operator, which means that $J(u) = \|u'\|_1$ formally becomes the total variation of u. Since, compared to the one-dimensional case, the proof that the subgradients of minimal norm are eigenvectors is much more technically involved in two or more space dimensions, we give the one-dimensional proof first to illustrate the ideas before proving the statement in arbitrary space dimension in the subsequent section.

5B. *L^1-flow.* We consider $\mathcal{H} = L^2(\Omega)$ and

$$J(u) = \begin{cases} \|u\|_{L^1(\Omega)}, & u \in L^1(\Omega) \cap L^2(\Omega), \\ +\infty, & \text{else.} \end{cases} \tag{5-1}$$

Then it can be easily seen that $K = \{p \in L^2(\Omega) : \|p\|_{L^\infty} \leq 1\}$ and for any $u \in \mathrm{dom}(\partial J)$ we have

$$\partial J(u) = \left\{ p \in K : \int_\Omega pu \,\mathrm{d}x = \int_\Omega |u| \,\mathrm{d}x \right\},$$

i.e., $p(x) = u(x)/|u(x)|$ for $u(x) \neq 0$ and $p(x) \in [-1, 1]$ else. The subgradient of minimal norm in $\partial J(u)$ fulfills $p(x) = 0$ in the latter case.

Proposition 5.2. *Let $u \in \mathrm{dom}(\partial J)$ and let $p \in \partial J(u)$ be the subgradient of minimal norm. Then it holds $\langle p, p - q\rangle \geq 0$ for all $q \in K$.*

Proof. We calculate with Cauchy–Schwarz

$$\langle p, p - q\rangle = \int_\Omega |p|^2 \,\mathrm{d}x - \int_\Omega pq \,\mathrm{d}x \geq \int_\Omega |p|^2 \,\mathrm{d}x - \int_\Omega |p| \,\mathrm{d}x \geq 0$$

since p only takes the values 0, -1, or $+1$ and thus satisfies $|p| = |p|^2$. □

Thus, we obtain the following result:

Theorem 5.3 (L^1-flow). *Let J be given by (5-1). Then the gradient flow $\partial_t u(t) = -p(t),\ p(t) \in \partial J(u(t))$, with $u(0) = f$, generates a sequence of eigenvectors $p(t)$ for $t > 0$.*

Of course this example is of limited practical relevance, but can yield some insight since we can explicitly compute the solutions of the gradient flow noticing that the subgradient just equals the sign of u pointwise; hence splitting $f = f_+ - f_-$ with nonnegative functions f_+ and f_- of disjoint support we have

$$u(x,t) = (f_+(x) - t)_+ - (f_-(x) - t)_+.$$

5C. *L^∞-flow.* As before we consider $\mathcal{H} = L^2(\Omega)$ and define

$$J(u) = \begin{cases} \|u\|_{L^\infty(\Omega)}, & u \in L^1(\Omega) \cap L^\infty(\Omega), \\ +\infty, & \text{else.} \end{cases} \tag{5-2}$$

It is obvious that $K = \{p \in L^2(\Omega) : \|p\|_{L^1(\Omega)} \le 1\}$ and in order to study $\partial J(u)$ we need to define the set of points where $|u|$ attains its essential maximum:

$$\Omega_{\max} := \{x \in \Omega : |u(x)| = J(u)\}, \tag{5-3}$$

which is defined up to Lebesgue measure zero. It is easy to see that for every $u \in \operatorname{dom}(\partial J)$ it holds

$$\partial J(u) = \{p \in L^2(\Omega) : \|p\|_{L^1(\Omega)} = 1,\ p = 0 \text{ a.e. in } \Omega \setminus \Omega_{\max},\ \operatorname{sgn}(p) = \operatorname{sgn}(u) \text{ a.e. in } \Omega_{\max}\}.$$

Note that the subgradient of minimal norm is given by

$$p(x) = \begin{cases} \operatorname{sgn}(u)/|\Omega_{\max}|, & x \in \Omega_{\max}, \\ 0, & \text{else,} \end{cases}$$

which follows from

$$1 = \|q\|^2_{L^1(\Omega)} \le |\Omega_{\max}|\,\|q\|^2_{L^2(\Omega)} \quad \text{for all } q \in \partial J(u)$$

together with the fact that $\|p\|^2_{L^2(\Omega)} = 1/|\Omega_{\max}|$.

Proposition 5.4. *Let $u \in \operatorname{dom}(\partial J)$ and let $p \in \partial J(u)$ be the subgradient of minimal norm. Then it holds $\langle p, p - q\rangle \ge 0$ for all $q \in K$.*

Proof. Using $\|q\|_{L^1(\Omega)} \le 1$ for $q \in K$ and the explicit form of the subgradient of minimal norm, we calculate

$$\langle p, p - q\rangle = \|p\|^2_{L^2(\Omega)} - \frac{1}{|\Omega_{\max}|}\int_{\Omega_{\max}} \operatorname{sgn}(u) q \,\mathrm{d}x \ge \frac{1}{|\Omega_{\max}|} - \frac{1}{|\Omega_{\max}|}\int_{\Omega_{\max}} |q|\,\mathrm{d}x \ge 0. \qquad \square$$

Just as before, we obtain the following result:

Theorem 5.5 (L^∞-flow). *Let J be given by (5-2). Then the gradient flow $\partial_t u(t) = -p(t),\ p(t) \in \partial J(u(t))$, with $u(0) = f$, generates a sequence of eigenvectors $p(t)$ for $t > 0$.*

5D. *One-dimensional total variation flow.* Let $I \subset \mathbb{R}$ be a bounded interval and J be the total variation defined on $L^2(I)$, i.e., extended by infinity outside $\mathrm{BV}(I)$. From [Briani et al. 2011] we infer that $K = \{-g' : g \in H_0^1(I, [-1, 1])\}$ and

$$\partial J(u) = \{-g' : g \in H_0^1(I, [-1, 1]),\ g \equiv \pm 1 \text{ on } \operatorname{supp}((Du)_\pm)\}, \tag{5-4}$$

where $Du = (Du)_+ - (Du)_-$ is the Jordan decomposition of Du. We start with some results further characterizing $\operatorname{dom}(\partial J)$ and subgradients of minimal norm in this case:

Lemma 5.6. *Let $u \in \operatorname{dom}(\partial J)$ and $p \in \partial J(u) \setminus \{0\}$. Moreover, let $Du = (Du)_+ - (Du)_-$ be the Jordan decomposition of Du. Then for each $x_\pm \in \operatorname{supp}((Du)_\pm)$ the estimate*

$$|x_+ - x_-| \geq \frac{4}{\|p\|_{L^2}^2} \tag{5-5}$$

holds. Moreover, the distance of $x_\pm$ from ∂I is bounded below by $1/\|p\|_{L^2}^2$.

Proof. We write $p = -g'$ according to (5-4). Thus, for $x_\pm \in \operatorname{supp}((Du)_\pm)$ we find with the Cauchy–Schwarz inequality

$$2 = |g(x_+) - g(x_-)| = \left| \int_{x_-}^{x_+} g'(x)\,\mathrm{d}x \right| \leq \sqrt{|x_+ - x_-|}\ \|p\|_{L^2},$$

which yields (5-5). The estimate for the distance to the boundary follows by an analogous argument noticing that g has zero boundary values. □

Lemma 5.7. *Let $u \in \operatorname{dom}(\partial J)$ and $p = -g'$ be the element of minimal L^2-norm in $\partial J(u)$. Then g is a piecewise linear spline with zero boundary values and a finite number of kinks where g attains the values $+1$ or -1.*

Proof. First of all, due to (5-5) there can only be a finite number of changes of g between $+1$ and -1. Let y, z be two points in I such that $g(y) = g(z) = 1$ and $(y, z) \cap \operatorname{supp}(Du)_- = \varnothing$. Let

$$\tilde{g}(x) = \begin{cases} \max(g(x), 1) & \text{if } x \in (y, z), \\ g(x) & \text{else.} \end{cases}$$

Then $-\tilde{g}' \in \partial J(u)$ and

$$\int_I (\tilde{g}')^2\,\mathrm{d}x - \int_I (g')^2\,\mathrm{d}x = -\int_y^z (g')^2\,\mathrm{d}x \leq 0,$$

with equality if and only if $p = -g'$ vanishes in (y, z). Due to the minimality of g' we find that $g \equiv 1$ on (y, z). By an analogous argument we can show that $g \equiv -1$ on each interval (y, z) such that $g(y) = g(z) = -1$ and $(y, z) \cap \operatorname{supp}((Du)_+) = \varnothing$. Now let $y, z \in I$ be such that $g(y) = 1$, $g(z) = -1$ and $(y, z) \cap \operatorname{supp}(Du) = \varnothing$. Then we can define

$$\tilde{g}(x) = \begin{cases} g(y) + \dfrac{x-y}{z-y}(g(z) - g(y)) & \text{if } x \in (y, z), \\ g(x) & \text{else,} \end{cases}$$

and see again that

$$\int_I (\tilde{g}')^2\,\mathrm{d}x \leq \int_I (g')^2\,\mathrm{d}x.$$

A similar argument shows that g is piecewise linear close to the boundary of I, changing from 0 to $+1$ or -1, which completes the assertion. □

Lemma 5.8. *Let $u \in \operatorname{dom}(\partial J)$ and p be the element of minimal L^2-norm in $\partial J(u)$. Then $\langle p, p-q\rangle \geq 0$ for all $q \in K$.*

