Nonlinear spectral decompositions by gradient flows of one-homogeneous functionals

Leon Bungert, Martin Burger, Antonin Chambolle, Matteo Novaga

Nonlinear Spectral Decompositions by Gradient Flows of One-Homogeneous Functionals

Leon Bungert* Martin Burger* Antonin Chambolle† Matteo Novaga‡

September 19, 2021

Abstract

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 both 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 semi-norms, where gradient flows can yield spectral decompositions into eigenvectors.

We provide a geometric characterization of eigenvectors via a dual unit ball and prove them to be 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 semi-norms – 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.

Keywords: Nonlinear spectral decompositions, gradient flows, nonlinear eigenvalue problems, one-homogeneous functionals, extinction profiles.

AMS Subject Classification: 35P10, 35P30, 47J10.

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 (cf. [34, 29]) or the eigenvalue problems for high-dimensional or nonlinear Schrödinger equations (cf. [37]) 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

*Department Mathematik, Universität Erlangen-Nürnberg, Cauerstrasse 11, 91058 Erlangen, Germany. {leon.bungert,martin.burger}@fau.de

†CMAP, École Polytechnique, CNRS, 91128 Palaiseau Cedex, France. antonin.chambolle@cmap.polytechnique.fr

‡Dipartimento di Matematica, Università di Pisa, Largo Pontecorvo 5, 56127 Pisa, Italy. matteo.novaga@unipi.it

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 (cf. [7, 18, 19, 22, 31]).

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

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

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

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

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 (cf. [3, 5]) and is hidden in early examples for the variational problem (cf. [32] and [6] for a detailed discussion). Secondly, for a general datum f the solution of the gradient flow satisfies

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

i.e., the datum is decomposed into subgradients $p(t)$ of the functional J. Here $\overline{f}$ denotes 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 [6] 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 (cf. [27]), 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 [16] 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 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 semi-norm 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 favourable 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 (cf. [17, 7]). 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, cf. [24, 23, 20]). 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 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 (cf. [15]). 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 [16]. We apply our theory to several examples: in particular, it matches the finite-dimensional theory for polyhedral regularizations in [16], 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 (cf. [3, 25, 8]). Our theory is general enough to allow for a direct generalizations 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 [8] 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 Brezis, 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.

2.1 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.,

$$J(cu) = |c|J(u), \quad \forall u \in \mathcal{H},\ c \in \mathbb{R} \setminus \{0\}, \tag{2.1}$$

$$J(0) = 0. \tag{2.2}$$

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

$$\mathrm{dom}(J) := \{u \in \mathcal{H} \ :\ J(u) < \infty\}, \tag{2.3}$$

$$\mathcal{N}(J) := \{u \in \mathcal{H} \ :\ J(u) = 0\}. \tag{2.4}$$

Note that $\mathrm{dom}(J)$ and $\mathcal{N}(J)$ are not empty due to (2.2) and that $\mathcal{N}(J)$ is a (strongly and weakly) closed linear space [14] whose orthogonal complement we denote by $\mathcal{N}(J)^\perp$. The effective domain $\mathrm{dom}(J)$ 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

$$\mathcal{V} = \mathrm{dom}(J) \cap \mathcal{N}(J)^\perp, \tag{2.7}$$

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 $u \in \mathrm{dom}(J)$, given by

$$\partial J(u) := \{p \in \mathcal{H} \ : \ \langle p, v\rangle \leq J(v), \ \forall v \in \mathcal{H}, \ \langle p, u\rangle = J(u)\}, \tag{2.8}$$

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

$$\partial J(0) = \{p \in \mathcal{H} \ : \ \langle p, v\rangle \leq J(v), \ \forall v \in \mathcal{H}\}, \tag{2.9}$$

which uniquely determines J as we will see. Using the definition of the dual norm (2.6), it can be easily seen that

$$\partial J(0) = \{p \in \mathcal{N}(J)^\perp \ : \ \|p\|_* \leq 1\}, \tag{2.10}$$

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 [4] for fundamental properties.

2.2 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 [16].

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

1. $J(u) \geq 0$,
2. $J(u + v) \leq J(u) + J(v)$ *and* $J(u) - J(v) \leq J(u - v)$,
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 semi-norms 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 meets $K = -K$. In this case, it holds

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

where $\overline{\text{conv}}$ 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 semi-continuous 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 meets $K = -K$ which can be seen from $\langle -p, u \rangle = \langle p, -u \rangle \leq J(-u) = J(u)$ 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 $J^* = \chi_K^{**} = \chi_{\overline{\text{conv}}(K)}$. Hence, by standard subdifferential calculus one concludes $\partial J(0) = \{p \in \mathcal{H} \,:\, J^*(p) = 0\} = \overline{\text{conv}}(K)$.

