Nonlinear Spectral Duality

Francesco Tudisco Dong Zhang

Abstract

Nonlinear eigenvalue problems for pairs of homogeneous convex functions are a generalization of standard matrix eigenvalue problems based on non-quadratic and non-differentiable generalized Rayleigh quotients. These types of eigenproblems are encountered in diverse fields, including graph mining, machine learning, and network science. By considering different notions of duality transforms from both classical and recent convex geometry theory, in this work we show that one can move from the primal to the dual nonlinear eigenvalue formulation maintaining the spectrum, the variational spectrum, as well as the corresponding multiplicities unchanged. These nonlinear spectral duality properties can be used to transform the original eigenvalue problem into various alternative and possibly more treatable dual problems. We illustrate the use of nonlinear spectral duality in a variety of example settings involving optimization problems on graphs, nonlinear Laplacians, and distances between convex bodies. For all these settings we obtain new characterizations and results.

2020 Mathematics Subject Classification: 47J10, 49N15, 05C50, 90C46, 52A41

1 Introduction and motivation

A wide variety of nonlinear eigenvalue problems can be formulated as critical point conditions for Rayleigh-type quotients $r(x) = f(x)/g(x)$ involving two real-valued homogeneous functions f and g. The homogeneity of f and g ensures that r is scale-invariant. Thus, one easily verifies that the critical points of r are defined only up to scaling and correspond to a nonlinear notion of eigenvectors, similar to the standard linear case, with eigenvalues given by the corresponding critical values. For this reason, we call the critical values and critical points of the ratio r of convex homogeneous functions f and g, the *nonlinear spectrum of the function pair* (f, g), and refer to this concept as *eigenvalue problem for a function pair.*

One of the most celebrated examples of eigenvalue problem for a function pair is the p-Laplacian eigenvalue problem. This nonlinear eigenvalue problem is related to the solution of a nonlinear divergence-form PDE and has been extensively studied in the last few decades [10,19,37,40,50] also due to its central role in imaging and machine learning [7,8,13,21,24,44,48]. Alongside the p-Laplacian, eigenvalue problems for function pairs appear in a range of settings, including the study of operator norms for linear and multilinear forms [25,27,43]; the solution of the Gross-Pitaevskii equation in quantum chemistry [12, 46, 55]; the identification and analysis of relevant mesoscopic structures in complex networks, such as central nodes, communities and core-periphery [6, 33, 51, 52, 54]; the optimization of polynomials and generalized polynomials on the unit sphere [26, 27, 58].

A number of complications arise when moving from the classical linear eigenvalue problem to the nonlinear one, starting from the fact that the number of eigenvalues and eigenvectors is no longer bounded by the space dimension. However, in most cases, one can use the Lusternik-Schnirelmann theory combined with the Krasnoselski genus to define a sequence of variational eigenvalues by means of a Courant-Fisher-like min-max characterization. This subset of variational eigenvalues has very useful properties in most application settings. However, unlike the linear case, evaluating, computing, or approximating the variational eigenvalues is in general a very challenging problem in the nonlinear case, which boils down to a nonsmooth optimization problem for pairs of homogeneous convex functions.

Another fundamental property of linear eigenvalue problems that is much more difficult to show and analyze in the nonlinear case is the spectral invariance under duality. For example, it is well-known that for a real linear operator $A : V \to W$, AA^* and A^*A share the same (nonzero) eigenvalues as well as their multiplicities. The two mappings AA^* and A^*A are related by a notion of duality inherited from the Rayleigh quotients that define their eigenvalues. In fact, the eigenvalues of $(A^*A)^{1/2}$ are critical values of $\|x\|_A/\|x\|$ while those of $(AA^*)^{1/2}$ correspond to the critical values of $\|x\|_A^*/\|x\|$, where $\|\cdot\|_A$ denotes the pseudo-norm $\|A\cdot\|$, and $\|\cdot\|_A^*$ its dual.

Duality is a fundamental concept in algebra, analysis, geometry, as well as mathematical optimization, and computational mathematics. However, while duality is widely used to e.g. connect critical equations of different optimization problems, the effects of duality transformations on the spectral properties associated with Rayleigh-like quotients of homogeneous functions are not well understood and studied. In this paper, we focus on the family of function pairs (f, g) that, on top of being homogeneous and convex, are nonnegative and thus have a linear kernel. These properties are very common in a range of applications, as we will further detail in Section 8.

For these types of functions, we define three duality transforms obtained by adapting three fundamental notions of duality from convex geometry and convex analysis: the norm duality, the Fenchel's convex conjugate (i.e., Legendre transform) and the polarity transform (or $\mathcal{P}$-transform) [1,2,45]. Thus, we provide three main results showing that the variational spectrum as well as its multiplicities are invariant under these duality transforms.

In particular, by applying our main results, Theorem 4.1 and Corollary 4.2, to the norms case (i.e., the case $f = \|\cdot\|_\alpha$ and $g = \|\cdot\|_\beta$ are generic norms), we obtain the following nontrivial generalization of the fact that AA^* and A^*A share the same nonzero spectrum:

Suppose $(X_\alpha, \|\cdot\|_\alpha)$ and $(X_\beta, \|\cdot\|_\beta)$ are two normed spaces with dual norms $\|\cdot\|_\alpha^*$ and $\|\cdot\|_\beta^*$ respectively. Let $A : X_\beta \to X_\alpha$ be a linear map and let ∂ denote the subgradient operator. Then, the nonzero eigenvalues of the two eigenproblems

$$0 \in \partial_x \|Ax\|_\alpha - \lambda \partial \|x\|_\beta \qquad \text{and} \qquad 0 \in \partial_y \|A^* y\|_\beta^* - \lambda \partial \|y\|_\alpha^*$$

coincide, and their multiplicities (resp., variational multiplicities) coincide as well. Moreover, their eigenvectors are related as follows: if (λ, x) is a nontrivial eigenpair of the primal eigenproblem $0 \in \partial_x \|Ax\|_\alpha - \lambda \partial \|x\|_\beta$, then any vector $y \in \partial \|Ax\|_\alpha \cap (A^*)^{-1}\big(\lambda \partial \|x\|_\beta\big)$ is an eigenvector of the dual eigenproblem $0 \in \partial_y \|A^* y\|_\beta^* - \lambda \partial \|y\|_\alpha^*$. Such a result is new even in the standard differentiable l^p-norm case and is shown to be optimal given the eigenvalues and their multiplicities.

Our approach is mostly theoretical, but our results can help guide computational aspects. In fact, they allow us to move from a given nonlinear eigenvalue problem to several new dual problems, which, depending on the particular setting, may result in a more treatable optimization problem or may reveal useful properties that are difficult to observe and prove using the primal eigenvalue formulation. For example, if $f = g$ are L^p-norms, convergence guarantees for the fixed point iteration method to compute $\max f(x)/g(x)$ may be obtained using the dual pair, while the same method may fail to converge for the primal problem [25]. Similarly, a variety of established algorithms for nonlinear eigenproblems such as the inverse iterations [34, 35], the family of RatioDCA methods [31, 54], the MBO energy landscape and active set search methods for graph total variation [6, 17, 33], or the continuous gradient-flow approach [9, 22], can be directly transferred to the dual eigenvalue equations. The resulting dual iteration or dual flow can be used to solve the optimization of the primal eigenvalue problem and may behave better in practice. Several more specific application settings where nonlinear spectral duality may be used are illustrated in Section 8. Some of the example settings discussed there contain new results we obtain as a consequence of our spectral duality theory.

Our **main contributions** are as follows:

- We develop a novel analytical framework solely based on convex and nonlinear analysis to establish various dual formulations of nonlinear eigenvalue problems (see Theorems 4.1, 5.1, 5.2 and 5.7). In particular, application of our framework to the l^p-norm case leads to the new spectral duality of the graph p-Laplacian, which has been only partially studied so far, see Subsections 8.1.1, 8.1.2, 8.1.3, and 8.1.4.
- Since the proofs for the linear case and the maximum eigenvalue in the nonlinear case do not apply to our nonlinear spectrum case, we develop an approach to estimate the Krasnoselskii genus of nonlinear eigenspaces in the nonsmooth case based on the Moreau-Yosida approximation and the partitions of unity, see Lemmas 4.3, 4.4 and 4.5 and the proof of Theorem 4.1.
- We fully characterize the relationship between the Krasnoselskii variational spectrum of a generic function pair and its several dual formulations, including their ordering and multiplicities, see Theorem 4.1; these include new techniques to overcome significant mathematical challenges in proving the nonlinear spectral invariance under duality, see Lemma 4.6.
- We highlight important consequences of the novel spectral duality invariance results to spectral theory for graphs and hypergraphs. Most of the obtained results on graphs and hypergraphs were not previously known, see Theorems 8.2 and 8.3.

Before entering the technical development, we briefly and somewhat informally state two representative results, postponing more general and precise formulations to later sections.

Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map, and let $f : \mathbb{R}^m \to [0, +\infty)$ and $g : \mathbb{R}^n \to [0, +\infty)$ be positively one-homogeneous convex functions such that $f(y) = 0$ if and only if $y = 0$, and $g(x) = 0$ if and only if $x = 0$. Consider the nonlinear eigenvalue problem

$$0 \in \partial f(Ax) - \lambda\, \partial g(x), \tag{1.1}$$

and its dual counterpart

$$0 \in \partial g^*(A^\top y) - \lambda\, \partial f^*(y), \tag{1.2}$$

where ∂ denotes the subdifferential (see Definition 2.1), and f^*, g^* are the standard norm duals of f and g (see Section 3). Our first main result asserts that the nonzero eigenvalues of

(1.1) coincide exactly with the nonzero eigenvalues of (1.2). Full statements may be found in Theorem 4.1 and Corollary 4.2.

Another key result is actually an application of the above theorem. Roughly speaking, we prove that the nonzero normalized eigenvalues of p-Laplacian on a hypergraph coincide with the nonzero normalized eigenvalues of p^*-Laplacian on the dual hypergraph, where p and p^* are Hölder conjugate, see Theorem 8.7.

The rest of the paper is structured as follows. In Section 2 we introduce the class of functions of interest and the associated notions of spectrum and variational spectrum. In Section 3 we introduce the notion of norm-like dual for the class of one-homogeneous functions of interest and we review several preliminary properties for this duality operator, including the concept of infimal postcomposition from convex analysis. In Section 4 we present our main results, showing the spectral invariance for one-homogeneous functions under norm-like duality. In Section 5 we then move on to the class of p-homogeneous functions, for $p \geq 1$. We introduce the Legendre and polarity duality mappings and we extend the nonlinear spectral duality theorem to these two alternative notions of duality. While we focus our analysis on the case of functions and operators acting on finite-dimensional real vector spaces, a number of our results transfer directly to the infinite-dimensional setting. We devote a brief Section 7 to discuss this setting. Finally, in Section 8 we illustrate a number of examples and problems from graph theory, network science, and convex geometry, where the new spectral duality theory can be used to provide new insight, including important applications to the spectral theory of (hyper-)graph p-Laplacians.