Proof. Let the jump set of $p = -g'$ be $\{x_i\}_{i=1}^N$ in ascending order. Then $g \in H_0^1(I, [-1, 1])$ satisfies either $g(x) = 1$ or $g(x) = -1$ in $x = x_i$ and is piecewise linear in between. Denoting by x_0 and x_{N+1} the boundary points of I we can use $g(x_0) = g(x_{N+1}) = 0$ and have g as the piecewise linear spline between all the x_i. Now we can write any $q \in \partial J(0)$ as $q = -h'$ for some $h \in H_0^1(I, [-1, 1])$. Thus, we have

$$\langle p, p-q\rangle = \int_I g'(g'-h')\,\mathrm{d}x = \sum_{i=0}^{N} \int_{x_i}^{x_{i+1}} g'(g'-h')\,\mathrm{d}x.$$

Since g' is piecewise constant we find

$$\int_{x_i}^{x_{i+1}} g'(g'-h')\,\mathrm{d}x = \frac{g(x_{i+1})-g(x_i)}{x_{i+1}-x_i}(g(x_{i+1})-g(x_i)-h(x_{i+1})+h(x_i)).$$

If $g(x_{i+1}) = g(x_i)$, then the integral vanishes. If $0 < i < N$ and $g(x_{i+1}) \neq g(x_i)$, then

$$\begin{aligned}\int_{x_i}^{x_{i+1}} g'(g'-h')\,\mathrm{d}x &= \frac{4}{x_{i+1}-x_i} - \frac{g(x_{i+1})-g(x_i)}{x_{i+1}-x_i}(h(x_{i+1})-h(x_i)) \\ &\geq \frac{4-2|h(x_{i+1})-h(x_i)|}{x_{i+1}-x_i} \geq 4\frac{1-\|h\|_\infty}{x_{i+1}-x_i} \geq 0.\end{aligned}$$

A similar argument for the boundary terms $i = 0$ and $i = N$ finally yields

$$\langle p, p-q\rangle \geq 0.$$

□

Remark 5.9. Analogously, similar computations can be made in the case of an unbounded interval I or the case $I = \mathbb{R}$ where the subgradients are given by $p = -g'$, where $|g(x)| \leq 1$ and $g' \in L^2(\mathbb{R})$.

Theorem 5.10 (one-dimensional total variation flow). *Let J be given by the one-dimensional total variation on an interval $I \subset \mathbb{R}$. Then the gradient flow $\partial_t u(t) = -p(t)$, $p(t) \in \partial J(u(t))$, with $u(0) = f$, yields a sequence of eigenvectors $p(t)$ for $t > 0$.*

Remark 5.11 (geometric structure of eigenvectors). From the characterization of the subdifferential (5-4) of the one-dimensional total variation, we conclude that p is an eigenvector if and only if $p = -g'$, where $g \equiv \pm 1$ on $\operatorname{supp}((-g'')_\pm)$. In particular, this implies that p is a step function which jumps at the kinks of g.

5E. *Divergence flow.* Let $n \geq 2$ and $\Omega \subset \mathbb{R}^n$ be an open and bounded set. We consider the gradient flow

$$\partial_t u = -p, \quad p \in \partial J(u), \tag{5-6}$$

where

$$J(u) := \int_\Omega |\operatorname{div} u| := \sup\left\{-\int_\Omega u \cdot \nabla v\,\mathrm{d}x : v \in C_c^\infty(\Omega, [-1, 1])\right\}, \quad u \in L^2(\Omega, \mathbb{R}^n), \tag{5-7}$$

and $u(0) \in L^2(\Omega, \mathbb{R}^n)$. It holds that $u \in \mathrm{dom}(J)$, i.e., $J(u) < \infty$, if and only if the distribution $\operatorname{div} u$ can be represented as a finite Radon measure; see [Briani et al. 2011]. The null-space $\mathcal{N}(J)$ of J is infinite-dimensional since it contains all vector fields of the type $\nabla v + w$, where v is an harmonic function and w is a divergence-free vector field. Note that J can be cast into the canonical form $J = \chi_K^*$ by setting

$$K := \{-\nabla v : v \in H_0^1(\Omega, [-1, 1])\}. \tag{5-8}$$

Hence, a subgradient p of J in u fulfills $p \in K$ and $\int_\Omega p \cdot u = \int_\Omega |\operatorname{div} u| \, \mathrm{d}x$. To understand the meaning of v in (5-8), we perform a integration by parts for smooth u to obtain

$$\int_\Omega p \cdot u \, \mathrm{d}x = -\int_\Omega \nabla v \cdot u \, \mathrm{d}x = \int_\Omega v \, \operatorname{div} u \, \mathrm{d}x.$$

Therefore, v should be chosen as

$$v(x) \in \begin{cases} \{1\}, & \operatorname{div} u(x) > 0, \\ \{-1\}, & \operatorname{div} u(x) < 0, \\ [-1, 1], & \operatorname{div} u(x) = 0, \end{cases} \quad x \in \Omega. \tag{5-9}$$

In the general case, one considers the polar decomposition of the measure $\operatorname{div} u$, given by $\operatorname{div} u = \theta_u \, |\operatorname{div} u|$, where $\theta_u(x) := (\operatorname{div} u / |\operatorname{div} u|)(x)$ denotes the Radon–Nikodym derivative, which exists $|\operatorname{div} u|$-almost everywhere and has values in $\{-1, 1\}$. This allows us to define the sets

$$\mathcal{E}_u^\pm := \{x \in \Omega : \theta_u(x) \text{ exists and } \theta_u(x) = \pm 1\}. \tag{5-10}$$

Note that trivially it holds $\mathcal{E}_u^+ \cup \mathcal{E}_u^- \subset \mathrm{supp}(\operatorname{div} u)$ and one can easily prove that $\mathrm{supp}(\operatorname{div} u) = \overline{\mathcal{E}_u^+ \cup \mathcal{E}_u^-}$. In [Briani et al. 2011] the authors showed that, in full analogy to (5-9), the subdifferential of J in $u \in \mathrm{dom}(J)$ can be characterized as

$$\partial J(u) = \{-\nabla v : v \in H_0^1(\Omega, [-1, 1]), \ v = \pm 1 |\operatorname{div} u|\text{-a.e. on } \mathcal{E}_u^\pm\}. \tag{5-11}$$

If $p = -\nabla v$ is a subgradient of J in u we refer to v as a *calibration* of u.

Remark 5.12. Since $v \in H_0^1(\Omega, [-1, 1])$ can be defined pointwise everywhere except from a set with zero H^1-capacity and the measure $\operatorname{div} u$ vanishes on such sets (see [Adams and Hedberg 1996, Chapter 6] and [Chen et al. 2009], respectively), the pointwise definition of v in (5-11) makes sense.

Since the gradient flow selects the subgradients with minimal norm, we are specifically interested in

$$p_u^0 = \arg\min \left\{ \int_\Omega |p|^2 \, \mathrm{d}x : p \in \partial J(u) \right\}.$$

Using (5-11) one has

$$p_u^0 = -\nabla \arg\min \left\{ \int_\Omega |\nabla v|^2 \, \mathrm{d}x : v \in H_0^1(\Omega, [-1, 1]), \ v = \pm 1 |\operatorname{div} u|\text{-a.e. on } \mathcal{E}_u^\pm \right\}, \tag{5-12}$$

which implies that $p_u^0 = -\nabla v$, where v minimizes the Dirichlet energy and, in particular, is harmonic on the open set $\Omega_H := \Omega \setminus (\overline{\mathcal{E}_u^+ \cup \mathcal{E}_u^-})$.

Now we show that the gradient flow (5-6) indeed generates eigenvectors. We have the following:

Proposition 5.13. *Let $u \in \operatorname{dom}(\partial J)$ and $p = p_u^0$. Then $p \in \partial J(p)$.*

To prove the proposition, by means of Proposition 3.2 it suffices to check that $\langle p, p-q\rangle \geq 0$ for all $q \in K$. We start with an approximation lemma.

Lemma 5.14. *Let $0 < \varepsilon < 1$ and define*

$$\psi_\varepsilon(t) = \begin{cases} 0, & |t| > 1, \\ 1, & |t| < 1-\varepsilon, \\ (1/\varepsilon)(1-|t|), & 1-\varepsilon \leq |t| \leq 1, \end{cases} \qquad \phi_\varepsilon(t) = \int_0^t \psi_\varepsilon(\tau)\,\mathrm{d}\tau. \tag{5-13}$$

If $v \in H_0^1(\Omega, [-1,1])$ then $v_\varepsilon := \phi_\varepsilon \circ v \in H_0^1(\Omega, [-1,1])$ and $v_\varepsilon \to v$ strongly in H^1 as $\varepsilon \searrow 0$.