□

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

$$J\left(\int_a^b u(t)\,\mathrm{d}t\right) \leq \int_a^b J(u(t))\,\mathrm{d}t \tag{2.11}$$

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

$$J(u) = \sup_{p \in K} \langle p, u \rangle,\ u \in \mathcal{H}, \qquad K := \partial J(0), \tag{2.12}$$

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

2.3 Subgradients and eigenvectors

In this section, we will define the non-linear eigenvalue problem under consideration and provide first insights into the 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 $u \in \text{dom}(\partial J)$. 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 $\partial_{\text{rel}} K$ of K, where $\partial_{\text{rel}} K = K \setminus \text{relint}(K)$ and $\text{relint}(K) = \{p \in K \,:\, \exists c > 1 : cp \in K\}$ denotes the relative interior of K.*

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

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

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. E.g., 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 6. in Proposition 2.2.*

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 proof 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 [6] so-called ground states where 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 $\partial_{\text{rel}}K$ 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 $\partial_{\text{rel}}K$ and an eigenvector.

Proof. It has been shown in [6] 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 per definitionem 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 $q \in \partial_{\text{rel}}K$ such that $\|q\| < \|p_0\| = \lambda_0$. This implies with the Cauchy-Schwarz inequality and the definition of λ_0

$$\langle q, u \rangle \leq \|q\| \, \|u\| < \lambda_0 \, \|u\| = J(u), \quad \forall u \in \mathcal{N}(J)^\perp.$$

Since this inequality is strict, we can conclude that $q \notin \partial_{\text{rel}}K$. 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 $\partial_{\text{rel}}K$ 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 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. *$u \in \mathcal{H}$ 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 6. in Proposition 2.2. □

A key geometric characterization is provided by the following result:

[1] E.g. if J meets (6.3)

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

$$\langle p, p - q \rangle \geq 0 \qquad \forall q \in K. \tag{3.1}$$

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 one. It is obvious from the picture that any other vector on the boundary does not meet the boundary orthogonally.

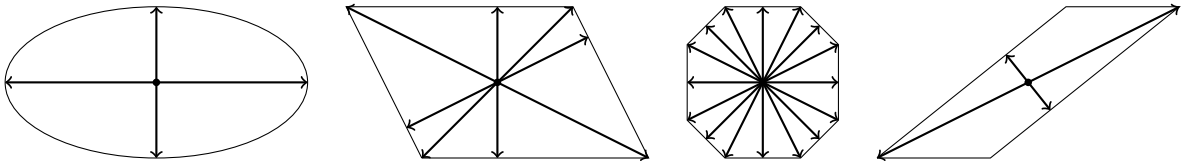

Figure 1: Four different dual unit balls K with all eigenvectors with eigenvalue one

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 semi-definite 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)**.** *Note that 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 it can be seen in Figure 1. The leftmost set corresponds to the linear case and has two different eigendirections. In contrast, in the non-linear 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 set in Figure 1 it also becomes clear that in the nonlinear case one cannot expect to have orthogonal eigenvectors as it 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*

$$\langle p, p - q \rangle = 0, \quad \forall q \in \partial J(u), \tag{3.2}$$

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

Proof. The non-negativity 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

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

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

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

which can be simplified to

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

Letting λ tend to zero and reordering shows $\langle p, p-q\rangle \leq 0$, 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 [16] 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*

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

from [16]. 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 $\langle p, p-q\rangle \geq 0$ 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 shows two polyhedrons together with a subgradient of minimal norm in the subdifferential marked in red. The left polyhedron meets (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.

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

4 Spectral decompositions by gradient flows

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

4.1 Gradient flow theory from maximal monotone evolution equations

In this section we give a concise recap of the theory of non-linear evolution equations due to Brezis [10], see also [30] 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 *non-linear* and *multi-valued* operator defined on a subset $\mathrm{dom}(A) := \{u \in \mathcal{H} \ : \ Au \neq \emptyset\}$ and is assumed to be *maximal monotone*. That is,

$$\langle p - q, u - v \rangle \geq 0, \quad \forall p \in Au, \ q \in Av, \tag{4.2}$$

and A cannot be extended to a monotone operator with larger domain (see [10] for a precise definition). Furthermore, one defines

$$A^0 u := \arg\min\{\|p\| \ : \ p \in Au\}, \ u \in \mathrm{dom}(A), \tag{4.3}$$

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

For these class of operators one has the following

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

1. *$u(t) \in \mathrm{dom}(A)$ 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*

$$\|\partial_t u\|_{L^\infty(0,\infty;\mathcal{H})} \leq \left\|A^0 f\right\| \tag{4.4}$$