We emphasise that throughout the manuscript, we primarily work under homogeneity assumptions. In particular, the norm duality operator inherently requires functions that are one-homogeneous. By contrast, the Legendre dual and polarity dual are defined without any homogeneity requirement. This suggests that some of the results developed here, especially those based on Legendre or polarity duality, may potentially be extended to more general non-homogeneous functions.

Such an extension, however, would necessitate overcoming substantial additional technical difficulties arising from the loss of homogeneity. We therefore leave a systematic treatment of the non-homogeneous case as an open direction for future research.

Acknowledgments. The authors are grateful to an anonymous referee for comments and suggestions, which greatly helped us improve the quality of the presentation of our paper. Dong Zhang is supported by grants from the National Natural Science Foundation of China (No. 12401443). FT is partially funded by the PRIN-MUR project MOLE (code 2022ZK5ME7) and the PRIN-PNRR project FIN4GEO (code P2022BNB97).

2 Convex homogeneous functions and their spectrum

Consider two real-valued functions $f, g : \mathbb{R}^n \to \mathbb{R}$ and suppose they are differentiable. The critical points and critical values of the ratio $r(x) = f(x)/g(x)$, i.e. the pairs (λ, x^*) such that $\nabla r(x^*) = 0$ and $r(x^*) = \lambda$, define what we call the (nonlinear) spectrum of the function pair (f, g). This is because, $\nabla r(x^*) = 0$ if and only if x^* is such that

$$\nabla f(x^*) = \lambda \nabla g(x^*) \, .$$

This definition still makes sense without the differentiability assumption. In that case, we can consider Clarke's sub-differential ∂ to show that if $0 \in \partial r(x^*)$ then $0 \in \partial f(x^*) - \lambda \partial g(x^*)$.

However, the reverse implication is in general not true without assuming the functions to be differentiable. Overall, we have

Definition 2.1. *Given locally Lipschitz functions $f, g : \mathbb{R}^n \to \mathbb{R}$, we call (λ, x) an eigenpair for the function pair (f, g) if*

$$0 \in \partial f(x) - \lambda \partial g(x)$$

where ∂ denotes Clarke's generalized derivative [15], defined as

$$\partial f(x) := \Big\{ \xi \in \mathbb{R}^n \;\Big|\; \langle \xi, v\rangle \leq \lim_{y\to x, h\downarrow 0} \frac{f(y+hv)-f(y)}{h}\,,\; \forall v \in \mathbb{R}^n \Big\}.$$

If f is further assumed to be convex, then the following equivalent definition holds

$$\partial f(x) = \{\xi \in \mathbb{R}^n \mid \langle \xi, v\rangle \leq f(x+v) - f(x)\,,\; \forall v \in \mathbb{R}^n\}\,.$$

In the linear setting, eigenvectors are defined up to scale. The same fundamental property holds when f and g are homogeneous functions. Recall that a function $f : \mathbb{R}^n \to \mathbb{R}$ is (positively) p-homogeneous if $f(\lambda x) = \lambda^p f(x)$ for all $x \in \mathbb{R}^n$ and all $\lambda \in \mathbb{R}$, $\lambda > 0$. We call p the homogeneity degree of f. For the special cases $p = 1$ and $p = 0$ we equivalently say that f is one-homogeneous and scale-invariant, respectively. In particular, in this work, we will focus on the class of homogeneous functions that are convex and have a linear kernel. This type of function appears frequently in a large number of applications, some of which are discussed in Sections 1 and 8. Precisely, we define

Definition 2.2. *For $p \geq 1$, let $CH_p^+(\mathbb{R}^n)$ denote the collection of all positively p-homogeneous functions $f : \mathbb{R}^n \to \mathbb{R}$ with the following properties:*

1. *f is convex and nonnegative, i.e. $f(x) \geq 0$ for all $x \in \mathbb{R}^n$;*
2. *$\mathrm{Ker}(f)$ is a linear subspace of $\mathbb{R}^n$.*

Here and in the following, we equivalently write $\mathrm{Ker}(f)$ and $f^{-1}(0)$ to denote the set $\{x : f(x) = 0\}$, while we let $f^{-1}(y) = \{x : f(x) = y\}$ denote the preimage of f at a generic point y.

The following remark is a direct consequence of the definition above.

Remark 2.3. *If $f \in CH_p^+(\mathbb{R}^n)$, then $f(x+z) = f(x)$ for any $z \in \mathrm{Ker}(f)$ and $x \in \mathbb{R}^n$.*

Proof of Remark 2.3. Assume the contrary holds: $f(x + z) > f(x)$ for some $x \in \mathbb{R}^n$ and $z \in \mathrm{Ker}(f)$. Fix such x and z, and let $\delta = f(x+z)-f(x) > 0$. By the convexity of f, for any $t \geq 0$, $\frac{1}{1+t} f(x + (1 + t)z) + \frac{t}{1+t} f(x) \geq f(x + z)$, which is equivalent to

$$f(x + (1 + t)z) \geq f(x + z) + t(f(x + z) - f(x)) = f(x + z) + t\delta. \tag{2.1}$$

Since $\mathrm{Ker}(f)$ is a vector space, $z \in \mathrm{Ker}(f)$ implies $(1 + t)z \in \mathrm{Ker}(f)$. Then, it follows from $(1 + t)z \in \mathrm{Ker}(f)$ and the convexity and p-homogeneity of f that

$$\frac{1}{2} f(x) = \frac{f(x) + f((1+t)z)}{2} \geq f\left(\frac{x + (1+t)z}{2}\right) = \frac{1}{2^p} f(x + (1 + t)z)$$

which yields $2^{p-1} f(x) \geq f(x+(1+t)z)$. Together with (2.1), we obtain $2^{p-1} f(x) \geq f(x+z)+t\delta$ for any $t > 0$, but it is impossible, because the right-hand-side tends to $+\infty$ when we take $t \to +\infty$. $\square$

In general, there can be infinitely many eigenvalues for a function pair, unless f and g are quadratic, in which case the corresponding eigenpairs are standard linear eigenvalue problems. One remarkable property of the spectrum of homogeneous function pairs is that when f and g are homogeneous with the same homogeneity degree and $g(\mathbb{R}^n \setminus 0) \subseteq \mathbb{R} \setminus 0$, similarly to the linear eigenvalue problem case, we can identify a set of n variational eigenvalues for the function pair (f, g) via the Lusternik-Schnirelmann theory. In fact, in that case the ratio $r(x) = f(x)/g(x) : \mathbb{R}^n \setminus 0 \to \mathbb{R}$ is scale-invariant and one has that $0 \in \partial r(x)$ implies that the pair $(r(x), x)$ is an eigenpair for (f, g). Hence, a set of n eigenvalues for (f, g) can be identified via the following variational characterization:

$$\lambda_k = \lambda_k(f, g) = \inf_{\substack{\text{genus}(S) \geq k \\ S \subset \mathbb{R}^n \setminus 0}} \sup_{x \in S} r(x), \qquad k = 1, \cdots, n, \tag{2.2}$$

where genus(S) denotes the Krasnoselski's genus of the closed, symmetric set S (see e.g. [39, Chapter 6] or [49, Page 94]), whose precise definition is recalled below.

Definition 2.4 (Krasnoselksii genus). *Let $\mathcal{A}$ be the class of closed symmetric subsets of $\mathbb{R}^n$, $\mathcal{A} = \{S \subseteq \mathbb{R}^n : S \text{ closed}, S = -S\}$. For any $S \in \mathcal{A}$, let $C_k(S) = \{\varphi : S \to \mathbb{R}^k \setminus \{0\}, \text{continuous, s.t. } \varphi(x) = -\varphi(-x)\}$. The Krasnoselskii genus of S is the number defined as*

$$\text{genus}(S) = \begin{cases} \inf\{k \in \mathbb{N} : \exists\, \varphi \in C_k(S)\}; \\ \infty, \text{ if there exists no such } k; \\ 0, \text{ if } S = \emptyset. \end{cases}$$

We provide a sketch of the proof for (2.2). Suppose that the value λ_k defined by (2.2) is not a critical value of r. Then, all the points in $\{x \in \mathbb{S}^{n-1} : r(x) = \lambda_k\}$ are not critical points, where $\mathbb{S}^{n-1}$ is the standard unit sphere in $\mathbb{R}^n$. Thus, we can use gradient flows to construct a deformation that transforms the sublevel set $\{x \in \mathbb{S}^{n-1} : r(x) \leq \lambda_k\}$ to an origin-symmetric compact subset S with genus$(S) \geq k$ and $r(x) < \lambda_k$, $\forall x \in S$. However, this leads to a contradiction with the definition of λ_k. This is the core of the min-max principle in nonlinear analysis [49].

This definition of variational eigenvalues (2.2) is a generalization of the Courant-Fisher min-max characterization of the eigenvalues $Ax = \lambda Bx$ of the pair of symmetric matrices (A, B). In fact, the Krasnoselski genus is a homeomorphism-invariant generalization to symmetric sets of the notion of dimension. In particular, genus$(S) \geq k$ for any linear subspace $S \subseteq \mathbb{R}^n$ of dimension greater than k. Thus, Courant-Fisher's characterization is retrieved from (2.2) when S is any linear subspace, the genus is replaced by the dimension of S and (f, g) are the quadratic functions $f(x) = x^\top Ax$ and $g(x) = x^\top Bx$. In particular, note that $\lambda_n(f, g) = \max_{x \neq 0} r(x)$, $\lambda_1(f, g) = \min_x r(x)$ and that, since $f^{-1}(0)$ is linear, the smallest nonzero eigenevalue of (f, g) always coincides with the smallest nonzero variational eigenvalue, i.e.,

$$\lambda_{d_f+1}(f, g) = \min\{\lambda \text{ eigenvalue of } (f, g) : \lambda > 0\}$$

where $d_f = \dim f^{-1}(0)$.