Proof. The membership to $H_0^1(\Omega, [-1,1])$ follows directly from the chain rule for compositions of Lipschitz and Sobolev functions. For strong convergence it suffices to show $\nabla v_\varepsilon \to \nabla v$ in $L^2(\Omega)$. Using $\nabla v_\varepsilon = \psi_\varepsilon(v)\nabla v$ this follows from

$$\begin{aligned} \|\nabla v_\varepsilon - \nabla v\|_{L^2}^2 = \|(\psi_\varepsilon(v)-1)\nabla v\|_{L^2}^2 &= \int_{\{1-\varepsilon \leq |v| \leq 1\}} \overbrace{\left[\frac{1-|v(x)|}{\varepsilon}\right]^2}^{\leq 1} |\nabla v|^2\,\mathrm{d}x \\ &\leq \int_{\{1-\varepsilon \leq |v| \leq 1\}} |\nabla v|^2\,\mathrm{d}x \to \int_{\{|v|=1\}} |\nabla v|^2\,\mathrm{d}x = 0, \quad \varepsilon \searrow 0. \end{aligned}$$

□

Proof of Proposition 5.13. We write $p = p_u^0 = -\nabla v$, with v as in (5-12), and define v_ε as in the previous lemma. Any $q \in K$ can be written as $q = -\nabla w$ with $w \in H_0^1(\Omega, [-1,1])$ and thanks to Lemma 5.14 we have

$$\langle p, p-q\rangle = \int_\Omega \nabla v \cdot \nabla(v-w)\,\mathrm{d}x = \lim_{\varepsilon \searrow 0} \int_\Omega \nabla v_\varepsilon \cdot \nabla(v-w)\,\mathrm{d}x. \tag{5-14}$$

As a first step, we replace $v-w$ by a smooth z with compact support to obtain with the product rule

$$\int_\Omega \nabla v_\varepsilon \cdot \nabla z\,\mathrm{d}x = \int_\Omega \psi_\varepsilon(v)\nabla v \cdot \nabla z\,\mathrm{d}x = \int_\Omega \nabla v \cdot (\nabla[\psi_\varepsilon(v)z] - \psi_\varepsilon'(v)z\nabla v)\,\mathrm{d}x = -\int_\Omega \psi_\varepsilon'(v)z|\nabla v|^2\,\mathrm{d}x.$$

For the last equality we used that $\psi_\varepsilon(v)z$ is a test function for the minimization of the Dirichlet energy in (5-12) and that the first variation of the Dirichlet integral in the direction of the test function vanishes consequently.

Strongly approximating $v-w$ by smooth and compactly supported functions in $H_0^1(\Omega)$ and using the above-noted calculation results in

$$\int_\Omega \nabla v_\varepsilon \cdot \nabla(v-w)\,\mathrm{d}x = -\int_\Omega \psi_\varepsilon'(v)(v-w)|\nabla v|^2\,\mathrm{d}x. \tag{5-15}$$

Note that

$$\psi_\varepsilon'(t) = \frac{1}{\varepsilon}\begin{cases} 1, & t \in [-1, -1+\varepsilon], \\ -1, & t \in [1-\varepsilon, 1], \\ 0, & \text{else}, \end{cases}$$

which we can use to calculate

$$\begin{aligned}-\int_\Omega \psi_\varepsilon'(v)(v-w)|\nabla v|^2\,\mathrm{d}x &= \int_{\{-1\le v\le -1+\varepsilon\}} \frac{1}{\varepsilon}(w-v)|\nabla v|^2\,\mathrm{d}x + \int_{\{1-\varepsilon\le v\le 1\}} \frac{1}{\varepsilon}(v-w)|\nabla v|^2\,\mathrm{d}x \\ &\ge -\int_{\{1-\varepsilon\le |v|\le 1\}} |\nabla v|^2\,\mathrm{d}x \\ &\to -\int_{\{|v|=1\}} |\nabla v|^2\,\mathrm{d}x = 0, \quad \varepsilon \searrow 0. \end{aligned} \tag{5-16}$$

Putting (5-14)–(5-16) together, we have shown that

$$\langle p, p-q\rangle = \lim_{\varepsilon\searrow 0}\int_\Omega \nabla v_\varepsilon\cdot\nabla(v-w)\,\mathrm{d}x = \lim_{\varepsilon\searrow 0} -\int_\Omega \psi_\varepsilon'(v)(v-w)|\nabla v|^2\,\mathrm{d}x \ge 0,$$

as desired. □

Theorem 5.15 (divergence flow). *Let J be given by (5-7). Then the gradient flow $\partial_t u(t) = -p(t)$, $p(t)\in\partial J(u(t))$, with $u(0)=f$, generates a sequence of eigenvectors $p(t)$ for $t>0$.*

Remark 5.16 (geometric structure of eigenvectors). From (5-11) we infer that $p\in\partial J(p)$ if and only if $p=-\nabla v$, where $v\in H_0^1(\Omega,[-1,1])$ and $v=\pm 1$ on $(\operatorname{div} u)_\pm = (-\Delta v)_\pm$. Note that $-\Delta v$ is a finite Radon measure in this situation whose support cannot contain an open set since otherwise v would be locally constant there and hence have zero Laplacian. Therefore, eigenvectors are gradients of functions v which are harmonic everywhere on $\Omega\setminus\Gamma$, where $\Gamma := \operatorname{supp}(-\Delta v)$ is a closed exceptional set, not containing any open set.

5F. *Rotation flow.* Now we fix the dimension $n=2$ and consider the functional

$$\tilde J(u) := \int_\Omega |\operatorname{rot} u| := \sup\left\{\int_\Omega u\cdot\nabla^\perp v\,\mathrm{d}x : v\in C_c^\infty(\Omega,[-1,1])\right\}, \quad u\in L^2(\Omega,\mathbb{R}^n), \tag{5-17}$$

where formally $\operatorname{rot} u = \partial_1 u_2 - \partial_2 u_1$ and $\nabla^\perp = (\partial_2, -\partial_1)^T$. Defining the rotation matrix $R = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0\end{smallmatrix}\right)$, which fulfills $\nabla^\perp = R\nabla$, it holds $\tilde J(u) = J(Ru)$, where J is given by (5-7). As before, $\tilde J$ can be expressed by duality as $\tilde J = \chi_{\widetilde K}^*$, where $\widetilde K := \{\nabla^\perp v : v\in H_0^1(\Omega,[-1,1])\}$ and it holds $\widetilde K = RK$. Due to the invertibility of R, the gradient flows with respect to J and $\tilde J$ are fully equivalent and the respective solutions are connected via the rotation R. In particular, the results from the previous section directly generalize to $\tilde J$, meaning that the rotation flow $\partial_t u = -p,\ p\in\partial\tilde J(u)$, also generates eigenfunctions $p\in\partial\tilde J(p)$.

Theorem 5.17 (rotation flow). *Let $\tilde J$ be given by (5-17). Then the gradient flow $\partial_t u(t) = -p(t),\ p(t)\in\partial\tilde J(u(t))$, with $u(0)=f$, generates a sequence of eigenvectors $p(t)$ for $t>0$.*

5G. *Divergence-rotation flow.* Now we define the functional

$$\mathcal{J}(u) := \int_\Omega |\operatorname{div} u| + \int_\Omega |\operatorname{rot} u| = J(u) + \tilde J(u), \tag{5-18}$$

which measures the sum of the distributional divergence and rotation of a vector field $u\in L^2(\Omega,\mathbb{R}^n)$. Its null-space still contains all gradients of harmonic functions and we define $\mathcal{K} := \partial\mathcal{J}(0)$.

Proposition 5.18. *It holds for all* $u \in \operatorname{dom}(\partial \mathcal{J})$

$$\partial \mathcal{J}(u) = \partial J(u) + \partial \tilde{J}(u). \tag{5-19}$$

Furthermore, the sets $\partial J(u)$ *and* $\partial \tilde{J}(u)$ *are orthogonal.*

Crucial for the proof of this sum rule is the *Helmholtz decomposition*, which can be phrased as follows.

Theorem 5.19 (Helmholtz decomposition). *Let* $\Omega \subset \mathbb{R}^n$ *be an arbitrary domain and let*

$$L^2_\sigma(\Omega, \mathbb{R}^n) := \overline{\{\varphi \in C_c^\infty(\Omega, \mathbb{R}^n) : \operatorname{div} \varphi = 0\}}^{\|\cdot\|_{L^2}},$$
$$G^2(\Omega, \mathbb{R}^n) := \{-\nabla v \in L^2(\Omega, \mathbb{R}^n) : v \in L^2_{\mathrm{loc}}(\Omega)\}.$$

Then the ***Helmholtz decomposition***

$$L^2(\Omega, \mathbb{R}^n) = L^2_\sigma(\Omega, \mathbb{R}^n) \oplus G^2(\Omega, \mathbb{R}^n)$$

holds and is orthogonal. The orthogonal projection onto the closed subspace $L^2_\sigma(\Omega, \mathbb{R}^n)$ *of* $L^2(\Omega, \mathbb{R}^n)$ *is called the* ***Helmholtz projection*** *and denoted by* Π_σ.