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*

$$\partial_t^+ u(t) + A^0 u(t) = 0, \quad \forall t \in (0, \infty) \tag{4.5}$$

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

Proof. For the proof see [10, Theorem 3.1]. □

A important instance of maximally monotone operators are subdifferentials ∂J of lower semi-continuous 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 (Brezis, 1973)**.** *Let* $A = \partial J$ *where* $J : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ *is lower semi-continuous, 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*

$$\|\partial_t u\|_{L^\infty(\delta,\infty;\mathcal{H})} \leq \left\|A^0 v\right\| + \frac{1}{\delta} \|f - v\|, \quad \forall v \in \operatorname{dom}(A),\ \forall \delta > 0 \tag{4.6}$$

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

$$\partial_t^+ u(t) + A^0 u(t) = 0, \quad \forall t \in (0, \infty) \tag{4.7}$$

4. *The function* $t \mapsto A^0 u(t)$ *is right-continuous for all* $t > 0$ *and the function* $t \mapsto \left\|A^0 u(t)\right\|$ *is non-increasing*
5. *The function* $t \mapsto J(u(t))$ *is convex, non-increasing, Lipschitz continuous on* $[\delta, \infty)$ *for all* $\delta > 0$ *and it holds*

$$\frac{\mathrm{d}^+}{\mathrm{d}t} J(u(t)) = -\left\|\partial_t^+ u(t)\right\|^2, \quad \forall t > 0 \tag{4.8}$$

Proof. For the proof see [10, 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 non-increasing. □

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* $\frac{\mathrm{d}}{\mathrm{d}t}$ *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* $\langle p(t), p(t) - q \rangle \geq 0, \quad \forall q \in K, \ \forall t > 0.$

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.

4.2 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

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

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

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) (cf. [14] 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 5. in Theorem 4.2,

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

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 Fejer mean $q(t) := \frac{1}{t}\int_0^t p(s)\, ds$ 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) posses 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*

$$w(\tau) := v\left(1/\tau\right) - \frac{1}{\tau} v'\left(1/\tau\right), \tag{4.9}$$

$$r(\tau) := q\left(1/\tau\right), \tag{4.10}$$

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 $q'(1/\tau) = \tau^2 \left[v(1/\tau) - f - \frac{1}{\tau} v'(1/\tau)\right]$. Using this we find with the chain rule

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

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

$$J(w(\tau)) = J\left(v(1/\tau) - \frac{1}{\tau} v'(1/\tau)\right) = J\left(v(1/\tau) + \frac{1}{\tau} p(1/\tau)\right) \leq J(v(1/\tau)) + \frac{1}{\tau} J(p(1/\tau)).$$

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

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

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

4.3 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* $0 < r \leq s \leq t$ *that*

$$\langle p(t), p(s) - p(r)\rangle = 0. \tag{4.11}$$

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 $s \leq t$. Using Theorem 4.7 above one can even show

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

$$p(s) \in \partial J(p(t)). \tag{4.12}$$

Proof. Equation (4.11) implies $\|p(t)\|^2 = \langle p(s), p(t)\rangle$ for all $0 \leq s \leq t$. This yields

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

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 $0 < s_1 \leq t_1 \leq s_2 \leq t_2$.

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 a orthogonal spectral representation of the data f in the sense of [16], i.e.,

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

Here, ϕ formally act 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 4.2 and 4.3 have already been proved in a finite dimensional setting [16] and assuming that* $K \subset \mathbb{R}^n$ *is a polyhedron that meets* (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.*

4.4 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) gives formally rise to a decomposition of an arbitrary datum into subgradients which takes the form

$$f - \bar{f} = \int_0^\infty p(t)\,\mathrm{d}t, \tag{4.13}$$

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 (4.13) 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)*, with $f \in \mathrm{dom}(J)$. 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 [13, Theorem 5] for even and hence, in particular, for absolutely one-homogeneous J. To see that $u_\infty = \bar{f}$ we observe

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

by using 7. in Proposition 2.2. 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*

$$\lim_{T\to\infty} \left\| \int_0^T p(t)\,\mathrm{d}t - (f - \bar{f}) \right\| = 0, \tag{4.14}$$

which implies (4.13)*.*

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 3. and 7. in Proposition 2.2. □

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

$$f = \int_0^\infty p(t)\,\mathrm{d}t. \tag{4.15}$$

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:

$$f = \int_{\sigma(T)} \mathrm{d}E_\lambda f. \tag{4.16}$$

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

$$E_\lambda = \sum_{k=1}^\infty \delta_{\lambda_k}(\lambda)\langle \cdot, e_k\rangle e_k \tag{4.17}$$

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:

$$f = \sum_{k=1}^\infty \langle f, e_k\rangle e_k. \tag{4.18}$$

Consequently, our aim is to manipulate the measure $p(t)\,\mathrm{d}t$ in (4.15) in such a way that it becomes a non-linear generalization of (4.16)-(4.18) 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 non-increasing and converges to zero according to Theorem 4.2 and [11, 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 and let $\tilde{\mu}$ be the measure

$$\tilde{\mu}(A) = \int_A p(t)\,\mathrm{d}t, \tag{4.19}$$

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.,

$$\mu(B) := \tilde{\mu}(\lambda^{-1}B), \tag{4.20}$$

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

$$\sigma(\mu) := \{\lambda(t) \ : \ t > 0\}. \tag{4.21}$$

Remark 4.15. *Note that $t \mapsto \lambda(t)$ is indeed a measurable map since it is non-increasing according to 4. in Theorem 4.2 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 (4.16). 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 (4.17).