Remark 2.5 (On the use of the Lusternik-Schnirelmann category index). *The Krasnoselski's genus is arguably the most popular index function in the context of variational eigenvalues for* *nonlinear function pairs. However, when r is not even, this index cannot be used and other set measures may be required. One possibility is to use the original Lusternik-Schnirelmann category index* $\mathrm{cat}(S)$ *[3, 16, 23, 41]. However, since $\mathbb{R}^n \setminus \{0\}$ is homotopy equivalent to $\mathbb{S}^{n-1}$ and $r = f/g$ is zero-homogeneous on $\mathbb{R}^n \setminus \{0\}$, it follows from $\mathrm{cat}(\mathbb{S}^{n-1}) = 2$ that the original Lusternik-Schnirelmann category can only characterize the minimum and maximum eigenvalues in general (in contrast, $\mathrm{genus}(\mathbb{S}^{n-1}) = n$ means that the genus can be used to characterize n variational eigenvalues when r is even). To characterize more variational eigenvalues for r not even, we need to add further assumptions on r. For example, if $r : (\mathbb{R}^{n_1} \setminus \{0\}) \times \cdots \times (\mathbb{R}^{n_m} \setminus \{0\}) \to \mathbb{R}$ is a locally Lipschitz function which is zero-homogeneous on each component, that is, $r(t_1 x^1, \cdots, t_m x^m) = r(x^1, \cdots, x^m)$ for any $t_i > 0$ and $x^i \in \mathbb{R}^{n_i}$, $i = 1, \cdots, m$, we may use the Lusternik-Schnirelmann category to define $m+1$ eigenvalues of (f, g), as $(\mathbb{R}^{n_1} \setminus \{0\}) \times \cdots \times (\mathbb{R}^{n_m} \setminus \{0\})$ is homotopy equivalent to $\mathbb{S}^{n_1-1} \times \cdots \times \mathbb{S}^{n_m-1}$ whose category is $m+1$. We emphasize that all the theorems of this paper hold unchanged if* genus *is replaced by* cat. *We omit the required straightforward adjustments to the corresponding proofs for the sake of brevity.*

In the next sections, we will consider three notions of duality transforms for functions in $CH_p^+(\mathbb{R}^n)$: the norm duality, the Legendre transform and the polarity transform [1, 2]. To ensure that the class of functions $CH_p^+(\mathbb{R}^n)$ is closed under such transforms, we make a small modification to these dual operations by composing them with the orthogonal projection onto $\mathrm{Ker}(f)^\perp$, as we will detail later. If one wants to study classes of convex and homogeneous functions where $\mathrm{Ker}(f)$ can be nonlinear and can take the value $+\infty$, one should instead use the standard versions of these dual operations. It is quite interesting that most of the results we present in this paper still hold, in a certain sense, if we use the standard versions of the three transforms, as we will briefly discuss in Section 6.

3 Norm-like duality

Any norm $\|\cdot\|$ on $\mathbb{R}^n$ is a convex, one-homogeneous, nonnegative function and admits a duality transform by means of which one defines the dual norm $\|x\|^* := \sup\{\langle y, x\rangle : \|y\| \leq 1\}$. The dual norm inherits many properties from the original norm $\|\cdot\|$ and moving from one norm to the other can be of help in many applications. For a review of properties, we refer to [5, 15, 45, 56, 57]. A similar dual operator $\mathcal{N}$ can be defined for general nonnegative one-homogeneous convex functions in $CH_1^+(\mathbb{R}^n)$, as we discuss below. Our main result shows that the considered norm-like duality transform preserves the eigenpairs of any nonnegative homogeneous function pair in $CH_1^+(\mathbb{R}^n)$, as well as the corresponding multiplicities, and their variational eigenvalues.

On $CH_1^+(\mathbb{R}^n)$, consider the dual operator $\mathcal{N} : CH_1^+(\mathbb{R}^n) \to CH_1^+(\mathbb{R}^n)$ defined by

$$\mathcal{N}f(x) := \sup\Big\{\langle y, x\rangle : f(y) \leq 1 \text{ and } y \perp \mathrm{Ker}(f)\Big\}$$

for any $f \in CH_1^+(\mathbb{R}^n)$. It is worth noting that one should be careful with the notation above, as $\mathcal{N}f(x)$ denotes the dual of f at x, which implicitly depends on the variable x itself.

Note that this dual operator is essentially a composition of the "standard" norm dual operator

$$f^*(x) := \sup\{\langle y, x\rangle : f(y) \leq 1\} = \sup_{y \neq 0} \frac{\langle y, x\rangle}{f(y)}$$

and a projection onto the orthogonal complement of $\operatorname{Ker}(f)$. In other words, if π_f denotes the orthogonal projection onto $\operatorname{Ker}(f)$, then it holds

$$\mathcal{N}f(x) = f^*(x - \pi_f x)\,. \tag{3.1}$$

A short proof of (3.1) is as follows. By definition, $\mathcal{N}f(x)$ equals

$$\sup_{f(y)\le 1, y\perp \operatorname{Ker}(f)} \langle y, x\rangle = \sup_{f(y)\le 1, y\perp \operatorname{Ker}(f)} \langle y, x-\pi_f x\rangle = \sup_{f(w)\le 1} \langle w, x-\pi_f x\rangle$$

and the right hand side is nothing but $f^*(x - \pi_f x)$, where the first equality displayed above is due to $\langle y, \pi_f x\rangle = 0$ for any $y\perp\operatorname{Ker}(f)$, and the second equality displayed above is due to that by taking $y = w - \pi_f w$ there hold $f(y) = f(w)$ and $y\perp\operatorname{Ker}(f)$.

We use the "modified" norm dual $\mathcal{N}f$ instead of the standard norm dual f^* because we want to work on the function space $CH_1^+(\mathbb{R}^n)$ and we want $CH_1^+(\mathbb{R}^n)$ to be closed under the dual operation. However, $f^*(x) = +\infty$ for $x \notin (\operatorname{Ker} f)^\perp$ and $f \in CH_1^+(\mathbb{R}^n)$.

A number of useful properties follow directly from the above definition of $\mathcal{N}$, we discuss some of them in the following.

Proposition 3.1. *For any $f \in CH_1^+(\mathbb{R}^n)$ it holds $\operatorname{Ker}(f) = \operatorname{Ker}(\mathcal{N}f)$, $\mathcal{N}\mathcal{N}f = f$ and, in particular, $\mathcal{N}f \in CH_1^+(\mathbb{R}^n)$.*

Proof. Clearly, $x \notin f^{-1}(0)$ if and only if there exists $y\perp f^{-1}(0)$ such that $\langle y, x\rangle > 0$. This means that $f(x) > 0 \Leftrightarrow \mathcal{N}f(x) > 0$ for any given x, which implies $f^{-1}(0) = (\mathcal{N}f)^{-1}(0)$.

By (3.1), $\mathcal{N}\mathcal{N}f(x) = \mathcal{N}f^*(x-\pi_f x) = f^{**}(x - \pi_f x - \pi_f(x-\pi_f x)) = f^{**}(x-\pi_f x) = f(x-\pi_f x) = f(x)$ for any $x\in\mathbb{R}^n$, where we used the well-known identity $f^{**} = f$. So, $\mathcal{N}\mathcal{N}f = f$.

For any $z \in (\mathcal{N}f)^{-1}(0) = f^{-1}(0)$, $Pz = z$, and $\mathcal{N}f(x+z) = f^*(x+z-\pi_f(x+z)) = f^*(x-\pi_f x) = \mathcal{N}f(x)$. Therefore, $\mathcal{N}f\in CH_1^+(\mathbb{R}^n)$.

□

Proposition 3.2. *Let $S\subseteq\mathbb{R}^n$ be a bounded set and let $\operatorname{conv}(S)$ be its convex hull. Suppose 0 is in the relative interior of $\operatorname{conv}(S)$, and consider the support function $f_S(x) := \sup_{v\in S}\langle x, v\rangle$. Then*

$$\mathcal{N}f_S(x) = \inf\Big\{\sum_i \alpha_i : f_S\Big(\sum \alpha_i v_i - x\Big) = 0 \text{ where } \alpha_i \ge 0, v_i\in S\Big\}.$$

Proof. We rewrite f_S as f for simplicity in the proof. Note that if $f(y)\le 1$ and $v\in S$, then $\langle y, v\rangle \le 1$. Hence, $\mathcal{N}f(v)\le 1$, $\forall v\in S$. Let

$$\mathcal{F} = \{f' \in CH_1^+(\mathbb{R}^n) : f'^{-1}(0) = f^{-1}(0) \text{ and } f'(v)\le 1,\ \forall v\in S\}.$$

Then, by Proposition 3.1, we obtain $\mathcal{N}f\in\mathcal{F}$. For any $f'\in\mathcal{F}$, it is clear that $S\subset (f^{-1}(0))^\perp = (f'^{-1}(0))^\perp = (\mathcal{N}f')^{-1}(0)^\perp$, and thus

$$f(y) = \sup_{v\in S}\langle y, v\rangle \le \sup_{v\perp f'^{-1}(0): f'(v)\le 1}\langle y, v\rangle = \mathcal{N}f'(y)$$

which implies $\mathcal{N}f\ge f'$. That is, $\mathcal{N}f$ is the largest function in $\mathcal{F}$.

Consider the function $\widetilde{f} : x\mapsto \inf\{\sum\alpha_i : f(\sum\alpha_i v_i - x) = 0 \text{ for some } \alpha_i\ge 0, v_i\in S\}$. Clearly, $\widetilde{f}\in CH_1^+(\mathbb{R}^n)$, $\widetilde{f}(v)\le 1$, $\forall v\in S$, and $\widetilde{f}^{-1}(0) = f^{-1}(0)$. That is, $\widetilde{f}\in\mathcal{F}$.

For any $f' \in \mathcal{F}$, $f'(x) \leq \sum \alpha_i f'(v_i) \leq \sum \alpha_i$ whenever $x - \sum \alpha_i v_i \in f^{-1}(0)$. Taking the infimum, we get $f'(x) \leq \widetilde{f}(x)$. In consequence, we have proved that $\widetilde{f}$ is also the largest function in $\mathcal{F}$. The proof of $\mathcal{N}f = \widetilde{f}$ is then completed. $\square$

Let f_S be defined as in Proposition 3.2. Clearly one has $f_S(x) = \sup_{v \in \mathrm{conv}(S)} \langle x, v \rangle$, thus we may assume without loss of generality that S is convex. In that case, if we assume S centrally symmetric, then f_S defines a semi-norm and

$$\mathcal{N}f_S(x) = \inf \left\{ \sum |\alpha_i| : x - \sum \alpha_i v_i \perp \mathrm{span}(S) \text{ where } v_i \in S \right\} .$$

In addition, given a norm $\|\cdot\|$ and a subset $S \subset \{v : \|v\| = 1\}$ with $\mathrm{conv}((-S) \cup S) = \{v : \|v\| \leq 1\}$, we have $\|x\| = \inf \{\sum |\alpha_i| : \sum \alpha_i v_i = x, v_i \in S\}$. For example, we can take S as the set of the extreme points of the unit ball $\{v : \|v\| \leq 1\}$, and this implies the known identity $\|A\|_{l^2 \to l^2} = \inf\{\sum |\alpha_i| : A = \sum \alpha_i U_i \text{ with } U_i \text{ unitary}\}$, for a square matrix A.

Finally, we remark that, given a norm $\|\cdot\|$ on $\mathbb{R}^n$ and a linear subspace X of $\mathbb{R}^n$, the map $x \mapsto \inf\{\|z\| : z - x \perp X\}$ defines a semi-norm on $\mathbb{R}^n$. In other terms, $[x] \mapsto \inf\{\|y\| : y - x \in X\}$ defines a norm on the quotient space $\mathbb{R}^n/X$ (we refer to Gromov's norm for this basic construction [29]).