Proof. The proof can be found in [Sohr 2001, Lemmas 2.5.1, 2.5.2]. □

Proof of Proposition 5.18. Let $u \in \operatorname{dom}(\partial \mathcal{J})$ and $p \in \partial \mathcal{J}(u)$. In particular, $p \in \mathcal{K}$ and hence it holds for arbitrary $v \in L^2(\Omega, \mathbb{R}^n)$ that

$$\langle p, v\rangle \le \mathcal{J}(v) = J(v) + \tilde{J}(v).$$

In particular, one has

$$\langle (1-\Pi_\sigma)p, v\rangle = \langle p, (1-\Pi_\sigma)v\rangle \le J((1-\Pi_\sigma)v) = J(v),$$
$$\langle \Pi_\sigma p, v\rangle = \langle p, \Pi_\sigma v\rangle \le \tilde{J}(\Pi_\sigma v) = \tilde{J}(v),$$

where we used Proposition 2.2(3) and the self-adjointness of projections. This shows $(1-\Pi_\sigma)p \in K$ and $\Pi_\sigma p \in \widetilde{K}$ and hence

$$\langle (1-\Pi_\sigma)p, u\rangle \le J(u),$$
$$\langle \Pi_\sigma p, u\rangle \le \tilde{J}(u).$$

If we assume that one of the inequalities is strict, then using $\langle p, u\rangle = \mathcal{J}(u)$ would imply

$$J(u) + \tilde{J}(u) = \mathcal{J}(u) = \langle p, u\rangle = \langle (1-\Pi_\sigma)p, u\rangle + \langle \Pi_\sigma p, u\rangle < J(u) + \tilde{J}(u),$$

which is a contradiction. Therefore, we have

$$\langle (1-\Pi_\sigma)p, u\rangle = J(u),$$
$$\langle \Pi_\sigma p, u\rangle = \tilde{J}(u),$$

which lets us conclude that $(1-\Pi_\sigma)p \in \partial J(u)$ and $\Pi_\sigma p \in \partial \tilde{J}(u)$. Hence, we have established $p \in \partial J(u) + \partial \tilde{J}(u)$, which concludes the first part of the proof. Orthogonality of $\partial J(u)$ and $\partial \tilde{J}(u)$ follows from the fact that any subgradient p_J and $p_{\tilde{J}}$ in these sets can be written as $p_J = -\nabla v$ and $p_{\tilde{J}} = \nabla^\perp w$, with $v, w \in H^1_0(\Omega, [-1, 1])$ respectively. Approximating w strongly by compactly supported and smooth

functions, it follows from an integration by parts using $\operatorname{div} \nabla^\perp w = 0$ and the zero boundary conditions of v that

$$\langle p_j, p_{\tilde{J}} \rangle = -\langle \nabla v, \nabla^\perp w \rangle = \langle v, \operatorname{div} \nabla^\perp w \rangle = 0. \qquad \square$$

Corollary 5.20. *Let $u \in \operatorname{dom}(\partial \mathcal{J})$. Then the subgradient of minimal norm in $\partial \mathcal{J}(u)$ is given by the sum of the subgradients of minimal norm in $\partial J(u)$ and $\partial \tilde{J}(u)$, respectively.*

Proof. The proof follows directly from Proposition 5.18. $\square$

Lemma 5.21. *Let p_J and $p_{\tilde{J}}$ be eigenvectors with eigenvalue 1 of ∂J and $\partial \tilde{J}$, respectively. Then it holds*:

(1) $\langle p_J, p_{\tilde{J}} \rangle = 0$.

(2) $p_J \in \mathcal{N}(\tilde{J})$, $p_{\tilde{J}} \in \mathcal{N}(J)$.

(3) $p := p_j + p_{\tilde{J}}$ *is an eigenvector with eigenvalue* 1 *of* $\partial \mathcal{J}$.

Proof. (1) According to Proposition 5.18 this holds for general subgradients.

(2) We only prove $p_J \in \mathcal{N}(\tilde{J})$, the proof of the second inclusion working analogously. We calculate

$$\tilde{J}(p_J) = J(Rp_J) = \sup_{p \in K} \langle p, Rp_J \rangle = \sup_{p \in K} \langle R^T p, p_J \rangle.$$

Due to the definition of K we can write $p_J = -\nabla v$ and $p = -\nabla w$, with $v, w \in H_0^1(\Omega, [-1, 1])$. This yields

$$\langle R^T p, p_J \rangle = \langle R \nabla w, -\nabla v \rangle = -\langle \nabla^\perp w, \nabla v \rangle.$$

By the same density argument as above we obtain $\langle R^T p, p_J \rangle = 0$ and hence $\tilde{J}(p_J) = 0$.

(3) Using that p_J and $p_{\tilde{J}}$ are eigenvectors, it holds

$$\begin{aligned} \langle p, p \rangle &= \langle p_J, p_J \rangle + \langle p_{\tilde{J}}, p_{\tilde{J}} \rangle + 2\overbrace{\langle p_J, p_{\tilde{J}} \rangle}^{=0} \\ &= J(p_J) + \tilde{J}(p_{\tilde{J}}) \\ &= J(p) + \tilde{J}(p) = \mathcal{J}(p). \end{aligned}$$

For the third equality we used (2) together with Proposition 2.2(3). $\square$

From here on we easily infer:

Theorem 5.22 (div-rot flow). *Let $\mathcal{J}$ be given by (5-18). Then the gradient flow $\partial_t u(t) = -p(t)$, $p \in \partial \mathcal{J}(u(t))$, with $u(0) = f$, yields a sequence of eigenvectors $p(t)$ for $t > 0$. Furthermore, $(1 - \Pi_\sigma) p(t)$ and $\Pi_\sigma p(t)$ are eigenvectors of ∂J and $\partial \tilde{J}$, respectively.*

Proof. The gradient flow selects subgradients of minimal norm, all of which are given by the sum of the subgradients of minimal norm in $\partial J(u)$ and $\partial \tilde{J}(u)$ according to Corollary 5.20. Due to Theorems 5.15 and 5.17, each subgradient of minimal norm is an eigenvector of ∂J and $\partial \tilde{J}$, respectively, and by Lemma 5.21(3) we conclude that their sum is an eigenvector of $\partial \mathcal{J}$. The uniqueness of the Helmholtz decomposition concludes the proof. $\square$

6. Finite extinction time and extinction profiles

Let us now consider a general gradient flow $\partial_t u(t) = -p(t),\ p(t) \in \partial J(u(t))$, with arbitrary $J \in \mathcal{C}$ and initial condition $u(0) = f$, where the class $\mathcal{C}$ is defined in Definition 2.1. It is well known (see [Benning and Burger 2013; Bungert and Burger 2019], for instance) that if the datum f is an eigenvector fulfilling $\lambda f \in \partial J(f)$, the corresponding solution of the gradient flow is given by $u(t) = \max(1 - \lambda t, 0) f$. Hence, there exists a time $T := 1/\lambda$, referred to as extinction time, such that $u(t) = 0 = \bar{f}$ for all $t \geq T$. In general, Theorem 4.11 tells us that $u(t)$ only converges to $\bar{f}$ strongly as $t \to \infty$. In the following, we will investigate under which conditions on the data f and the functional $J \in \mathcal{C}$ this limit is attained in finite time, i.e.,

$$T^*(f) := \inf\{t > 0 : u(t) = \bar{f}\} < \infty.$$

Furthermore, we derive lower and upper bounds of the extinction time which will depend on the (dual) norm of the data f. Scale-invariant upper bounds in the special case of total variation flow and related equations were shown in [Giga and Kohn 2011; Giga et al. 2015].

6A. ***Finite extinction time.*** First, we give a statement that can be interpreted as conservation of mass under the gradient flow.

Lemma 6.1 (conservation of mass). *Let u solve (GF). Then it holds $\overline{u(t)} = \bar{f}$ for all $t > 0$.*

Proof. Since $p(t)$ lies in $\mathcal{N}(J)^\perp$ for all $t > 0$, the same holds for $u(t) - f = -\int_0^t p(s)\,\mathrm{d}s$. Therefore, $0 = \overline{u(t) - f} = \overline{u(t)} - \bar{f}$, by the linearity of the projection. □

Theorem 6.2 (upper bound of extinction time). *Let u solve (GF). If there is $C > 0$ such that*

$$\|u(t) - \bar{f}\| \leq C J(u(t)) \quad \textit{for all } t > 0, \tag{6-1}$$

then it holds

$$T^*(f) \leq C\|f - \bar{f}\|. \tag{6-2}$$

Proof. The proof is analogous to the finite-dimensional case, treated in [Burger et al. 2016]. Using (6-1), it follows

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(t) - \bar{f}\|^2 &= \langle u(t) - \bar{f}, \partial_t u(t)\rangle = -\langle u(t) - \bar{f}, p(t)\rangle \\ &= -\langle u(t), p(t)\rangle = -J(u(t)) \leq -\frac{1}{C}\|u(t) - \bar{f}\|.\end{aligned}$$

This readily implies

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u(t) - \bar{f}\| \leq -\frac{1}{C},$$

and, integrating this equation,

$$\|u(t) - \bar{f}\| \leq \|f - \bar{f}\| - \frac{t}{C}.$$

Therefore, for $t \geq \|f - \bar{f}\|C$ it holds $u(t) = \bar{f}$, which concludes the proof. □

Remark 6.3. By Lemma 6.1, a sufficient condition for (6-1) to hold is the validity of the Poincaré-type inequality

$$\|u - \bar{u}\| \leq C J(u) \quad \text{for all } u \in \mathcal{H}. \tag{6-3}$$

Remark 6.4. Revisiting the example of a datum fulfilling $\lambda f \in \partial J(f)$, we observe that the upper bound (6-2) is sharp. It holds $u(t) = \max(1-\lambda t, 0) f$ and consequently we have

$$\frac{\|u(t)-\bar{f}\|}{J(u(t))} = \frac{\|u(t)\|}{J(u(t))} = \frac{\|f\|}{J(f)},$$

and, hence, the smallest possible constant in (6-1) is given by $C = \|f\|/J(f)$. This implies

$$C\|f-\bar{f}\| = C\|f\| = \frac{\|f\|^2}{J(f)} = \frac{1}{\lambda} = T^*(f).$$

Example 6.5. Since any $J \in \mathcal{C}$ is a norm on the subspace $\mathcal{V} = \operatorname{dom}(J) \cap \mathcal{N}(J)^\perp \subset \mathcal{H}$, the Poincaré inequality (6-3) always holds if $\mathcal{H}$ is finite-dimensional.