4.5 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 (4.12) derived in Corollary 4.8 is also sufficient. This condition generalizes the (SUB0) + orthogonality condition defined in [35], which appears to be the only sufficient condition in that line.

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

$$f = \int_0^\infty p(s)\,\mathrm{d}s, \quad p(s) \in \partial J(p(t)), \quad \forall 0 < s \leq t. \tag{4.22}$$

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

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

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

Definition 4.17 ((SUB0) + orthogonality [35])**.** We say that $f \in \mathcal{H}$ meets the (SUB0) + orthogonality condition if there are $N \in \mathbb{N}$ and positive numbers γ_i, λ_i for $i = 1, \ldots, N$ such that

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

where $p_i := \lambda_i u_i$ meet

- $p_i \in \partial J(p_i)$ for $i = 1, \ldots, N$,

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

In [35] the authors proved that the inverse scale space flow (ISS) yields a decomposition into eigenvectors if the datum meets the conditions above. In the following we prove that their condition is a special case of our necessary and sufficient condition (4.22).

Theorem 4.18. *Any datum f fulfilling the conditions from Definition 4.17 also meets* (4.22).

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 meeting $\lambda_i u_i \in \partial J(u_i)$ can be rewritten as

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

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

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

for $i = 1, \dots, N$ 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 [35, Proposition 3.4] one can easily calculate that $p(s) \in \partial J(p(t))$ holds for $s \leq t$, which is condition (4.22). □

4.6 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 semi-norm J which generalizes known results for the one-dimensional ROF problem [33].

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 [9] 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 $t > 0$.

5.1 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 [16].

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 [16] 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 1D proof first to illustrate the ideas before proving the statement in arbitrary space dimension in the subsequent section.

5.2 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 as 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

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

5.3 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 $K = \{p \in L^2(\Omega) \,:\, \|p\|_{L^1(\Omega)} \leq 1\}$ 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 $u \in \mathrm{dom}(\partial J)$ it holds

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

Note that the subgradient of minimal norm is given by

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

which follows from

$$1 = \|q\|^2_{L^1(\Omega)} \leq |\Omega_{\max}| \, \|q\|^2_{L^2(\Omega)}, \quad \forall q \in \partial J(u)$$

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 $\langle p, p - q\rangle \geq 0$ for all $q \in K$.*

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

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

□

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 $t > 0$.*

5.4 1D 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 $BV(I)$. From [12] we infer that $K = \{-g' \,:\, g \in H^1_0(I, [-1, 1])\}$ and

$$\partial J(u) = \{-g' \,:\, g \in H^1_0(I, [-1, 1]),\ g \equiv \pm 1 \text{ on } \operatorname{supp}((Du)_\pm)\} \tag{5.4}$$

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*

$$|x_+ - x_-| \geq \frac{4}{\|p\|^2_{L^2}} \tag{5.5}$$

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

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 $(y, z) \cap \operatorname{supp}(Du)_- = \emptyset$. 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 $(y, z) \cap \operatorname{supp}((Du)_+) = \emptyset$. Now let $y, z \in I$ be such that $g(y) = 1$, $g(z) = -1$ and $(y, z) \cap \operatorname{supp}(Du) = \emptyset$. Then we can define

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

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 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.*

5.5 Divergence flow

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

$$\partial_t u = -p,\ p \in \partial J(u), \tag{5.6}$$

where

$$J(u) := \int_\Omega |\operatorname{div} u| := \sup \left\{ -\int_\Omega u \cdot \nabla v \,\mathrm{d}x \ :\ v \in C_c^\infty(\Omega, [-1,1]) \right\},\ u \in L^2(\Omega, \mathbb{R}^n), \tag{5.7}$$

and $u(0) \in L^2(\Omega, \mathbb{R}^n)$. It holds that $u \in \operatorname{dom}(J)$, i.e., $J(u) < \infty$, if and only if the distribution $\operatorname{div} u$ can be represented as a finite Radon measure (cf. [12]). 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

$$K := \left\{ -\nabla v \ :\ v \in H_0^1(\Omega, [-1,1]) \right\}. \tag{5.8}$$

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 $\theta_u(x) := \frac{\operatorname{div} u}{|\operatorname{div} u|}(x)$ 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 $\mathcal{E}_u^+ \cup \mathcal{E}_u^- \subset \operatorname{supp}(\operatorname{div} u)$ and one can easily prove that $\operatorname{supp}(\operatorname{div} u) = \overline{\mathcal{E}_u^+ \cup \mathcal{E}_u^-}$. In [12] the authors showed that, in full analogy to (5.9), the subdifferential of J in $u \in \operatorname{dom}(J)$ can be characterized as

$$\partial J(u) = \left\{ -\nabla v \ : \ v \in H_0^1(\Omega, [-1,1]),\, v = \pm 1 \ |\operatorname{div} u|\text{-a.e. on } \mathcal{E}_u^\pm \right\}. \tag{5.11}$$