3.1 Linear transformation of homogeneous functions

Given a matrix $A \in \mathbb{R}^{m \times n}$, i.e. a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$, let $\mathcal{M}_A : CH_1^+(\mathbb{R}^n) \to CH_1^+(\mathbb{R}^m)$ be defined as

$$\mathcal{M}_A f(x) := f(A^\top x), \quad \forall f \in CH_1^+(\mathbb{R}^n), \forall x \in \mathbb{R}^m$$

where $A^\top$ denotes the transpose of A. Let π_A denote the orthogonal projection onto $\mathrm{Ran}\, A$. As $\mathbb{R}^m = \mathrm{Ran}\, A \oplus \mathrm{Ker}\, A^\top$ we can uniquely define the operator $\Pi_A : CH_1^+(\mathbb{R}^n) \to CH_1^+(\mathbb{R}^m)$ as the composition of the so-called infimal postcomposition $A \triangleright f$ (see e.g. [4]) and the orthogonal projection π_A. Precisely, we set

$$\Pi_A f(x) := A \triangleright f(\pi_A x)$$

where

$$A \triangleright f(x) = \inf_{y : Ay = x} f(y) .$$

We use this slightly modified version of the infimal postcomposition because $A \triangleright f(x) = +\infty$ for $x \notin \mathrm{Ran}\, A$.

Proposition 3.3. *Given $f \in CH_1^+(\mathbb{R}^n)$, if $\mathrm{Ker}\, f \subseteq \mathrm{Ker}\, A$, then $\Pi_A = \mathcal{N}\mathcal{M}_A\mathcal{N}$. In particular, if f is positive (i.e. $f(x) > 0$ whenever $x \neq 0$) then $\Pi_A = \mathcal{N}\mathcal{M}_A\mathcal{N}$ holds for any matrix A.*

Proof. Keeping the assumption $f^{-1}(0) \subset \mathrm{Ker}(A)$ in mind, and by employing Sion's min-max theorem [47], we have

$$\begin{aligned}
\Pi_A \mathcal{N} f(x) &= \inf_{y \in A^{-1}(x)} \sup_{u \perp f^{-1}(0) : f(u) \leq 1} \langle y, u \rangle = \sup_{u \perp f^{-1}(0) : f(u) \leq 1} \inf_{y \in A^{-1}(x)} \langle y, u \rangle \\
&= \sup_{u \perp \mathrm{Ker}(A) : f(u) \leq 1} \langle \hat{y}, u \rangle = \sup_{v : f(A^\top v) \leq 1} \langle \hat{y}, A^\top v \rangle \\
&= \sup_{v \perp \mathrm{Ker}(A^\top) : f(A^\top v) \leq 1} \langle Ay, v \rangle = \sup_{v \perp (f \circ A^\top)^{-1}(0) : f(A^\top v) \leq 1} \langle x, v \rangle = \mathcal{N}\mathcal{M}_A f(x)
\end{aligned}$$

where $\hat{y}$ is any given vector in $A^{-1}(x)$. In the above equalities, we should note that the condition $f^{-1}(0) \subset \mathrm{Ker}(A)$ implies $\mathrm{Ker}(A^\top) = (f \circ A^\top)^{-1}(0)$. In fact, $A^\top z = 0 \Rightarrow f(A^\top z) = 0 \Rightarrow AA^\top z = 0 \Rightarrow A^\top z = 0$ which means $A^\top z = 0 \Leftrightarrow f(A^\top z) = 0$. Then, the second equality from below is proved.

Replacing f by $\mathcal{N}f$, we have $\Pi_A f = \Pi_A \mathcal{N}\mathcal{N} f = \mathcal{N}\mathcal{M}_A\mathcal{N} f$. □

Before moving on, we collect in the next remark an interesting geometric interpretation of $\mathcal{N}$, $\mathcal{M}_A$ and Π_A.

Remark 3.4. *Considering a convex body K in $\mathbb{R}^n$, it is well-known that the Minkowski functional of K equals the support function of its dual convex body K°. The dual operator transforms the Minkowski functional of K to its support function, while Π_A maps the Minkowski functional of K to the Minkowski functional of $A(K) \times \mathrm{Ker}(A)$, and $\mathcal{M}_A$ maps the support function of K to the support function of $A(K)$. If A is further assumed to be a projection, then $\mathcal{M}_A$ maps the Minkowski functional of K to the Minkowski functional of $K \cap \mathrm{Ker}(A)^\perp$, while Π_A transforms the support function of K to the support function of $K \cap \mathrm{Ker}(A)^\perp$.*

Note that, as a consequence of Proposition 3.3, if $f \in CH_1^+(\mathbb{R}^n)$ is positive, $n = m$ and A is an invertible matrix, we have $\mathcal{N}\mathcal{M}_A\mathcal{N}f(x) = f(A^{-1}x)$, and therefore, $\mathcal{N}\mathcal{M}_A\mathcal{N}f(x) = \mathcal{M}_A f(x)$ whenever A is an orthogonal matrix. Moreover, for a general $f \in CH_1^+(\mathbb{R}^n)$, we have the identities $\mathcal{M}_A\mathcal{N}f = \mathcal{N}\Pi_A f$ and $\mathcal{N}\mathcal{M}_A f = \Pi_A\mathcal{N}f$. The equality $\mathcal{M}_A\mathcal{N}f = \mathcal{N}\Pi_A f$ means that "the section of the dual equals the dual of the projection", which is a useful observation with direct implications in convex geometry. On the other hand, the equality $\mathcal{N}\mathcal{M}_A f = \Pi_A\mathcal{N}f$ has a similar geometrical meaning, and it has an interesting additional consequence, which we summarize in the following proposition.

Proposition 3.5. *Let $\|\cdot\|$ be a monotonic norm on $\mathbb{R}^d$, i.e., $\|(t_1, \cdots, t_d)\| = \|(|t_1|, \cdots, |t_d|)\|$ for any $(t_1, \cdots, t_d) \in \mathbb{R}^d$. Let $g_i \in CH_1^+(\mathbb{R}^{n_i})$ be positive-definite, and let $A_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$ be a linear map, i.e., $A_i \in \mathbb{R}^{n_i \times n}$, $i = 1, \cdots, d$. Denote by $\hat{g}(x) = \|(g_1(A_1 x), \cdots, g_d(A_d x))\|$. Then*

$$\mathcal{N}\hat{g}(x) = \inf_{\sum_{i=1}^d A_i^\top x_i = x} \|(\mathcal{N}g_1(x_1), \cdots, \mathcal{N}g_d(x_d))\|_*,$$

where $\|\cdot\|_*$ is the dual norm induced by $\|\cdot\|$.

Note that, by letting $g_1, \cdots, g_d$ be norms, we immediately obtain Theorem 6 in [25], which has implications in the design of converging iterations for general matrix norm computations.

Proof. Let $\tilde{g}(x_1, \cdots, x_d) = \|(g_1(x_1), \cdots, g_d(x_d))\|$, $\forall (x_1, \cdots, x_d) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_d}$. Then,

$$\begin{aligned}
\|(\mathcal{N}g_1(x_1), \cdots, \mathcal{N}g_d(x_d))\| &= \sup_{\|(t_1,\cdots,t_d)\| \le 1} \sum_{i=1}^d t_i \mathcal{N}g_i(x_i) \\
&= \sup_{\|(|t_1|,\cdots,|t_d|)\| \le 1} \sum_{i=1}^d |t_i| \sup_{g_i(y_i) \le 1} \langle x_i, y_i \rangle \\
&= \sup_{\|(|t_1|,\cdots,|t_d|)\| \le 1} \sum_{i=1}^d \sup_{g_i(y_i) \le |t_i|} \langle x_i, y_i \rangle
\end{aligned}$$

$$
\begin{aligned}
&= \sup_{\|(|t_1|,\cdots,|t_d|)\|\le 1} \sup_{g_i(y_i)\le |t_i|,\forall i} \sum_{i=1}^{d} \langle x_i, y_i\rangle \\
&= \sup_{\|(g_1(y_1),\cdots,g_d(y_d))\|\le 1} \sum_{i=1}^{d} \langle x_i, y_i\rangle \\
&= \sup_{\tilde{g}(y_1,\cdots,y_d)\le 1} \langle (x_1,\cdots,x_d),(y_1,\cdots,y_d)\rangle \\
&= \mathcal{N}\tilde{g}(x_1,\cdots,x_d).
\end{aligned}
$$

Note that $\hat{g}(x) = \tilde{g}(A^\top x)$, where $A := [A_1^\top, \cdots, A_d^\top] \in \mathbb{R}^{n\times(n_1+\cdots+n_d)}$. The proof is then completed by the identity $\mathcal{N}\hat{g} = \mathcal{N}\mathcal{M}_A\tilde{g} = \Pi_A\mathcal{N}\tilde{g}$. □

4 Main results: spectral invariance for norm-like duality

We state here our main theorem showing that nonzero eigenvalues of function pairs, as well as their multiplicities and their variational eigenvalues (2.2), are invariant under the norm-like duality and suitable combinations of $\mathcal{M}_A$ and Π_A, for any matrix A. The relatively long proofs of this theorem and its main corollary cover the entire section.

Throughout the remainder of this paper, the 'eigenspace' of λ with respect to the function pair (f, g) is the set $S_\lambda(f, g)$ defined by

$$S_\lambda(f,g) = \left\{x : 0 \in \partial f(x) - \lambda\partial g(x)\right\}.$$

Note that when f and g are even functions, $S_\lambda(f, g)$ is a symmetric set. In this case, we define the multiplicity of the eigenvalue λ for (f, g) as

$$\mathrm{mult}_{f,g}(\lambda) = \mathrm{genus}\big(S_\lambda(f,g)\big).$$

For a set S, let $\mathrm{cone}(S) = \{\lambda v : \lambda > 0, v \in S\}$ denote the cone generated by S. The following main spectral invariance theorem holds,

Theorem 4.1. *Let $f, g \in CH_1^+(\mathbb{R}^n)$. Then*

P1. The nonzero eigenvalues of (f, g) and $(\mathcal{N}g, \mathcal{N}f)$ coincide. Moreover, for any eigenpair (λ, x) of (f, g) with $\lambda \neq 0$, and for any $u \in \mathrm{cone}(\partial f(x)) \cap \mathrm{cone}(\partial g(x))$, (λ, u) is an eigenpair of $(\mathcal{N}g, \mathcal{N}f)$.

P2. If f and g are even functions, then $\mathrm{mult}_{f,g}(\lambda) = \mathrm{mult}_{\mathcal{N}g,\mathcal{N}f}(\lambda)$, for any nonzero eigenvalue λ of (f, g).

P3. If f and g are even functions, then the variational eigenvalues of (f, g) and $(\mathcal{N}g, \mathcal{N}f)$ coincide exactly, up to reordering. Precisely, it holds

$$\lambda_k(f,g) = \lambda_{k-d_f+d_g}(\mathcal{N}g,\mathcal{N}f), \qquad k = d_f - d_{fg} + 1, \cdots, n - d_g$$

where $d_{fg} := \dim f^{-1}(0) \cap g^{-1}(0)$, $d_f := \dim f^{-1}(0)$ and $d_g := \dim g^{-1}(0)$.

Moreover, combining the norm-like duality operator $\mathcal{N}$ with $\mathcal{M}_A$ and Π_A for a matrix A, we obtain the following main consequence of the theorem above.