Example 6.6. Since for $n = 1, 2$ and $u \in \mathrm{BV}(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is a bounded domain, one has the Poincaré inequality [Ambrosio et al. 2000]

$$\|u-\bar{u}\| \le C\,\mathrm{TV}(u),$$

where $\bar{u} := |\Omega|^{-1}\int_\Omega u\,\mathrm{d}x$ and $C > 0$, the one- and two-dimensional TV-flow $\partial_t u(t) = -p(t),\ p(t) \in \partial\mathrm{TV}(u(t))$, becomes extinct in finite time for all initial conditions $u(0) = f \in L^2(\Omega)$.

Example 6.7. Since not even on bounded domains is there an embedding from $L^1(\Omega)$ to $L^2(\Omega)$, one cannot expect finite extinction time for the L^1-flow introduced in Section 5B. Indeed one has

$$u(t) = f - \int_0^t p(s)\,\mathrm{d}s,$$

where $\|p(s)\|_{L^\infty} \le 1$. This yields $u(t) \ge f - t$ almost everywhere in Ω for all $t \ge 0$. Hence, if f is an unbounded L^2-function, the extinction time is infinite. Hence, the quite strong property of a flow yielding a sequence of nonlinear eigenvectors still does not imply a finite extinction time.

We can also give a converse statement of Theorem 6.2, namely that the existence of a finite extinction time implies the validity of a Poincaré-type inequality.

Proposition 6.8. *Let u solve (GF) and assume that there exists $T > 0$ such that the solution of the gradient flow satisfies $u(t) = \bar{f}$ for all $t \ge T$. Then, one has the estimate*

$$\|u(t)-\bar{f}\| \le \sqrt{T-t}\sqrt{J(u(t))} \quad \text{for all } 0 \le t \le T. \tag{6-4}$$

In particular, for $t = 0$ and initial data satisfying $\|f-\bar{f}\| = 1$ this implies

$$\|f-\bar{f}\| \le T J(f).$$

Proof. It holds $u(t) - \bar{f} = \int_t^T p(s)\,\mathrm{d}s$, which implies by the Hölder inequality and Theorem 4.2(5)

$$\begin{aligned}\|u(t)-\bar{f}\| &\le \int_t^T \|p(s)\|\,\mathrm{d}s \le \sqrt{T-t}\sqrt{\int_t^T \|p(s)\|^2\,\mathrm{d}s}\\ &= \sqrt{T-t}\sqrt{\int_T^t \frac{\mathrm{d}}{\mathrm{d}s} J(u(s))\,\mathrm{d}s} = \sqrt{T-t}\sqrt{J(u(t))}.\end{aligned}$$

Noticing $\|f-\bar{f}\| = 1 = \|f-\bar{f}\|^2$, we obtain the assertion for $t = 0$. □

While upper bounds of the extinction time like (6-2) which depend on some Poincaré constant have already been considered in the literature [Andreu et al. 2002; Hauer and Mazón 2019], we can also give a sharp lower bound of the extinction time which is novel to the best of our knowledge.

Proposition 6.9 (lower bound of extinction time). *It holds*

$$T^*(f) \geq \|f\|_*, \tag{6-5}$$

where $\|f\|_* := \sup_{p\in\mathcal{N}(J)^\perp} \langle f, p\rangle / J(p)$ *denotes the dual norm of* f *with respect to* J.

Proof. Using $f - \bar{f} = \int_0^{T^*(f)} p(t)\,\mathrm{d}t$ it holds

$$\|f\|_* = \sup_{p\in\mathcal{N}(J)^\perp} \frac{\langle f, p\rangle}{J(p)} = \sup_{p\in\mathcal{N}(J)^\perp} \frac{1}{J(p)} \int_0^{T^*(f)} \langle p(t), p\rangle\,\mathrm{d}t \leq T^*(f),$$

where we used that due to (2-9) it holds $\langle p(t), p\rangle \leq J(p)$ for all $t > 0$. □

Remark 6.10. Since according to [Bungert and Burger 2019] the variational problem (VP) has the minimal extinction time $\|f\|_*$, we conclude that the extinction time of the gradient flow is always larger than or equal to the extinction time of the variational problem. Furthermore, this lower bound is sharp since in the case of a spectral decomposition the gradient flow and the variational problem have the same primal solution according to Theorem 4.5.

6B. *Extinction profiles: the general case.* Let us now assume that the gradient flow becomes extinct in finite time, and to simplify notation we consider the case of a datum f fulfilling $\bar{f} = 0$ without loss of generality (see Lemma 4.13). In [Andreu et al. 2002] it was shown for the special case of two-dimensional total variation flow with initial condition $f \in L^\infty(\Omega)$ that $T^*(f) < \infty$ and there is an increasing sequence of times t_n such that $t_n \to T^*(f)$ and

$$\lim_{n\to\infty} w(t_n) = p^* \tag{6-6}$$

strongly in $L^2(\Omega)$, where

$$w(t) := \begin{cases} u(t)/(T^*(f) - t), & 0 < t < T^*(f), \\ 0, & \text{else.} \end{cases} \tag{6-7}$$

Here $p^* \neq 0$ is an eigenvector and is referred to as the *extinction profile*. Note that $w(t)$ approximates the negative of the left derivative of $t \mapsto u(t)$ at $t = T^*(f)$ which — opposed to the right derivative — is not guaranteed to exist. The proof in [Andreu et al. 2002] is highly technical and uses a lot of structure that comes from the explicit form of the functional $J = \mathrm{TV}$. However, in our general framework this result can be harvested very easily and for general functionals $J \in \mathcal{C}$. First, we show that, if the strong limit (6-6) exists, it is an eigenvector.

Theorem 6.11. *Let* $u(t)$ *solve the gradient flow* (GF) *with extinction time* $T^*(f) < \infty$, *and assume that there is an increasing sequence* $t_n \to T^*(f)$ *such that* $w(t_n) \to p^*$ *as* $n \to \infty$ *holds for some* $p^* \in \mathcal{H}$. *Then* $p^* \in \partial J(p^*)$ *and if the Poincaré inequality* (6-1) *holds, one has* $p^* \neq 0$.

Proof. Let us show $p^* \in K$ first. Due to $u(T^*(f)) = 0$ and $p(t) = -\partial_t u(t)$ for $t > 0$, it holds

$$u(t_n) = \int_{t_n}^{T^*(f)} p(t)\,\mathrm{d}t,$$

and, hence, we calculate for arbitrary $v \in \mathcal{H}$

$$\left\langle \frac{u(t_n)}{T^*(f)-t_n}, v \right\rangle = \frac{1}{T^*(f)-t_n} \int_{t_n}^{T^*(f)} \langle p(t), v\rangle \, \mathrm{d}t \le J(v) \quad \text{for all } n \in \mathbb{N},$$

where we used that $p(t) \in K$ for all $t > 0$. By the closedness of K we infer that also $p^* \in K$ holds.

To show that $p^* \in \partial J(p^*)$, we calculate, using the lower semicontinuity of J,

$$J(p^*) \le \liminf_{n\to\infty} \frac{1}{T^*(f)-t_n} J(u(t_n)) = \liminf_{n\to\infty} \frac{1}{T^*(f)-t_n} \langle p(t_n), u(t_n)\rangle = \liminf_{n\to\infty} \langle p(t_n), w(t_n)\rangle.$$

Now we claim (see Lemma 6.12 below) that $\lim_{t\nearrow T^*(f)} \|p(t)-w(t)\| = 0$ which, together with the strong convergence of $w(t_n)$ to p^*, immediately implies

$$J(p^*) \le \liminf_{n\to\infty} \langle p(t_n), w(t_n)\rangle = \liminf_{n\to\infty} \langle p(t_n)-w(t_n), w(t_n)\rangle + \|w(t_n)\|^2 = \lim_{n\to\infty} \|w(t_n)\|^2 = \|p^*\|^2,$$

and, hence, p^* is an eigenvector. Here we used that $\langle p(t_n)-w(t_n), w(t_n)\rangle \le \|p(t_n)-w(t_n)\|\|w(t_n)\|$, which converges to zero since $\|w(t_n)\|$ is uniformly bounded in n.

To show that $p^* \ne 0$ in the case of (6-1), we observe that

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{2} \|u(t)\|^2 = \langle \partial_t u(t), u(t)\rangle = -\langle p(t), u(t)\rangle = -J(u(t))$$

holds, which implies with (6-1) that

$$\frac{\mathrm{d}}{\mathrm{d}t} \|u(t)\| = \frac{-J(u(t))}{\|u(t)\|} \le -\frac{1}{C}.$$

Integrating this inequality from $0 \le t < T^*(f)$ to $T^*(f)$ yields

$$-\|u(t)\| \le -\frac{1}{C}(T^*(f)-t)$$

and, hence,

$$\|w(t)\| = \frac{\|u(t)\|}{T^*(f)-t} \ge \frac{1}{C} > 0 \quad \text{for all } 0 \le t < T^*(f).$$

This implies that $p^* \ne 0$ holds. □

Lemma 6.12. *Under the conditions of Theorem 6.11 it holds $\|p(t)-w(t)\| \to 0$ as $t \nearrow T^*(f)$.*

Proof. Using the relation $\partial_t w(t) = (\partial_t u(t) + w(t))/(T^*(f)-t)$, the chain rule, and $p(t) \in \partial J(u(t))$, we obtain the estimate