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 (cf.[1, Ch. 6] and [21], 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*

$$\psi_\varepsilon(t) = \begin{cases} 0, & |t| > 1, \\ 1, & |t| < 1 - \varepsilon, \\ \frac{1}{\varepsilon}(1 - |t|), & 1 - \varepsilon \leq |t| \leq 1, \end{cases} \qquad \phi_\varepsilon(t) = \int_0^t \psi_\varepsilon(\tau) \,\mathrm{d}\tau. \tag{5.13}$$

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

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

□

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:

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

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 direction of the test function vanishes consequently.

Strongly approximating $v - w$ by smooth and compactly supported functions in $H_0^1(\Omega)$ and using 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 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.*

5.6 Rotation flow

Now we fix the dimension $n = 2$ and consider the functional

$$\tilde{J}(u) := \int_\Omega |\operatorname{rot} u| := \sup \left\{ \int_\Omega u \cdot \nabla^\perp v \, \mathrm{d}x \ : \ v \in C_c^\infty(\Omega, [-1,1]) \right\}, \ u \in L^2(\Omega, \mathbb{R}^n), \tag{5.17}$$

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 $R = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ 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 $\tilde{J} = \chi_{\tilde{K}}^*$ where $\tilde{K} := \{\nabla^\perp v \ : \ v \in H_0^1(\Omega, [-1,1])\}$ and it holds $\tilde{K} = RK$. 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* $t > 0$*.*

5.7 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 denote $\mathcal{K} := \partial \mathcal{J}(0)$.

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*

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

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* Helmholtz projection *and denoted by* Π_σ*.*

Proof. The proof can be found in [36, Lemma 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

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

In particular, one has

$$\begin{aligned}\langle (1-\Pi_\sigma)p, v\rangle &= \langle p, (1-\Pi_\sigma)v\rangle \leq J((1-\Pi_\sigma)v) = J(v),\\ \langle \Pi_\sigma p, v\rangle &= \langle p, \Pi_\sigma v\rangle \leq \tilde{J}(\Pi_\sigma v) = \tilde{J}(v),\end{aligned}$$

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

$$\begin{aligned}\langle (1-\Pi_\sigma)p, u\rangle &\leq J(u),\\ \langle \Pi_\sigma p, u\rangle &\leq \tilde{J}(u).\end{aligned}$$

If we assumed that one of the inequalities was strict, then using $\langle p, u\rangle = \mathcal{J}(u)$ implied

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

which is a contradiction. Therefore, we have

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

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

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

□

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. □

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. Ad 1.: According to Proposition 5.18 this holds for general subgradients.

Ad 2.: We only proof $p_J \in \mathcal{N}(\tilde{J})$, 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 $\tilde{J}(p_J) = 0$.

Ad 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 3. in Proposition 2.2. □

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 Thms 5.15 and 5.17, each subgradient of minimal norm is an eigenvector of ∂J and $\partial \tilde{J}$, respectively, and by 3. in Lemma 5.21 we conclude that their sum is an eigenvector of $\partial \mathcal{J}$. The uniqueness of the Helmholtz decomposition concludes the proof. □

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 (cf. [6, 14], 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 [25, 26].

6.1 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*

$$\left\|u(t) - \bar{f}\right\| \leq CJ(u(t)), \quad \forall t > 0, \tag{6.1}$$

then it holds

$$T^*(f) \leq C\left\|f - \bar{f}\right\| \tag{6.2}$$

Proof. The proof is analogous to the finite dimensional case, treated in [16]. Using (6.1), it follows

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

This readily implies

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

and, integrating this equation,

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

Therefore, for $t \geq \left\|f - \bar{f}\right\| C$ 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*

$$\|u - \bar{u}\| \leq CJ(u), \quad \forall u \in \mathcal{H}. \tag{6.3}$$

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*

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

and, hence, the smallest possible constant in (6.1) *is given by $C = \frac{\|f\|}{J(f)}$. This implies that*

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

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 $u \in BV(\Omega)$ where $\Omega \subset \mathbb{R}^n$ is a bounded domain one has the Poincaré inequality [2]

$$\|u - \bar{u}\| \leq C \operatorname{TV}(u),$$

where $\bar{u} := |\Omega|^{-1} \int_\Omega u \, \mathrm{d}x$ and $C > 0$, the one and two-dimensional TV-flow $\partial_t u(t) = -p(t),\ p(t) \in \partial \operatorname{TV}(u(t))$ 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 there is an embedding from $L^1(\Omega)$ to $L^2(\Omega)$, one cannot expect finite extinction time for the L^1-flow introduced in Section 5.2. Indeed one has