Corollary 4.2. *Let $f \in CH_1^+(\mathbb{R}^m)$, $g \in CH_1^+(\mathbb{R}^n)$ and $A \in \mathbb{R}^{n\times m}$. Then, the nonzero eigenvalues of $(\mathcal{M}_{A^\top} f, g)$, $(\mathcal{N}g, \mathcal{N}\mathcal{M}_{A^\top} f)$, $(\mathcal{M}_A\mathcal{N}g, \mathcal{N}f)$, $(f, \Pi_A g)$ and $(\mathcal{M}_{A^\top} f, \mathcal{M}_{A^\top}\Pi_A g)$ coincide. Moreover, if f and g are even functions, then the multiplicities of the nonzero eigenvalues coincide and the nonzero variational eigenvalues of all these function pairs coincide exactly, up to reordering.*

We subdivide the relatively long proof of the main results above into several separate parts, as well as a number of smaller preliminary results that are of independent interest.

First, we prove that nonzero eigenvalues are preserved under $\mathcal{N}$.

Proof of Theorem 4.1 point P1. For an eigenpair (λ, x) of (f, g) with $\lambda \neq 0$ and $x \neq 0$, it is easy to see that $f(x) = 0 \Leftrightarrow g(x) = 0$, and in this case, we have $\mathcal{N}f(x) = 0$, $\mathcal{N}g(x) = 0$, and $0 \in \partial\mathcal{N}f(x) \cap \partial\mathcal{N}g(x)$ which implies $0 \in \partial\mathcal{N}g(x) - \lambda\partial\mathcal{N}f(x)$. Hence, (λ, x) is also an eigenpair of $(\mathcal{N}g, \mathcal{N}f)$. In fact, from this proof, we obtain that if $f^{-1}(0) \cap g^{-1}(0) \neq \{0\}$, then the spectra of (f, g) and $(\mathcal{N}g, \mathcal{N}f)$ are $\mathbb{R}$. Therefore, without loss of generality, we assume that $f^{-1}(0) \cap g^{-1}(0) = \{0\}$, $g(x) = 1$ and $f(x) = \lambda \neq 0$. Thus, there exists $u \in \partial g(x)$ such that $\lambda u \in \partial f(x)$. Clearly, $u \neq 0$. It follows from the fact $\partial g(x) \subset (g^{-1}(0))^\perp = ((\mathcal{N}g)^{-1}(0))^\perp$ that $\mathcal{N}g(u) \neq 0$. Moreover, we have $\langle u, x\rangle = g(x) = 1$ by Euler's identity for positively one-homogeneous convex functions, and $\langle u, x'\rangle - 1 = \langle u, x' - x\rangle \leq g(x') - g(x) = g(x') - 1$, $\forall x' \in \mathbb{R}^n$ by the definition of the subgradient. Accordingly, $\mathcal{N}g(u) = 1$, and for any $u' \in \mathbb{R}^n$, $\langle u' - u, x\rangle = \langle u', x\rangle - 1 \leq \mathcal{N}g(u') - 1 = \mathcal{N}g(u') - \mathcal{N}g(u)$, which implies that $x \in \partial\mathcal{N}g(u)$. By $f(x/\lambda) = 1$ and $\lambda u \in \partial f(x) = \partial f(x/\lambda)$, we similarly derive that $x/\lambda \in \partial\mathcal{N}f(\lambda u) = \partial\mathcal{N}f(u)$ according to the zero-homogeneity of ∂f and $\partial\mathcal{N}f$. As a consequence, $0 = x - \lambda\cdot x/\lambda \subset \partial\mathcal{N}g(u) - \lambda\partial\mathcal{N}f(u)$, i.e., (λ, u) is an eigenpair of $(\mathcal{N}g, \mathcal{N}f)$. The converse also holds. And since $\partial\mathcal{N}f$ and $\partial\mathcal{N}g$ are scaling invariant, we indeed obtain that $\forall u \in \mathrm{cone}(\partial f(x)) \cap \mathrm{cone}(\partial g(x))$, (λ, u) is an eigenpair of $(\mathcal{N}g, \mathcal{N}f)$. ☐

Then, we move on to studying their multiplicities. To this end, we first observe that the genus of a compact set grows under the action of the subgradient of even functions. Here and throughout, we say a function is C^1-smooth if it has continuous gradient on $\mathbb{R}^n \setminus \{0\}$.

Lemma 4.3. *Let $g \in CH_1^+(\mathbb{R}^n)$ be an even function. Then, the Krasnoselskii genus of a compact subset S is smaller than or equal to that of the subset $\partial g(S) := \cup_{x\in S}\partial g(x)$.*

Proof. The proof is based on the deformation nondecreasing property and the continuity of the Krasnoselskii genus. We divide the proof into two steps:

Step 1. Suppose that $g \in CH_1^+(\mathbb{R}^n)$ is C^1-smooth on $\mathbb{R}^n \setminus \{0\}$. Since the vector field induced by $\partial g : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n \setminus \{0\}$ is continuous, for any compact subset $S \subset \mathbb{R}^n \setminus \{0\}$ with $\mathrm{genus}(S) = k$, the map $x \mapsto \partial g(x)$ is continuous and if g is even, then ∂g is odd, i.e., $\partial g(-x) = -\partial g(x)$, $\forall x \in \mathbb{R}^n$. Therefore, by the deformation nondecreasing property, $\partial g(S)$ is a subset of $\mathbb{R}^n$ with $\mathrm{genus}(\partial g(S)) \geq k$. That is, for a even, convex and smooth function g, we have $\mathrm{genus}(\partial g(S)) \geq \mathrm{genus}(S)$.

Step 2. Suppose that g is not C^1-smooth on $\mathbb{R}^n \setminus \{0\}$.

In this case, we take the Moreau-Yosida approximation of g, which is defined by

$$g_\alpha(x) = \inf_{y\in\mathbb{R}^n} g(y) + \frac{1}{2\alpha}\|y - x\|_2^2,\ \alpha > 0,$$

where we use the l^2-norm $\|\cdot\|_2$. It is known that g_α is C^1-smooth and convex. In fact, for sufficiently small $\epsilon > 0$, the size of the ϵ-neighborhood of $\partial g(S)$ equals genus$(\partial g(S))$, and for sufficiently small α, $\partial g_\alpha(S)$ lies in the ϵ-neighborhood of $\partial g(S)$. Therefore, genus$(\partial g(S)) \geq$ genus$(\partial g_\alpha(S))$, which is larger than or equal to genus(S) by Step 1.

□

Next, we show that for smooth functions the subgradient maps the eigenspace of λ as an eigenvalue of (f, g) into the eigenspace of λ as an eigenvalue of the dual pair $(\mathcal{N}g, \mathcal{N}f)$.

Lemma 4.4. *Let $f, g \in CH_1^+(\mathbb{R}^n)$ and let λ be an eigenvalue of (f, g). If g is differentiable on $\mathbb{R}^n \setminus \{0\}$, then $\partial g(S_\lambda(f, g)) \subset S_\lambda(\mathcal{N}g, \mathcal{N}f)$. Similarly, if f is differentiable, then $\partial f(S_\lambda(f, g)) \subset S_\lambda(\mathcal{N}g, \mathcal{N}f)$.*

Proof. By point P1 of Theorem 4.1 we have that

$$\varnothing \neq \bigcup_{x \in S_\lambda(f,g)} \mathrm{cone}(\partial f(x)) \cap \mathrm{cone}(\partial g(x)) \subset S_\lambda(\mathcal{N}g, \mathcal{N}f)$$

for any eigenvalue λ of (f, g). If g is differentiable at any eigenvector $x \in S_\lambda(f, g)$, then $\partial g(x) \subset \mathrm{cone}(\partial f(x)) \cap \partial g(x)$. Thus,

$$\partial g(S_\lambda(f, g)) := \bigcup_{x \in S_\lambda(f,g)} \partial g(x) \subset S_\lambda(\mathcal{N}g, \mathcal{N}f).$$

The proof of $\partial f(S_\lambda(f, g)) \subset S_\lambda(\mathcal{N}g, \mathcal{N}f)$ is similar. □

Finally, we need the following two technical properties.

Lemma 4.5. *Let $f, g \in CH_1^+(\mathbb{R}^n)$ and let λ be an eigenvalue of the function pair (f, g). It holds*

1. *The map $x \mapsto \mathrm{cone}(\partial f(x)) \cap \partial g(x)$ is upper semi-continuous, i.e., $\forall x$, $\forall \epsilon > 0$, there exists $\delta > 0$ such that for any $y \in \mathbb{B}_\delta(x)$, $\mathrm{cone}(\partial f(y)) \cap \partial g(y) \subset \mathbb{B}_\epsilon(\mathrm{cone}(\partial f(x)) \cap \partial g(x))$, where $\mathbb{B}_\epsilon(S)$ is the ϵ-neighborhood of a subset S.*

2. *For any $x \in S_\lambda(f, g)$, and for any $\epsilon > 0$, there exists an even, C^1-smooth function $g_x \in CH_1^+(\mathbb{R}^n)$ with $g_x^{-1}(0) = g^{-1}(0)$ and $\delta > 0$ such that $\partial g_x(\mathbb{B}_\delta(x)) \subset \mathbb{B}_\epsilon(\mathrm{cone}(\partial f(x)) \cap \partial g(x))$.*

Proof. Point 1 follows directly from the upper semi-continuity of ∂f and ∂g. Let us discuss point 2. We only need to deal with the case that g is positive-definite. For any $v \in \partial g(x)$, $\langle x, v\rangle = g(x) > 0$. Then, by a standard argument in linear algebra, there exists a positive-definite matrix A such that $Ax = v$. Then, we take $g_x(y) = \sqrt{\langle x, Ax\rangle \cdot \langle y, Ay\rangle}$. It is clear that g_x is smooth, positive-definite and convex and one-homogeneous. And it is not difficult to check that $\partial g_x(x) = Ax = v$. Now, suppose that the vector v lies in $\mathrm{cone}(\partial f(x)) \cap \partial g(x)$. By the above discussion, we immediately obtain that $\forall \epsilon > 0$, $\exists \delta > 0$ such that $\partial g_x(\mathbb{B}_\delta(x)) \subset \mathbb{B}_\epsilon(v) \subset \mathbb{B}_\epsilon(\mathrm{cone}(\partial f(x)) \cap \partial g(x))$. □

Proof of Theorem 4.1 point P2. From Lemmas 4.3 and 4.4 it is clear to see that genus$(S_\lambda(f, g)) \leq$ genus$(S_\lambda(\mathcal{N}g, \mathcal{N}f))$ if f or g is differentiable. Conversely, if $\mathcal{N}f$ or $\mathcal{N}g$ is differentiable, then genus$(S_\lambda(f,g)) \geq$ genus$(S_\lambda(\mathcal{N}g,\mathcal{N}f))$. Thus, we obtain that the multiplicity of λ as an eigenvalue of (f,g) coincides with the multiplicity of λ as an eigenvalue of $(\mathcal{N}g,\mathcal{N}f)$. Next, we prove that the same property holds without the differentiability condition.