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\Big(J(w(t)) - \tfrac{1}{2}\|w(t)\|^2\Big) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{J(u(t))}{T^*(f)-t} - \tfrac{1}{2}\|w(t)\|^2\right) = \frac{\langle p(t), \partial_t u(t)\rangle}{T^*(f)-t} + \frac{J(u(t))}{(T^*(f)-t)^2} - \langle w(t), \partial_t w(t)\rangle \\
&= \frac{\langle p(t), \partial_t u(t)\rangle}{T^*(f)-t} + \frac{\langle p(t), u(t)\rangle}{(T^*(f)-t)^2} - \langle w(t), \partial_t w(t)\rangle \\
&= \langle p(t), \partial_t w(t)\rangle - \langle w(t), \partial_t w(t)\rangle = -\langle \partial_t u(t) + w(t), \partial_t w(t)\rangle \\
&= -(T^*(f)-t)\|\partial_t w(t)\|^2 = -\frac{\|p(t)-w(t)\|^2}{T^*(f)-t} \le 0.
\end{aligned} \tag{6-8}$$

Since $u(t) \in \operatorname{dom}(J)$ for all $t > 0$, we can without loss of generality assume that $u(0) = f \in \operatorname{dom}(J)$, as well. This implies

$$J(w(t)) - \tfrac{1}{2}\|w(t)\|^2 \le J(w(0)) - \tfrac{1}{2}\|w(0)\|^2 \le C, \quad 0 \le t < T^*(f),$$

for a constant $C > 0$ since $w(0) = u(0)/T^*(f) = f/T^*(f)$. Integrating (6-8) yields

$$\begin{aligned}\int_s^t \frac{\|p(\tau) - w(\tau)\|^2}{T^*(f) - \tau}\,\mathrm{d}\tau &= J(w(s)) - \tfrac{1}{2}\|w(s)\|^2 - J(w(t)) + \tfrac{1}{2}\|w(t)\|^2 \\ &\le C + \tfrac{1}{2}\|w(t)\|^2 \le C', \quad 0 \le s < t < T^*(f),\end{aligned}$$

for a constant $C' > 0$. Letting t tend to $T^*(f)$, this implies

$$t \mapsto \frac{\|p(t) - w(t)\|^2}{T^*(f) - t} \in L^1(0, T^*(f)).$$

Hence, necessarily $\lim_{t \nearrow T^*(f)} \|p(t) - w(t)\| = 0$ has to hold. □

The weak limit in (6-6) always exists, as the following proposition states.

Proposition 6.13. *There exists a increasing sequence $t_n \to T^*(f)$ and $p^* \in \mathcal{H}$ such that*

$$w(t_n) \rightharpoonup p^*, \quad n \to \infty,$$

weakly in $\mathcal{H}$.

Proof. From $u(t) = \int_t^{T^*(f)} p(s)\,\mathrm{d}s$ it follows, using Theorem 4.2(2) and that $s \mapsto \|p(s)\|$ is nonincreasing,

$$\|w(t)\| = \left\| \frac{u(t)}{T^*(f) - t} \right\| \le \frac{1}{T^*(f) - t} \int_t^{T^*(f)} \|p(s)\|\,\mathrm{d}s \le \|p(t)\| \le \frac{2}{T^*(f)}\|f\|,$$

which holds for all $T^*(f)/2 < t < T^*(f)$. Hence, by the weak compactness of the closed unit ball in Hilbert spaces, there is a increasing sequence (t_n) converging to $T^*(f)$ and some $p^* \in \mathcal{H}$ such that $w(t_n) \rightharpoonup p^*$ weakly in $\mathcal{H}$. □

To ensure the existence of the strong limit one needs additional regularity, as the following corollary states.

Corollary 6.14. *Assume that there is some compactly embedded space $\mathcal{X} \Subset \mathcal{H}$ and $c > 0$ such that $\sup_{n\in\mathbb{N}} \|w(t_n)\|_{\mathcal{X}} \le c$. Then the convergence of $w(t_n)$ to p^* is strong in $\mathcal{H}$ and Theorem 6.11 is applicable.*

Using our abstract framework, we can obtain the results in [Andreu et al. 2002] for higher-dimensional total variation flow very straightforwardly.

Example 6.15 (total variation flow in arbitrary dimension). We consider the TV-flow in dimension $n \ge 3$ with bounded domain $\Omega \subset \mathbb{R}^n$. In this case, there is no general Poincaré inequality available and, hence, no finite extinction time can be expected for arbitrary initial datum $f \in L^2(\Omega)$. However, for $f \in L^\infty(\Omega)$ one can show a finite extinction time. It is well known that in this case the essential supremum of the solution $u(t)$ of the TV-flow remains uniformly bounded in t. Hence, we show that the Poincaré inequality

$$\|u - \bar{u}\|_{L^2(\Omega)} \le C\mathrm{TV}(u)$$

holds for functions $u \in \{u \in \mathrm{BV}(\Omega) \cap L^\infty(\Omega) : \|u\|_{L^\infty(\Omega)} \le c\}$. Assume there is a sequence of functions $u_k \in \mathrm{BV}(\Omega) \cap L^\infty(\Omega)$ with zero mean such that

$$\|u_k\|_{L^2(\Omega)} > k\mathrm{TV}(u_k), \quad \|u_k\|_{L^\infty(\Omega)} \le c \quad \text{for all } k \in \mathbb{N}.$$

Since $\|u_k\|_{L^2(\Omega)} \le \sqrt{|\Omega|}\|u_k\|_{L^\infty(\Omega)} \le \sqrt{|\Omega|}c < \infty$, we can set $\|u_k\|_{L^2(\Omega)} = 1$. Passing to a subsequence, we can furthermore assume that u_k converges strongly to some u in $L^1(\Omega)$. Since $\mathrm{TV}(u) \le \liminf_{k\to\infty} \mathrm{TV}(u_k) = 0$, by the lower semicontinuity of the total variation, we infer that u is in $\mathrm{BV}(\Omega)$ and is constant almost everywhere in Ω. Having zero mean, $u \equiv 0$ has to hold. Furthermore, from $0 \le \|u_k\|_{L^2(\Omega)} \le \sqrt{c\|u_k\|_{L^1(\Omega)}} \to 0$ as $k \to \infty$ we infer that $u_k \to 0$ in $L^2(\Omega)$, which is a contradiction to $\|u_k\|_{L^2(\Omega)} = 1$ for all $k \in \mathbb{N}$. Now Theorem 4.11 shows the existence of a finite extinction time.

To show that the strong limit $\lim_{n\to\infty} w(t_n)$ exists on some increasing sequence $t_n \to T^*(f)$ and is a nontrivial eigenvector of TV, it suffices to observe that the space $\mathcal{X} := L^\infty(\Omega) \cap \mathrm{BV}(\Omega)$ is compactly embedded in $\mathcal{H} := L^2(\Omega)$ and use [Andreu et al. 2002, Equation (5.4)] in order to apply Corollary 6.14 and Theorem 6.11.

Remark 6.16 (sharp convergence rate). The results of this section imply that the solution of the gradient flow $u(t)$ converges with rate $T^*(f) - t$ and if the Poincaré inequality (6-1) holds, this rate is sharp. The upper rate is established through the boundedness of $\|w(t)\|$ as shown in the proof of Proposition 6.13. The sharpness follows from the lower bound on $\|w(t)\|$ as established in the end of the proof of Theorem 6.11.

6C. *Extinction profiles: the spectral decomposition case.* Now we again specialize on the case that the gradient flow generates a sequence of eigenvectors and assume that it becomes extinct at time $T^*(f) < \infty$. In this case, we prove that extinction profiles are always eigenvectors without any compactness assumptions. In addition, since the gradient flow is equivalent to the variational problem, it holds $T^*(f) = \|f\|_*$ according to [Bungert and Burger 2019]. However, we can give another formula in terms of the extinction profile which allows us to classify all possible extinction profiles as maximizers of a certain functional.

Theorem 6.17. *Let the gradient flow (GF) generate a sequence of eigenvectors $p(t)$ for $t > 0$ and have the minimal extinction time $0 < T^*(f) < \infty$. Furthermore, let $p^* \in \mathcal{H}$ be such that $w(t_n) \rightharpoonup p^*$, where (t_n) is an increasing sequence converging to $T^*(f)$ as $n \to \infty$ (see Proposition 6.13). Then the following statements are true*:

(1) *If (6-1) holds, then $p^* \neq 0$.*

(2) *$p(t) \in \partial J(p^*)$ for all $0 < t < T^*(f)$.*

(3) *$p^* \in \partial J(p^*)$.*

(4) *$T^*(f) = \langle f, p^*\rangle / J(p^*) = \sup_{p \in \mathcal{N}(J)^\perp} \langle f, p\rangle / J(p)$.*

Proof. For concise notation we abbreviate $T := T^*(f)$.