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

where $\|p(s)\|_{L^\infty} \leq 1$. This yields $u(t) \geq f - t$ almost everywhere in Ω for all $t \geq 0$. 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 non-linear 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 solutions of the gradient flow satisfies $u(t) = \bar{f}$ for all $t \geq T$. Then, one has the estimate*

$$\left\|u(t) - \bar{f}\right\| \leq \sqrt{T - t} \sqrt{J(u(t))}, \quad \forall 0 \leq t \leq T. \tag{6.4}$$

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

$$\left\|f - \bar{f}\right\| \leq T J(f).$$

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

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

Noticing $\left\|f - \bar{f}\right\| = 1 = \left\|f - \bar{f}\right\|^2$ 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 [3, 28], 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 $\|f\|_* := \sup_{p \in \mathcal{N}(J)^\perp} \frac{\langle f,p\rangle}{J(p)}$ *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 [14] the variational problem* (VP) *has the minimal extinction time* $\|f\|_*$*, we conclude that the extinction time of the gradient flow is always larger or equal 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.*

6.2 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 (cf. Lemma 4.13). In [3] it was shown for the special case of 2D 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

$$w(t) := \begin{cases} \frac{u(t)}{T^*(f)-t}, & 0 < t < T^*(f), \\ 0, & \text{else.} \end{cases} \tag{6.7}$$

Here $p^* \neq 0$ is an eigenvector and is referred to as *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 [3] is highly technical and uses a lot of structure that comes from the explicit form of the functional $J = TV$. 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 $v \in \mathcal{H}$

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

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 lower semi-continuity of J:

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

Now we claim (cf. 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

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

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 $p^* \neq 0$ in 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,

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

This implies that $p^* \neq 0$ 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

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(J(w(t)) - \frac{1}{2}\|w(t)\|^2\right) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{J(u(t))}{T^*(f)-t} - \frac{1}{2}\|w(t)\|^2\right) \\
&= \frac{\langle p(t), \partial_t u(t)\rangle}{T^*(f)-t} + \frac{J(u(t))}{(T^*(f)-t)^2} - \langle w(t), \partial_t w(t)\rangle \\
&= \frac{\langle p(t), \partial_t u(t)\rangle}{T^*(f)-t} + \frac{\langle p(t), u(t)\rangle}{(T^*(f)-t)^2} - \langle w(t), \partial_t w(t)\rangle \\
&= \langle p(t), \partial_t w(t)\rangle - \langle w(t), \partial_t w(t)\rangle = -\langle \partial_t u(t) + w(t), \partial_t w(t)\rangle \\
&= -(T^*(f)-t)\,\|\partial_t w(t)\|^2 = -\frac{\|p(t)-w(t)\|^2}{T^*(f)-t} \leq 0
\end{aligned} \tag{6.8}$$

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

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

for a constant $C > 0$ since $w(0) = u(0)/T^*(f) = f/T^*(f)$. Integrating (6.8) yields

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

for a constant $C' > 0$. Letting t tend to $T^*(f)$, this implies that $t \mapsto \frac{\|p(t)-w(t)\|^2}{T^*(f)-t} \in L^1(0, T^*(f))$. 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*

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

weakly in $\mathcal{H}$.

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

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

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 $\sup_{n\in\mathbb{N}} \|w(t_n)\|_{\mathcal{X}} \leq c$. 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 [3] for more-dimensional total variation flow very straightforwardly.

Example 6.15 (Total variation flow in arbitrary dimension)**.** We consider the TV-flow in dimension $n \geq 3$ 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

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

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

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

Since $\|u_k\|_{L^2(\Omega)} \leq \sqrt{|\Omega|}\,\|u_k\|_{L^\infty(\Omega)} \leq \sqrt{|\Omega|}c < \infty$, 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 $\mathrm{TV}(u) \leq \liminf_{k\to\infty} \mathrm{TV}(u_k) = 0$, by the lower semi-continuity of the total variation, we infer that u is in $BV(\Omega)$ and is constant almost everywhere in Ω. Having zero mean, $u \equiv 0$ has to hold. Furthermore, from $0 \leq \|u_k\|_{L^2(\Omega)} \leq \sqrt{c\,\|u_k\|_{L^1(\Omega)}} \to 0$ 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 non-trivial eigenvector of TV, it suffices to observe that the space $\mathcal{X} := L^\infty(\Omega) \cap BV(\Omega)$ is compactly embedded in $\mathcal{H} := L^2(\Omega)$ and use [3, Eq. (5.4)] in order to apply Corollary 6.14 and Theorem 6.11.