Let $\hat{S}_\lambda(f,g) = S_\lambda(f,g) \cap \{x : \|x\|_2 = 1\}$ be the 'unit sphere' of the eigenspace corresponding to λ. Then, the multiplicity of λ coincides with genus$(\hat{S}_\lambda(f,g))$. Fix an $\epsilon > 0$ such that

$$\text{genus}\, \mathbb{B}_\epsilon\left(\bigcup_{x\in\hat{S}_\lambda(f,g)} \text{cone}(\partial f(x)) \cap \partial g(x)\right) = \text{genus}\left(\bigcup_{x\in\hat{S}_\lambda(f,g)} \text{cone}(\partial f(x)) \cap \partial g(x)\right).$$

Take $\epsilon' < \frac{1}{2}\epsilon$. Due to Lemma 4.5, we can consider a family of open sets $\{\mathbb{B}_{\delta_x}(x) : x \in \hat{S}_\lambda(f,g)\}$ and the corresponding smooth function family $\{g_x : x \in \hat{S}_\lambda(f,g)\}$ such that for any $y \in \mathbb{B}_{2\delta_x}(x)$, we have $\text{cone}(\partial f(y)) \cap \partial g(y) \subset \mathbb{B}_{\epsilon'}(\text{cone}(\partial f(x)) \cap \partial g(x))$ and $\partial g_x(\mathbb{B}_{2\delta_x}(x)) \subset \mathbb{B}_{\epsilon'}(\text{cone}(\partial f(x)) \cap \partial g(x))$, for a sufficiently small δ_x.

Since $\hat{S}_\lambda(f,g)$ is compact and $\{\mathbb{B}_{\delta_x}(x) : x \in \hat{S}_\lambda(f,g)\}$ induces an open cover of $\hat{S}_\lambda(f,g)$, we can take a finite subfamily $\{\mathbb{B}_{\delta_i}(x_i)\}$ of $\{\mathbb{B}_{\delta_x}(x) : x \in \hat{S}_\lambda(f,g)\}$ such that the centers $\{x_i\}$ of these open balls are distributed centrally symmetrically in $\mathbb{R}^n$, and $\partial g_i(\mathbb{B}_{\delta_i}(x_i)) \subset \mathbb{B}_{\epsilon'}(\text{cone}(\partial f(x_i)) \cap \partial g(x_i))$, where we simply write g_{x_i} as g_i. Then, there exist partitions of unity $\{\psi_i\}$ subordinate to the open cover $\{\mathbb{B}_{\delta_i}(x_i)\}$, i.e., $\text{supp}(\psi_i) \subset \mathbb{B}_{\delta_i}(x_i)$, $\psi_i \geq 0$, $\sum_i \psi_i = 1$ and $\psi_i = \psi_{i'}$ whenever $x_i = -x_{i'}$. For example, we can simply take

$$\psi_i(y) = \frac{\max\{0, \delta_i - \|y - x_i\|_2\}}{\sum_j \max\{0, \delta_j - \|y - x_j\|_2\}}, \quad \forall y \in \mathbb{R}^n.$$

Taking $\Psi(x) = \sum_i \psi_i(x)\partial g_i(x)$, then Ψ is a continuous map.

Given $x \in \hat{S}_\lambda(f,g)$, let $I(x) = \{i : x \in \mathbb{B}_{\delta_i}(x_i)\}$ be the index set of x. Note that $\psi_i(x) > 0$ implies $x \in \mathbb{B}_{\delta_i}(x_i)$, and thus it holds $\Psi(x) = \sum_{i\in I(x)} \psi_i(x)\partial g_i(x)$ and $\partial g_i(x) \in \mathbb{B}_{\epsilon'}(\text{cone}(\partial f(x_i)) \cap \partial g(x_i))$, whenever $x \in \mathbb{B}_{\delta_i}(x_i)$. Moreover, there exists a bijection $\tau : I(x) \to I(-x)$ such that $x_i = -x_{\tau(i)}$, which implies and $\psi_i(x) = \psi_{\tau(i)}(-x)$ and $\partial g_i(x) = -\partial g_{\tau(i)}(-x)$. This implies that

$$\begin{aligned}\Psi(-x) &= \sum_{i\in I(-x)} \psi_i(-x)\partial g_i(-x) = \sum_{i\in I(x)} \psi_{\tau(i)}(-x)\partial g_{\tau(i)}(-x) \\ &= \sum_{i\in I(x)} -\psi_i(x)\partial g_i(x) = -\Psi(x).\end{aligned}$$

Let $i(x) = \text{argmax}\{\delta_i : i \in I(x)\}$. Then, for any $i \in I(x)$, $x_i \in \mathbb{B}_{\delta_i}(x) \subset \mathbb{B}_{\delta_i}(\mathbb{B}_{\delta_{i(x)}}(x_{i(x)})) = \mathbb{B}_{\delta_i+\delta_{i(x)}}(x_{i(x)}) \subset \mathbb{B}_{2\delta_{i(x)}}(x_{i(x)})$. Thus, $\forall i \in I(x)$, $\text{cone}(\partial f(x_i)) \cap \partial g(x_i) \subset \mathbb{B}_{\epsilon'}(\text{cone}(\partial f(x_{i(x)})) \cap \partial g(x_{i(x)}))$. Therefore, $\partial g_i(x) \in \mathbb{B}_{2\epsilon'}(\text{cone}(\partial f(x_{i(x)})) \cap \partial g(x_{i(x)}))$ for any $i \in I(x)$. Consequently, we have

$$\begin{aligned}\Psi(x) &= \sum_{i\in I(x)} \psi_i(x)\partial g_i(x) \in \mathbb{B}_{2\epsilon'}(\text{cone}(\partial f(x_{i(x)})) \cap \partial g(x_{i(x)})) \\ &\subset \mathbb{B}_\epsilon\left(\bigcup_{x\in\hat{S}_\lambda(f,g)} \text{cone}(\partial f(x)) \cap \partial g(x)\right)\end{aligned}$$

which implies that $\Psi(\hat{S}_\lambda(f,g)) \subset \mathbb{B}_\epsilon\left(\cup_{x\in\hat{S}_\lambda(f,g)}\mathrm{cone}(\partial f(x))\cap\partial g(x)\right)$. Thus,

$$\begin{aligned}\mathrm{genus}(\hat{S}_\lambda(f,g)) &\le \mathrm{genus}(\Psi(\hat{S}_\lambda(f,g)) \le \mathrm{genus}\ \mathbb{B}_\epsilon\left(\bigcup_{x\in\hat{S}_\lambda(f,g)}\mathrm{cone}(\partial f(x))\cap\partial g(x)\right)\\ &= \mathrm{genus}\left(\bigcup_{x\in\hat{S}_\lambda(f,g)}\mathrm{cone}(\partial f(x))\cap\partial g(x)\right)\end{aligned}$$

where the first inequality is due to the fact that Ψ is odd continuous, the second inequality is based on the nondecreasing property of the genus, and the last equality follows from the continuity of the genus.

In summary, we have proved that for any $f,g\in CH_1^+(\mathbb{R}^n)$ and any (λ,x) eigenpair of (f,g) there always holds

$$\mathrm{genus}\big(S_\lambda(\mathcal{N}g,\mathcal{N}f)\big) = \mathrm{genus}\left(\bigcup_{x\in S_\lambda(f,g)}\mathrm{cone}(\partial f(x))\cap\partial g(x)\right) \ge \mathrm{genus}\big(S_\lambda(f,g)\big).$$

The reverse inequality follows from replacing f with $\mathcal{N}f$, g with $\mathcal{N}g$ and using Proposition 3.1. □

Proof of Theorem 4.1 point P3. It is straightforward to verify that

$$\begin{aligned}&0=\lambda_1(f,g)=\cdots=\lambda_{d_f-d_{fg}}(f,g)<\lambda_{d_f-d_{fg}+1}(f,g)\le\cdots\le\lambda_{n-d_g}(f,g),\\ &0=\lambda_1(\mathcal{N}g,\mathcal{N}f)=\cdots=\lambda_{d_g-d_{fg}}(\mathcal{N}g,\mathcal{N}f)<\lambda_{d_g-d_{fg}+1}(\mathcal{N}g,\mathcal{N}f)\le\cdots\le\lambda_{n-d_f}(\mathcal{N}g,\mathcal{N}f).\end{aligned}$$

Without loss of generality, we may assume that $f^{-1}(0)\cap g^{-1}(0)=\{0\}$, and in this case, we shall prove that $\lambda_{k-d_f+d_g}(\mathcal{N}g,\mathcal{N}f)\le\lambda_k(f,g)$, $k=d_f+1,\cdots,n-d_g$. For any subset $S\subset g^{-1}(1)$ realizing $\lambda_k(f,g)$ with $\mathrm{genus}(S)\ge k$, i.e., a set such that $\sup_{x\in S}f(x)/g(x)=\lambda_k(f,g)$, we have $\lambda_k(f,g)\ge f(x)/g(x)=f(x)$, $\forall x\in S$. Let $\mathbb{S}$ be the unit sphere in the linear subspace $g^{-1}(0)$ centered at the origin 0. Let $W=\partial g(S)*\mathbb{S}$ be the geometric join of $\partial g(S)$ and $\mathbb{S}$, i.e., $W=\{tu+(1-t)y: u\in\partial g(S), y\in\mathbb{S}, 0\le t\le 1\}$. Since $\partial g(S):=\cup_{x\in S}\partial g(x)\subset(g^{-1}(0))^\perp$ is orthogonal to the sphere $\mathbb{S}$ in the linear subspace $g^{-1}(0)$, it holds $\mathrm{genus}(W)=\mathrm{genus}(\partial g(S))+\mathrm{genus}(\mathbb{S})$.