(1) We have already seen in the proof of Theorem 6.11 that the Poincaré inequality (6-1) implies that $\|w(t)\| \ge 1/C > 0$ for all $0 \le t < T^*(f)$. Since, however, $w(t)$ converges only weakly to p^* this is not

yet sufficient to prove $p^* \neq 0$. Instead we consider the scalar product $\langle w(t), f\rangle$ and use the identities

$$w(t) = \frac{1}{T-t}\int_t^T p(s)\,\mathrm{d}s, \qquad f = \int_0^T p(r)\,\mathrm{d}r$$

to infer

$$\begin{aligned}\langle w(t), f\rangle &= \frac{1}{T-t}\int_0^T\int_t^T \langle p(s), p(r)\rangle\,\mathrm{d}s\,\mathrm{d}r \\ &= \frac{1}{T-t}\left[\int_0^t\int_t^T \langle p(s), p(r)\rangle\,\mathrm{d}s\,\mathrm{d}r + \int_t^T\int_t^T \langle p(s), p(r)\rangle\,\mathrm{d}s\,\mathrm{d}r\right] \\ &= \frac{1}{T-t}\left[\int_0^t\int_t^T \|p(s)\|^2\,\mathrm{d}s\,\mathrm{d}r + \int_t^T\int_t^T \langle p(s), p(r)\rangle\,\mathrm{d}s\,\mathrm{d}r\right], \end{aligned} \tag{6-9}$$

where we used the orthogonality of increments (see Theorem 4.7). Now we observe that by the Cauchy–Schwarz inequality

$$\|w(t)\| \le \frac{1}{T-t}\int_t^T \|p(s)\|\,\mathrm{d}s \le \frac{1}{\sqrt{T-t}}\left(\int_t^T \|p(s)\|^2\,\mathrm{d}s\right)^{\frac{1}{2}},$$

and, hence,

$$\frac{1}{T-t}\int_t^T \|p(s)\|^2\,\mathrm{d}s \ge \|w(t)\|^2 \ge \frac{1}{C^2}. \tag{6-10}$$

Furthermore, we calculate the following integral by splitting the square integration domain $[t, T]^2$ into integration over two triangles:

$$\begin{aligned}\int_t^T\int_t^T \langle p(s), p(r)\rangle\,\mathrm{d}s\,\mathrm{d}r &= \int_t^T\int_t^s \langle p(s), p(r)\rangle\,\mathrm{d}r\,\mathrm{d}s + \int_t^T\int_t^r \langle p(s), p(r)\rangle\,\mathrm{d}s\,\mathrm{d}r \\ &= \int_t^T (s-t)\|p(s)\|^2\,\mathrm{d}s + \int_t^T (r-t)\|p(r)\|^2\,\mathrm{d}r \ge 0, \end{aligned} \tag{6-11}$$

where we used Theorem 4.7 again. Combining (6-9), (6-10), and (6-11), we infer

$$\langle w(t), f\rangle \ge \frac{t}{C^2}, \qquad 0 \le t < T,$$

and, thus, $w(t_n)$ cannot converge weakly to zero for $n \to \infty$.

(2) According to Theorem 4.5 it holds $p(t) \in \partial J(u(s))$ for $t \le s$. This implies

$$\begin{aligned}J(p^*) &\le \liminf_{n\to\infty} J(w(t_n)) = \liminf_{n\to\infty} \frac{1}{T^*(f)-t_n} J(u(t_n)) = \liminf_{n\to\infty} \frac{1}{T^*(f)-t_n}\langle p(t), u(t_n)\rangle \\ &= \liminf_{n\to\infty}\langle p(t), w(t_n)\rangle = \langle p(t), p^*\rangle \quad \text{for all } 0 \le t < T^*(f),\end{aligned}$$

which shows $p(t) \in \partial J(p^*)$.

(3) The equality $J(p^*) = \langle p(t), p^*\rangle$ implies

$$J(p^*) = \frac{1}{T^*(f)-s}\int_s^{T^*(f)} \langle p(t), p^*\rangle\,\mathrm{d}t = \langle w(s), p^*\rangle \quad \text{for all } 0 \le s < T.$$

Using this with $s = t_n$ and taking the limit $n \to \infty$ yields $J(p^*) = \|p^*\|^2$ due to the weak convergence of $w(t_n)$ to p^*.

(4) We use (2) together with $f = \int_0^{T^*(f)} p(t)\,\mathrm{d}t$ to compute

$$\langle f, p^*\rangle = \int_0^{T^*(f)} \langle p(t), p^*\rangle\,\mathrm{d}t = \int_0^{T^*(f)} J(p^*)\,\mathrm{d}t = T^*(f)J(p^*),$$

which is implies $T^*(f) = \langle f, p^*\rangle / J(p^*)$. On the other hand, by Theorem 4.5 the solution $u(t)$ of the gradient flow coincides with the solution of the variational problem (VP) which has the extinction time $T^*(f) = \|f\|_* = \sup_{p\in\mathcal{N}(J)^\perp} \langle f, p\rangle / J(p)$ according to [Bungert and Burger 2019]. □

Corollary 6.18. *Under the conditions of Theorem 6.17 any extinction profile p^* is a maximizer of the functional*

$$\mathcal{N}(J)^\perp \ni p \mapsto \frac{\langle f, p\rangle}{J(p)}$$

or, equivalently,

$$\frac{f}{\|f\|_*} \in \partial J(p^*).$$

In particular, the extinction profile is uniquely determined if $f/\|f\|_*$ lies in the intersection of $\partial J(0)$ and exactly one other subdifferential.

Corollary 6.19. *Under the conditions of Theorem 6.17 any extinction profile p^* satisfies*

$$J\left(\frac{p^*}{\|p^*\|}\right) \le \frac{\|f\|}{\|f\|_*}.$$

This estimate is sharp, which can be seen by choosing f to be an eigenvector.

Example 6.20 (one-dimensional total variation flow). Since according to Section 5D, the one-dimensional total variation flow yields a spectral decomposition into eigenvectors, Theorem 6.17 and Corollary 6.18 are applicable. Bonforte and Figalli [2012] studied the case of a nonnegative initial datum $f \in \mathrm{BV}(\mathbb{R})$ which lives on the (minimally chosen) interval $[a, b]$. They showed that in this case, the extinction time is given by $T = \frac{1}{2}\int_a^b f\,\mathrm{d}x$ and that $w(t) = u(t)/(T-t)$ converges to $p^* := (2/(b-a))\chi_{[a,b]}$ for $t \to T$. This follows directly from our results after observing that $f/\|f\|_*$ can only lie in the subdifferential of a constant function on $[a, b]$, which therefore coincides with p^*. The extinction time is consequently given by

$$T = \frac{\langle f, p^*\rangle}{\mathrm{TV}(p^*)} = \frac{1}{2}\int_a^b f\,\mathrm{d}x.$$

To see that $g := f/\|f\|_*$ cannot lie in the subdifferential of a nonconstant function, let us assume that there is p such that $g \in \partial\mathrm{TV}(p)$ and $p'(x_0) \neq 0$ (in the sense of measures) for some $x_0 \in (a, b)$. Writing $g = -v'$ with $|v| \le 1$ and $v = \pm 1$, where p' is positive/negative, we find that $v(x_0) = \pm 1$ and v is nonincreasing since $g = -v'$ is nonnegative on $\mathbb{R}$. Hence, it holds $v = 1$ on $(-\infty, x_0)$ or $v = -1$ on (x_0, ∞), which is a contradiction to the fact that $g = -v'$ takes on strictly positive values on a subset in $[a, x_0]$ and $[x_0, b]$ that has positive Lebesgue measure. Otherwise, the interval $[a, b]$ would not be minimally chosen.

Acknowledgements

The work by Bungert and Burger has been supported by ERC via EU FP7-ERC Consolidator Grant 615216 LifeInverse. Burger and Chambolle acknowledge further support by the EU under grant 2020 NoMADS-DLV-777826. Novaga was partially supported by the INDAM/GNAMPA and by the University of Pisa Project PRA 2017-18.

References

[Adams and Hedberg 1996] D. R. Adams and L. I. Hedberg, *Function spaces and potential theory*, Grundlehren der Math. Wissenschaften **314**, Springer, 1996. MR Zbl

[Ambrosio et al. 2000] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Univ. Press, 2000. MR Zbl

[Andreu et al. 2002] F. Andreu, V. Caselles, J. I. Díaz, and J. M. Mazón, "Some qualitative properties for the total variation flow", *J. Funct. Anal.* **188**:2 (2002), 516–547. MR Zbl

[Bauschke and Combettes 2011] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, Springer, 2011. MR Zbl

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

[Benning and Burger 2013] M. Benning and M. Burger, "Ground states and singular vectors of convex variational regularization methods", *Methods Appl. Anal.* **20**:4 (2013), 295–334. MR Zbl

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

[Bonforte and Figalli 2012] M. Bonforte and A. Figalli, "Total variation flow and sign fast diffusion in one dimension", *J. Differential Equations* **252**:8 (2012), 4455–4480. MR Zbl

[Brézis 1971] H. Brézis, "Propriétés régularisantes de certains semi-groupes non linéaires", *Israel J. Math.* **9** (1971), 513–534. MR Zbl

[Brézis 1973] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Math. Studies **5**, North-Holland, Amsterdam, 1973. MR Zbl

[Brézis 1975] H. Brézis, "Monotone operators, nonlinear semigroups and applications", pp. 249–255 in *Proc. Int. Congress Math., II* (Vancouver, 1974), edited by R. D. James, 1975. MR Zbl

[Briani et al. 2011] A. Briani, A. Chambolle, M. Novaga, and G. Orlandi, "On the gradient flow of a one-homogeneous functional", *Confluentes Math.* **3**:4 (2011), 617–635. MR Zbl

[Bruck 1975] R. E. Bruck, Jr., "Asymptotic convergence of nonlinear contraction semigroups in Hilbert space", *J. Funct. Anal.* **18**:1 (1975), 15–26. MR Zbl

[Bungert and Burger 2019] L. Bungert and M. Burger, "Solution paths of variational regularization methods for inverse problems", *Inverse Problems* **35**:10 (2019), art. id. 105012. MR Zbl

[Burger and Osher 2004] M. Burger and S. Osher, "Convergence rates of convex variational regularization", *Inverse Problems* **20**:5 (2004), 1411–1421. MR Zbl

[Burger and Osher 2013] M. Burger and S. Osher, "A guide to the TV zoo", pp. 1–70 in *Level set and PDE based reconstruction methods in imaging* (Cetraro, Italy, 2008), edited by M. Burger and S. Osher, Lecture Notes in Math. **2090**, Springer, 2013. MR Zbl