Remark 6.16 (Sharp convergence rate)**.** *We would like to remark that 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.*

6.3 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 [14]. 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 $n \to \infty$ (cf. 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. *$T^*(f) = \frac{\langle f, p^* \rangle}{J(p^*)} = \sup_{p \in \mathcal{N}(J)^\perp} \frac{\langle f, p \rangle}{J(p)}$.*

Proof. For a concise notation we abbreviate $T := T^*(f)$. Ad 1.: We have already seen in the proof of Theorem 6.11 that the Poincaré inequality (6.1) implies that $\|w(t)\| \geq \frac{1}{C} > 0$ for all $0 \leq t < T^*(f)$. 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

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

to infer

$$\begin{aligned}
\langle w(t), f \rangle &= \frac{1}{T-t} \int_0^T \int_t^T \langle p(s), p(r) \rangle \, \mathrm{d}s \, \mathrm{d}r \\
&= \frac{1}{T-t} \left[\int_0^t \int_t^T \langle p(s), p(r) \rangle \, \mathrm{d}s \, \mathrm{d}r + \int_t^T \int_t^T \langle p(s), p(r) \rangle \, \mathrm{d}s \, \mathrm{d}r \right] \\
&= \frac{1}{T-t} \left[\int_0^t \int_t^T \|p(s)\|^2 \, \mathrm{d}s \, \mathrm{d}r + \int_t^T \int_t^T \langle p(s), p(r) \rangle \, \mathrm{d}s \, \mathrm{d}r \right], \qquad (6.9)
\end{aligned}$$

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

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

and, hence,

$$\frac{1}{T-t} \int_t^T \|p(s)\|^2 \, \mathrm{d}s \geq \|w(t)\|^2 \geq \frac{1}{C^2}. \qquad (6.10)$$

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

$$\begin{aligned}
\int_t^T \int_t^T \langle p(s), p(r) \rangle \, \mathrm{d}s \, \mathrm{d}r &= \int_t^T \int_t^s \langle p(s), p(r) \rangle \, \mathrm{d}r \, \mathrm{d}s + \int_t^T \int_t^r \langle p(s), p(r) \rangle \, \mathrm{d}s \, \mathrm{d}r \\
&= \int_t^T (s-t) \|p(s)\|^2 \, \mathrm{d}s + \int_t^T (r-t) \|p(r)\|^2 \, \mathrm{d}r \\
&\geq 0, \qquad (6.11)
\end{aligned}$$

where we used Theorem 4.7 again. Combining (6.9), (6.10), and (6.11), we infer

$$\langle w(t), f\rangle \geq \frac{t}{C^2}, \quad 0 \leq t < T,$$

and, thus, $w(t_n)$ cannot converge weakly to zero for $n \to \infty$.
Ad 2.: According to Theorem 4.5 it holds $p(t) \in \partial J(u(s))$ for $t \leq s$. This implies

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

which shows $p(t) \in \partial J(p^*)$.
Ad 3.: The equality $J(p^*) = \langle p(t), p^*\rangle$ implies

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

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^*.
Ad 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 [14]. □

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^* meets*

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

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 5.4, the one-dimensional total variation flow yields a spectral decomposition into eigenvectors, Theorem 6.17 and Corollary 6.18 are applicable. Bonforte and Figalli in [8] studied the case of a non-negative initial datum $f \in BV(\mathbb{R})$ 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 $p^* := \frac{2}{b-a}\chi_{[a,b]}$ 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 non-constant 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 $|v| \leq 1$ and $v = \pm 1$ where p' is positive / negative, we find that $v(x_0) = \pm 1$ and v is non-increasing since $g = -v'$ is non-negative 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 LB and MB has been supported by ERC via Grant EU FP7 – ERC Consolidator Grant 615216 LifeInverse. MB and AC acknowledge further support by the EU under grant 2020 NoMADS - DLV-777826. MN was partially supported by the INDAM/GNAMPA and by the University of Pisa Project PRA 2017-18.

References

[1] D. R. Adams and L. I. Hedberg. *Function spaces and potential theory*, volume 314. Springer Science & Business Media, 2012.

[2] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*, volume 254. Clarendon Press Oxford, 2000.

[3] F. Andreu, V. Caselles, J. Diaz, and J. M. Mazón. Some qualitative properties for the total variation flow. *Journal of functional analysis*, 188(2):516–547, 2002.

[4] H. H. Bauschke, P. L. Combettes, et al. *Convex analysis and monotone operator theory in Hilbert spaces*, volume 408. Springer, 2011.

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

[6] M. Benning and M. Burger. Ground states and singular vectors of convex variational regularization methods. *Methods and Applications of Analysis*, 20(4):295–334, 2013.

[7] M. Benning and M. Burger. Modern regularization methods for inverse problems. *Acta Numerica*, 27:1–111, 2018.