For any $y\in W\cap(f^{-1}(0))^\perp$, there exist $0\le t\le 1$, $u\in\partial g(S)$ and $-v\in(1-t)\mathbb{S}$, such that $y=tu-v$. And there exists $x\in S$ such that $u\in\partial g(x)$. Therefore, $x\in\partial\mathcal{N}g(u)=\partial\mathcal{N}g(tu)$, $\mathcal{N}g(tu)=t\mathcal{N}g(u)=t$ and $\mathcal{N}g(v)=0$. Note that by the definition of subgradients, $x\in\partial\mathcal{N}g(tu)$ implies $\mathcal{N}g(v)-\mathcal{N}g(tu)\ge\langle v-tu,x\rangle$. Thus, we have

$$\begin{aligned}\mathcal{N}f(tu-v) &= \sup_{z\perp f^{-1}(0)}\frac{\langle tu-v,z\rangle}{f(z)} = \sup_{z\neq 0}\frac{\langle tu-v,z\rangle}{f(z)} \ge \frac{\langle tu-v,x\rangle}{f(x)}\\ &\ge \frac{\mathcal{N}g(tu)-\mathcal{N}g(v)}{f(x)} = \frac{t}{f(x)} \ge \frac{t}{\lambda_k(f,g)}\end{aligned}$$

and $\mathcal{N}g(tu-v)=\mathcal{N}g(tu)=t$. This implies that $\mathcal{N}g(tu-v)/\mathcal{N}f(tu-v)\le\lambda_k(f,g)$. Hence $\sup_{y\in W}\mathcal{N}g(y)/\mathcal{N}f(y)\le\lambda_k(f,g)$. Now, note that

$$\mathrm{genus}(W\cap(f^{-1}(0))^\perp)\ge\mathrm{genus}(\partial g(S))+\mathrm{genus}(\mathbb{S})-\dim f^{-1}(0)$$

$$\geq \operatorname{genus}(S) + \dim g^{-1}(0) - \dim f^{-1}(0) \geq k + d_g - d_f$$

in which we used the claim $\operatorname{genus}(\partial g(S)) \geq \operatorname{genus}(S)$. Thus, for $k = d_f + 1, \cdots, n - d_g$ we obtain $\lambda_{k+d_g-d_f}(\mathcal{N}g, \mathcal{N}f) \leq \lambda_k(f, g)$. Analogously, for $k' = d_g + 1, \cdots, n - d_f$, we have $\lambda_{k'+d_f-d_g}(f, g) \leq \lambda_{k'}(\mathcal{N}g, \mathcal{N}f)$. Substituting $k' = k + d_g - d_f$ into the latter inequality, we get $\lambda_k(f, g) \leq \lambda_{k+d_g-d_f}(\mathcal{N}g, \mathcal{N}f)$, and therefore, we derive $\lambda_{k+d_g-d_f}(\mathcal{N}g, \mathcal{N}f) = \lambda_k(f, g)$, $k = d_f + 1, \cdots, n - d_g$. □

We now move on to the proof of Corollary 4.2. We need one more preliminary lemma.

Lemma 4.6. *For g and A as in Corollary 4.2, define $g_{\operatorname{Ker}(A)}(x) = \inf_{x' \in x + \operatorname{Ker}(A)} g(x') = \mathcal{M}_{A^\top} \Pi_A g(x)$ and let $S = \{x \in \mathbb{R}^n : g(x) = g_{\operatorname{Ker}(A)}(x)\}$. Then, x is an eigenvector corresponding to a nonzero eigenvalue of $(\mathcal{M}_{A^\top} f, g)$ only if $x \in S$.*

Proof. If $x \notin S$, we shall prove that $\partial g(x) \cap \operatorname{Ker}(A)^\perp = \varnothing$. Otherwise, there exists $v \in \partial g(x)$ such that $v \perp \operatorname{Ker}(A)$. Then taking $y \in x + \operatorname{Ker}(A)$ such that $g(y) = \inf_{x' \in x + \operatorname{Ker}(A)} g(x')$, we have $0 > g(y) - g(x) \geq \langle v, y - x\rangle = 0$ which leads to a contradiction. Thus, we have shown that $\partial g(x) \cap \operatorname{Ker}(A)^\perp = \varnothing$. On the other hand, $\partial_x f(Ax) = A^\top \partial f(Ax) \subset \operatorname{Range}(A^\top) = \operatorname{Ker}(A)^\perp$. This implies that, for any $\lambda \neq 0$, $\partial_x f(Ax) \cap \lambda \partial g(x) \subset \operatorname{Ker}(A)^\perp \cap \lambda \partial g(x) = \varnothing$, which means that x is not an eigenvector of any nonzero eigenvalue of $(\mathcal{M}_{A^\top} f, g)$. The proof is completed. □

Proof of Corollary 4.2. We organize the proof as illustrated by the diagram below

$$\begin{array}{ccccc}
(\mathcal{M}_{A^\top} f, \mathcal{M}_{A^\top}\Pi_A g) & \overset{?}{\Longleftrightarrow} & (\mathcal{M}_{A^\top} f, g) & \overset{\text{Thm 4.1}}{\Longleftrightarrow} & (\mathcal{N}g, \mathcal{N}\mathcal{M}_{A^\top} f) \\
 & & \Big\Updownarrow{\scriptstyle ?} & & \\
 & & (\mathcal{M}_A \mathcal{N} g, \mathcal{N} f) & \overset{\text{Thm 4.1}}{\Longleftrightarrow} & (f, \mathcal{N}\mathcal{M}_A \mathcal{N} g)
\end{array}$$

Here, '⇔' denotes 'spectral equivalence', i.e., the thesis holds for the two nonlinear eigenvalue problems connected by '⇔'. Note that $\mathcal{M}_{A^\top} f = f \circ A \in CH_1^+(\mathbb{R}^n)$ and $g \in CH_1^+(\mathbb{R}^n)$. Thus, by Theorem 4.1, the thesis holds for $(\mathcal{N}g, \mathcal{N}\mathcal{M}_{A^\top} f)$ and $(\mathcal{M}_{A^\top} f, g)$. The same is true for $(\mathcal{M}_A \mathcal{N} g, \mathcal{N} f)$ and $(f, \mathcal{N}\mathcal{M}_A \mathcal{N} g)$. In the remainder of the proof, we will show that the two relations marked with a '?' hold.

We first prove that the set of nonzero eigenvalues of $(\mathcal{M}_{A^\top} f, g)$ coincides with the set of nonzero eigenvalues of $(\mathcal{M}_A \mathcal{N} g, \mathcal{N} f)$. For an eigenpair (λ, x) of $(\mathcal{M}_{A^\top} f, g)$ with $g(x) = 1$, we have $0 \in \partial_x f(Ax) - \lambda \partial g(x) = A^\top \partial f(Ax) - \lambda \partial g(x)$. Hence, there exists $u \in \partial g(x)$ such that $\lambda u = A^\top v$ for some $v \in \partial f(Ax)$. Thus, $Ax/\lambda \in \partial \mathcal{N} f(v)$ and $x \in \partial \mathcal{N} g(u) = \partial \mathcal{N} g(A^\top v/\lambda) = \partial \mathcal{N} g(A^\top v)$. Therefore, $Ax \in A \partial \mathcal{N} g(A^\top v) = \partial_v \mathcal{N} g(A^\top v) = \partial \mathcal{M}_A \mathcal{N} g(v)$, which implies $Ax \in \partial \mathcal{M}_A \mathcal{N} g(v) \cap \lambda \partial \mathcal{N} f(v)$ and (λ, v) is an eigenpair of $(\mathcal{M}_A \mathcal{N} g, \mathcal{N} f)$. Since $(\mathcal{M}_{A^\top} \mathcal{N}\mathcal{N} f, \mathcal{N}\mathcal{N} g) = (\mathcal{M}_{A^\top} f, g)$, the converse also holds. In summary, we have shown that

$$\varnothing \neq \bigcup_{x \in S_\lambda(\mathcal{M}_{A^\top} f, g)} \partial f(Ax) \cap (A^\top)^{-1}(\lambda \partial g(x)) \subset S_\lambda(\mathcal{M}_A \mathcal{N} g, \mathcal{N} f).$$

Together with Lemma 4.3, this shows that the multiplicity is also maintained.

Next, we show that $(\mathcal{M}_{A^\top} f, \mathcal{M}_{A^\top} \Pi_A g)$ and $(\mathcal{M}_{A^\top} f, g)$ have the same nonzero eigenvalues. By the definitions of the operators $\mathcal{M}_{A^\top}$ and Π_A, we have

$$\mathcal{M}_{A^\top} \Pi_A g(x) = \inf_{y \in A^{-1}(Ax)} g(y) = \inf_{z \in \operatorname{Ker}(A)} g(x + z).$$

Let $g_{\mathrm{Ker}(A)}$ and S be as in Lemma 4.6. For any $x \in S$, we have $\partial_x f(Ax) = A^\top \partial f(Ax) \subset \mathrm{Im}(A^\top) = \mathrm{Ker}(A)^\perp$ and thus

$$\begin{aligned}\partial_x f(Ax) \cap \lambda \partial_x g(x) &= A^\top \partial f(Ax) \cap \lambda \partial g(x) = A^\top \partial f(Ax) \cap \mathrm{Ker}(A)^\perp \cap \lambda \partial g(x) \\ &= \partial_x f(Ax) \cap \lambda \partial g_{\mathrm{Ker}(A)}(x)\end{aligned}$$

where we used the fact $\partial g_{\mathrm{Ker}(A)}(x) = \partial g(x) \cap \mathrm{Ker}(A)^\perp$. In addition, for any x,

$$\partial_x f(Ax) \cap \lambda \partial g_{\mathrm{Ker}(A)}(x) = \partial_x f(Ax_{\mathrm{ker}}) \cap \lambda \partial g_{\mathrm{Ker}(A)}(x_{\mathrm{ker}}) = \partial_x f(Ax_{\mathrm{ker}}) \cap \lambda \partial g(x_{\mathrm{ker}})$$

where $x_{\mathrm{ker}} \in S \cap (x + \mathrm{Ker}(A))$. Hence, together with Lemma 4.6 for $\lambda \neq 0$, we further obtain

$$\partial_x f(Ax) \cap \lambda \partial_x g(x) \neq \varnothing \Longrightarrow \partial_x f(Ax) \cap \lambda \partial g_{\mathrm{Ker}(A)}(x) \neq \varnothing$$

$$\partial_x f(Ax) \cap \lambda \partial g_{\mathrm{Ker}(A)}(x) \neq \varnothing \Longrightarrow \partial_x f(Ax_{\mathrm{ker}}) \cap \lambda \partial g(x_{\mathrm{ker}}) \neq \varnothing$$

implying that λ is a nonzero eigenvalue of $(\mathcal{M}_{A^\top} f, g)$ if and only if λ is a nonzero eigenvalue of $(\mathcal{M}_{A^\top} f, \mathcal{M}_{A^\top} \Pi_A g)$, with the same multiplicity.