[Burger et al. 2015] M. Burger, L. Eckardt, G. Gilboa, and M. Moeller, "Spectral representations of one-homogeneous functionals", pp. 16–27 in *Scale space and variational methods in computer vision* (Lège-Cap Ferret, France, 2015), edited by J.-F. Aujol et al., Lecture Notes in Comput. Sci. **9087**, Springer, 2015. MR

[Burger et al. 2016] M. Burger, G. Gilboa, M. Moeller, L. Eckardt, and D. Cremers, "Spectral decompositions using one-homogeneous functionals", *SIAM J. Imaging Sci.* **9**:3 (2016), 1374–1408. MR Zbl

[Caselles et al. 2011] V. Caselles, A. Chambolle, and M. Novaga, "Total variation in imaging", pp. 1–39 in *Handbook of mathematical methods in imaging*, edited by O. Scherzer, Springer, 2011. Zbl

[Chambolle et al. 2017] A. Chambolle, V. Duval, G. Peyré, and C. Poon, "Geometric properties of solutions to the total variation denoising problem", *Inverse Problems* **33**:1 (2017), art. id. 015002. MR Zbl

[Chen et al. 2009] G.-Q. Chen, M. Torres, and W. P. Ziemer, "Gauss–Green theorem for weakly differentiable vector fields, sets of finite perimeter, and balance laws", *Comm. Pure Appl. Math.* **62**:2 (2009), 242–304. MR Zbl

[Donoho 2006] D. L. Donoho, "Compressed sensing", *IEEE Trans. Inform. Theory* **52**:4 (2006), 1289–1306. MR Zbl

[Duval and Peyré 2015] V. Duval and G. Peyré, "Exact support recovery for sparse spikes deconvolution", *Found. Comput. Math.* **15**:5 (2015), 1315–1355. MR Zbl

[Fornasier and Rauhut 2011] M. Fornasier and H. Rauhut, "Compressive sensing", pp. 187–228 in *Handbook of mathematical methods in imaging*, edited by O. Scherzer, Springer, 2011. Zbl

[Giga and Kohn 2011] Y. Giga and R. V. Kohn, "Scale-invariant extinction time estimates for some singular diffusion equations", *Discrete Contin. Dyn. Syst.* **30**:2 (2011), 509–535. MR Zbl

[Giga et al. 2015] Y. Giga, H. Kuroda, and H. Matsuoka, "Fourth-order total variation flow with Dirichlet condition: characterization of evolution and extinction time estimates", preprint 1064, Hokkaido Univ., 2015, available at http://doi.org/10.14943/84208.

[Gilboa 2014] G. Gilboa, "A total variation spectral framework for scale and texture analysis", *SIAM J. Imaging Sci.* **7**:4 (2014), 1937–1961. MR Zbl

[Hauer and Mazón 2019] D. Hauer and J. M. Mazón, "Kurdyka–Łojasiewicz–Simon inequality for gradient flows in metric spaces", *Trans. Amer. Math. Soc.* **372**:7 (2019), 4917–4976. MR Zbl

[Kato 1966] T. Kato, *Perturbation theory for linear operators*, Grundlehren der Math. Wissenschaften **132**, Springer, 1966. MR Zbl

[Kōmura 1967] Y. Kōmura, "Nonlinear semi-groups in Hilbert space", *J. Math. Soc. Japan* **19**:4 (1967), 493–507. MR Zbl

[Mairal et al. 2014] J. Mairal, F. Bach, and J. Ponce, "Sparse modeling for image and vision processing", *Found. Trends Comput. Graph. Vision* **8**:2-3 (2014), 85–283. Zbl

[Meyer 2001] Y. Meyer, *Oscillating patterns in image processing and nonlinear evolution equations*, Univ. Lecture Series **22**, Amer. Math. Soc., Providence, RI, 2001. MR Zbl

[Overgaard 2018] N. C. Overgaard, "On the taut string interpretation of the one-dimensional Rudin–Osher–Fatemi model", pp. 233–244 in *Proc. Seventh Int. Conf. Pattern Recognition Appl. Methods* (Funchal, Portugal, 2018), SciTePress, Setúbal, Portugal, 2018. Zbl

[Reed and Simon 1972] M. Reed and B. Simon, *Methods of modern mathematical physics, I: Functional analysis*, Academic Press, New York, 1972. MR Zbl

[Schmidt et al. 2018] M. F. Schmidt, M. Benning, and C.-B. Schönlieb, "Inverse scale space decomposition", *Inverse Problems* **34**:4 (2018), art. id. 045008. MR Zbl

[Sohr 2001] H. Sohr, *The Navier–Stokes equations: an elementary functional analytic approach*, Birkhäuser, Basel, 2001. MR Zbl

[Weinstein 1985] M. I. Weinstein, "Modulational stability of ground states of nonlinear Schrödinger equations", *SIAM J. Math. Anal.* **16**:3 (1985), 472–491. MR Zbl

Received 9 Feb 2019. Revised 25 Sep 2019. Accepted 2 Dec 2019.

LEON BUNGERT: leon.bungert@fau.de
Department Mathematik, Universität Erlangen-Nürnberg, Erlangen, Germany

MARTIN BURGER: martin.burger@fau.de
Department Mathematik, Universität Erlangen-Nürnberg, Erlangen, Germany

ANTONIN CHAMBOLLE: antonin.chambolle@cmap.polytechnique.fr
CMAP, École Polytechnique, CNRS, Palaiseau, France

MATTEO NOVAGA: matteo.novaga@unipi.it
Dipartimento di Matematica, Università di Pisa, Pisa, Italy

Analysis & PDE

msp.org/apde

EDITORS

EDITOR-IN-CHIEF

Patrick Gérard
patrick.gerard@math.u-psud.fr
Université Paris Sud XI
Orsay, France

BOARD OF EDITORS

Massimiliano Berti	Scuola Intern. Sup. di Studi Avanzati, Italy berti@sissa.it	Werner Müller	Universität Bonn, Germany mueller@math.uni-bonn.de
Michael Christ	University of California, Berkeley, USA mchrist@math.berkeley.edu	Gilles Pisier	Texas A&M University, and Paris 6 pisier@math.tamu.edu
Charles Fefferman	Princeton University, USA cf@math.princeton.edu	Igor Rodnianski	Princeton University, USA irod@math.princeton.edu
Ursula Hamenstaedt	Universität Bonn, Germany ursula@math.uni-bonn.de	Yum-Tong Siu	Harvard University, USA siu@math.harvard.edu
Vadim Kaloshin	University of Maryland, USA vadim.kaloshin@gmail.com	Terence Tao	University of California, Los Angeles, USA tao@math.ucla.edu
Herbert Koch	Universität Bonn, Germany koch@math.uni-bonn.de	Michael E. Taylor	Univ. of North Carolina, Chapel Hill, USA met@math.unc.edu
Izabella Laba	University of British Columbia, Canada ilaba@math.ubc.ca	Gunther Uhlmann	University of Washington, USA gunther@math.washington.edu
Richard B. Melrose	Massachussets Inst. of Tech., USA rbm@math.mit.edu	András Vasy	Stanford University, USA andras@math.stanford.edu
Frank Merle	Université de Cergy-Pontoise, France Frank.Merle@u-cergy.fr	Dan Virgil Voiculescu	University of California, Berkeley, USA dvv@math.berkeley.edu
William Minicozzi II	Johns Hopkins University, USA minicozz@math.jhu.edu	Steven Zelditch	Northwestern University, USA zelditch@math.northwestern.edu
Clément Mouhot	Cambridge University, UK c.mouhot@dpmms.cam.ac.uk	Maciej Zworski	University of California, Berkeley, USA zworski@math.berkeley.edu

PRODUCTION

production@msp.org

Silvio Levy, Scientific Editor

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2021 is US $340/year for the electronic version, and $550/year (+$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

© 2021 Mathematical Sciences Publishers

ANALYSIS & PDE

Volume 14 No. 3 2021

Analysis of the linear sampling method for imaging penetrable obstacles in the time domain 667
FIORALBA CAKONI, PETER MONK and VIRGINIA SELGAS

Variations of a class of Monge–Ampère-type functionals and their applications 689
HAODI CHEN, SHIBING CHEN and QI-RUI LI

L^2-boundedness of gradients of single-layer potentials and uniform rectifiability 717
LAURA PRAT, CARMELO PULIATTI and XAVIER TOLSA

Recurrence for smooth curves in the moduli space and an application to the billiard flow on nibbled ellipses 793
KRZYSZTOF FRĄCZEK

Nonlinear spectral decompositions by gradient flows of one-homogeneous functionals 823
LEON BUNGERT, MARTIN BURGER, ANTONIN CHAMBOLLE and MATTEO NOVAGA

Restriction of toral eigenfunctions to totally geodesic submanifolds 861
XIAOQI HUANG and CHENG ZHANG

A gap theorem for α-harmonic maps between two-spheres 881
TOBIAS LAMM, ANDREA MALCHIODI and MARIO MICALLEF

Stable self-similar blow-up for a family of nonlocal transport equations 891
TAREK M. ELGINDI, TEJ-EDDINE GHOUL and NADER MASMOUDI

Potential well theory for the derivative nonlinear Schrödinger equation 909
MASAYUKI HAYASHI

Monomial convergence on ℓ_r 945
DANIEL GALICER, MARTÍN MANSILLA, SANTIAGO MURO and PABLO SEVILLA-PERIS