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

[9] H. Brézis. Propriétés régularisantes de certains semi-groupes non linéaires. *Israel Journal of Mathematics*, 9(4):513–534, 1971.

[10] H. Brezis. *Operateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, volume 5. Elsevier, 1973.

[11] H. Brezis. Monotone operators, nonlinear semigroups and applications. In *Proceedings of the International Congress of Mathematicians (Vancouver, BC, 1974)*, volume 2, pages 249–255, 1974.

[12] A. Briani, A. Chambolle, M. Novaga, and G. Orlandi. On the gradient flow of a one-homogeneous functional. *Confluentes Mathematici*, 3(04):617–635, 2011.

[13] R. E. Bruck Jr. Asymptotic convergence of nonlinear contraction semigroups in hilbert space. *Journal of Functional Analysis*, 18(1):15–26, 1975.

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

[15] M. Burger, L. Eckardt, G. Gilboa, and M. Moeller. Spectral representations of one-homogeneous functionals. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 16–27. Springer, 2015.

[16] M. Burger, G. Gilboa, M. Moeller, L. Eckardt, and D. Cremers. Spectral decompositions using one-homogeneous functionals. *SIAM Journal on Imaging Sciences*, 9(3):1374–1408, 2016.

[17] M. Burger and S. Osher. Convergence rates of convex variational regularization. *Inverse problems*, 20(5):1411, 2004.

[18] M. Burger and S. Osher. A guide to the tv zoo. In *Level set and PDE based reconstruction methods in imaging*, pages 1–70. Springer, 2013.

[19] V. Caselles, A. Chambolle, and M. Novaga. Total variation in imaging. *Handbook of Mathematical Methods in Imaging*, pages 1–39, 2011.

[20] A. Chambolle, V. Duval, G. Peyré, and C. Poon. Geometric properties of solutions to the total variation denoising problem. *Inverse Problems*, 33(1):015002, 2016.

[21] G.-Q. Chen, W. P. Ziemer, and M. Torres. Gauss-green theorem for weakly differentiable vector fields, sets of finite perimeter, and balance laws. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 62(2):242–304, 2009.

[22] D. L. Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.

[23] V. Duval and G. Peyré. Exact support recovery for sparse spikes deconvolution. *Foundations of Computational Mathematics*, 15(5):1315–1355, 2015.

[24] M. Fornasier and H. Rauhut. Compressive sensing. In *Handbook of mathematical methods in imaging*, pages 187–228. Springer, 2011.

[25] Y. Giga and R. V. Kohn. Scale-invariant extinction time estimates for some singular diffusion equations. *Discrete Continuous Dynam. Systems*, pages 509–535, 2011.

[26] Y. Giga, H. Kuroda, and H. Matsuoka. Fourth-order total variation flow with dirichlet condition: characterization of evolution and extinction time estimates. *Hokkaido University Preprint Series in Mathematics*, 1064:1–36, 2015.

[27] G. Gilboa. A total variation spectral framework for scale and texture analysis. *SIAM journal on Imaging Sciences*, 7(4):1937–1961, 2014.

[28] D. Hauer and J. Mazón. Kurdyka-{\L} ojasiewicz-simon inequality for gradient flows in metric spaces. *arXiv preprint arXiv:1707.03129*, 2017.

[29] T. Kato. *Perturbation theory for linear operators*, volume 132. Springer Science & Business Media, 2013.

[30] Y. Komura. Nonlinear semi-groups in hilbert space. *Journal of the Mathematical Society of Japan*, 19(4):493–507, 1967.

[31] J. Mairal, F. Bach, J. Ponce, et al. Sparse modeling for image and vision processing. *Foundations and Trends® in Computer Graphics and Vision*, 8(2-3):85–283, 2014.

[32] Y. Meyer. *Oscillating patterns in image processing and nonlinear evolution equations: the fifteenth Dean Jacqueline B. Lewis memorial lectures*, volume 22. American Mathematical Soc., 2001.

[33] N. C. Overgaard. On the taut string interpretation of the one-dimensional rudin-osher-fatemi model: A new proof, a fundamental estimate and some applications. *arXiv preprint arXiv:1710.10985*, 2017.

[34] M. Reed and B. Simon. *Methods of Modern Mathematical Physics: Functional Analysis.-1972.-(RU-idnr: M103448034)*. Academic Press, 1972.

[35] M. F. Schmidt, M. Benning, and C.-B. Schönlieb. Inverse scale space decomposition. *Inverse Problems*, 34(4):045008, 2018.

[36] H. Sohr. *The Navier-Stokes equations: An elementary functional analytic approach*. Springer Science & Business Media, 2012.

[37] M. I. Weinstein. Modulational stability of ground states of nonlinear schrödinger equations. *SIAM journal on mathematical analysis*, 16(3):472–491, 1985.