Finally, we need to show that the variational eigenvalues are preserved. For any subset $S \subset g^{-1}(1)$ realizing $\lambda_k(\mathcal{M}_{A^\top} f, g)$ with $\mathrm{genus}(S) \geq k$, we have $\lambda_k(\mathcal{M}_{A^\top} f, g) \geq f(Ax)/g(x) = f(Ax)$, $\forall x \in S$. Let $\mathbb{S}$ be the unit sphere in the linear subspace $\mathrm{Ker}(A^\top)$ centered at the origin 0. Let $\xi : \mathbb{R}^n \to \mathrm{Ker}(A^\top)^\perp$ be a linear map induced by $\xi(x) = (A^\top)^{-1}(x) \cap \mathrm{Ker}(A^\top)^\perp$. Clearly, ξ is an odd continuous map. Define the geometric join

$$W := \xi(\partial g(S) \cap \mathrm{Ker}(A)^\perp) * \mathbb{S}.$$

For any $y \in W$, there exist $0 \leq t \leq 1$, $u \in \xi(\partial g(S) \cap \mathrm{Ker}(A)^\perp)$ and $-v \in (1-t)\mathbb{S}$, such that $y = tu - v$. Thus, $A^\top u \in \partial g(S) \cap \mathrm{Ker}(A)^\perp$. So, there exists $x \in S$ such that $A^\top u \in \partial g(x)$. Therefore, $x \in \partial \mathcal{N} g(A^\top u) = \partial \mathcal{N} g(tA^\top u)$, $\mathcal{N} g(tA^\top u) = t\mathcal{N} g(A^\top u) = t$ and $\mathcal{N} g(A^\top v) = 0$. Note that $\langle u, Ax\rangle = \langle A^\top u, x\rangle = g(x) = 1$, which implies $x \notin \mathrm{Ker}(A)$. Again, by the definition of subgradients, it follows from $x \in \partial \mathcal{N} g(tA^\top u)$ that $\mathcal{N} g(A^\top v) - \mathcal{N} g(tA^\top u) \geq \langle A^\top v - tA^\top u, x\rangle$. Then, we have

$$\begin{aligned}\mathcal{N} f(tu - v) &= \sup_{z \neq 0} \frac{\langle tu - v, z\rangle}{f(z)} \geq \frac{\langle tu - v, Ax\rangle}{f(Ax)} = \frac{\langle tA^\top u - A^\top v, x\rangle}{f(Ax)} \\ &\geq \frac{\mathcal{N} g(tA^\top u) - \mathcal{N} g(A^\top v)}{f(Ax)} = \frac{t}{f(Ax)} \geq \frac{t}{\lambda_k(\mathcal{M}_{A^\top} f, g)}\end{aligned}$$

and $\mathcal{N} g(A^\top(tu - v)) = \mathcal{N} g(tA^\top u) = t$. Accordingly, we obtain

$$\frac{\mathcal{N} g(A^\top(tu - v))}{\mathcal{N} f(tu - v)} \leq \lambda_k(\mathcal{M}_{A^\top} f, g)$$

and then

$$\sup_{y \in W} \frac{\mathcal{N} g(A^\top y)}{\mathcal{N} f(y)} \leq \lambda_k(\mathcal{M}_{A^\top} f, g).$$

Let $d_A = \dim \mathrm{Ker}(A)$ and $d_{A^\top} = \dim \mathrm{Ker}(A^\top)$. We estimate the Krasnoselskii genus of W as

$$\mathrm{genus}(W) = \mathrm{genus}(\xi(\partial g(S) \cap \mathrm{Ker}(A)^\perp)) + \mathrm{genus}(\mathbb{S})$$

$$
\begin{aligned}
&\geq \text{genus}(\partial g(S) \cap \text{Ker}(A)^{\perp}) + \text{dim Ker}(A^{\top}) \\
&\geq \text{genus}(\partial g(S)) - \text{dim Ker}(A) + \text{dim Ker}(A^{\top}) \\
&\geq \text{genus}(S) - d_A + d_{A^{\top}} \geq k - d_A + d_{A^{\top}}
\end{aligned}
$$

where the first equality uses the fact that $\xi(\partial g(S) \cap \text{Ker}(A)^{\perp}) \subset \text{Ker}(A^{\top})^{\perp}$ and $\mathbb{S}$ is the unit sphere of the linear subspace $\text{Ker}(A^{\top})$. Therefore, we obtain that

$$
\lambda_{k-d_A+d_{A^{\top}}}(\mathcal{M}_A\mathcal{N}g, \mathcal{N}f) \leq \lambda_k(\mathcal{M}_{A^{\top}}f, g)\,. \tag{4.1}
$$

As the converse holds by a similar argument, we conclude that the identity holds in (4.1).

To conclude, we prove that $\lambda_k(\mathcal{M}_{A^{\top}}f, g) = \lambda_{k-d_A}(\mathcal{M}_{A^{\top}}f, \mathcal{M}_{A^{\top}}\Pi_A g)$. Let again S be defined as in Lemma 4.6. We know that $\text{genus}(S) = n - \dim \text{Ker}(A)$. For any W with $\text{genus}(W) > \dim \text{Ker}(A)$, $\text{genus}(W \cap S) \geq \text{genus}(W) - \dim \text{Ker}(A)$. It is not difficult to check that

$$
\begin{aligned}
\lambda_k(\mathcal{M}_{A^{\top}}f, g) &= \inf_{\text{genus}(W)\geq k} \sup_{x\in W} \frac{f(Ax)}{g(x)} \geq \inf_{\text{genus}(W)\geq k} \sup_{x\in W\cap S} \frac{f(Ax)}{g(x)} \\
&= \inf_{\text{genus}(W')\geq k-d_A, W'\subset S} \sup_{x\in W'} \frac{f(Ax)}{g(x)} \\
&= \inf_{\text{genus}(W')\geq k-d_A, W'\cap \text{Ker}(A)=\varnothing} \sup_{x\in W'} \frac{f(Ax)}{g_{\text{Ker}(A)}(x)} \\
&= \lambda_{k-d_A}(\mathcal{M}_{A^{\top}}f, \mathcal{M}_{A^{\top}}\Pi_A g).
\end{aligned}
$$

On the other hand, for any W realizing $\lambda_{k-d_A}(\mathcal{M}_{A^{\top}}f, \mathcal{M}_{A^{\top}}\Pi_A g)$, there is an eigenvector in W, and every nontrivial eigenvector lies in S. Fix such a subset W, consider a family of subsets defined by $\{(W\cap S)*(r\mathbb{S})\}_{r>1}$, where $r\mathbb{S}$ is the sphere with radius r in the linear subspace $\text{Ker}(A)$ centered at the origin 0. It is easy to check that $\text{genus}(W\cap S)*(r\mathbb{S}) = \text{genus}(W\cap S)+\text{genus}(r\mathbb{S}) \geq k - d_A + d_A = k$ for sufficiently large r. And one can verify that

$$
\lim_{r\to+\infty} \sup_{x\in(W\cap S)*(r\mathbb{S})} \frac{f(Ax)}{g(x)} = \sup_{x\in W\cap S} \frac{f(Ax)}{g(x)}
$$

which implies $\lambda_k(\mathcal{M}_{A^{\top}}f, g) \leq \lambda_{k-d_A}(\mathcal{M}_{A^{\top}}f, \mathcal{M}_{A^{\top}}\Pi_A g)$. Consequently, the proof of $\lambda_k(\mathcal{M}_{A^{\top}}f, g) = \lambda_{k-d_A}(\mathcal{M}_{A^{\top}}f, \mathcal{M}_{A^{\top}}\Pi_A g)$ is completed and we can conclude. □

5 Legendre and Polarity transforms

In this section, we use the Legendre and the Polarity transform to provide nonlinear spectral duality results for function pairs in CH_p^+ with $p \geq 1$, and not just $p = 1$.

First, we recall the notion of the two transforms for general functions. The Legendre transform of a function $f : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$
\hat{\mathcal{L}}f(x) := \sup_{y\in\mathbb{R}^n} \langle x, y\rangle - f(y) = \inf\{s \in \mathbb{R} : \langle x, y\rangle \leq f(y) + s, \forall y \in \mathbb{R}^n\},
$$

and the Polarity transform of a function $f : \mathbb{R}^n \to [0, +\infty)$ is defined as

$$
\hat{\mathcal{P}}f(x) := \sup_{y:f(y)>0} \frac{\langle x, y\rangle - 1}{f(y)} = \inf\{c \in \mathbb{R} : \langle x, y\rangle \leq cf(y) + 1, \forall y \in \mathbb{R}^n\}.
$$

It should be noted that the polarity transform for functions was originally introduced by Rockafellar in his celebrated book [45, Chapter 15], and has been recently "rediscovered" and further investigated in [1, 2].

Similar to the norm-like dual, we now consider a modified version of the two transforms that is better suited for the function family $CH_p^+(\mathbb{R}^n)$. Precisely, we define the Legendre and the Polarity transforms of a function $f \in CH_p^+(\mathbb{R}^n)$ respectively as

$$\mathcal{L}f(x) := \sup_{y \perp f^{-1}(0)} \langle x, y\rangle - f(y) = \inf\{s \in \mathbb{R} : \langle x, y\rangle \le f(y) + s, \forall y \in (f^{-1}(0))^{\perp}\}$$

$$\mathcal{P}f(x) := \sup_{y \perp f^{-1}(0)} \frac{\langle x, y\rangle - 1}{f(y)} = \inf\{c \in \mathbb{R} : \langle x, y\rangle \le cf(y) + 1, \forall y \in (f^{-1}(0))^{\perp}\}.$$

Just like the norm dual operator, we note that $\mathcal{L}f(x) = \hat{\mathcal{L}}f(x - \pi_f x)$ and $\mathcal{P}f(x) = \hat{\mathcal{P}}f(x - \pi_f x)$, where P denotes the orthogonal projection onto $\mathrm{Ker}(f)$. We emphasize that, as for the norm-like duality, we use these modified transforms instead of the standard ones because $\hat{\mathcal{L}}f(x) = \hat{\mathcal{P}}f(x) = +\infty$ for $x \notin (\mathrm{Ker}\, f)^{\perp}$ and $f \in CH_p^+(\mathbb{R}^n)$. Nonetheless, it is quite surprising that several of the results of the main theorems in this paper still hold in a certain sense if we use the standard concepts of infimal postcomposition, norm dual, Legendre transform and Polarity transform, instead of our modified versions. For the sake of clarity, we postpone this observation to the discussion in Section 6.

The next two theorems, Theorems 5.1 and 5.2, show spectral invariance under the two duality transforms for pairs of convex p-homogeneous functions $f \in CH_p^+(\mathbb{R}^n)$ and $g \in CH_q^+(\mathbb{R}^n)$, with $p, q \ge 1$. Then, in Theorem 5.7 we will present our main result of this section, which corresponds to the Legendre and polarity transforms' version of the norm-like duality in Theorem 4.1 and Corollary 4.2 from the previous section. In particular, Theorem 5.7 fully characterizes the spectral duality equivalence under the action of $\mathcal{L}$, $\mathcal{P}$, Π_A and $\mathcal{M}_A$, for a function pair $f \in CH_p^+(\mathbb{R}^n)$ and $g \in CH_q^+(\mathbb{R}^n)$, with $p, q \ge 1$. Here, and in the rest of the section, for a $p > 1$ we let p^* be its Hölder conjugate exponent $1/p + 1/p^* = 1$.

Theorem 5.1. *For any $f \in CH_p^+(\mathbb{R}^n)$ and $g \in CH_q^+(\mathbb{R}^n)$ with some $p, q > 1$, the nonzero eigenvalues of (f, g) and $(\mathcal{L}g, \mathcal{L}f)$ coincide up to a power factor. Precisely, for any eigenpair (λ, x) of (f, g) with $\lambda \neq 0$ and $g(x) \neq 0$, and for any $u \in \mathrm{cone}(\partial f(x)) \cap \partial g(x)$, (λ^{p^*-1}, u) is an eigenpair of $(\mathcal{L}g, \mathcal{L}f)$.*

Proof. It is known that ∂f is homogeneous of degree $(p-1)$, and ∂g is homogeneous of degree $(q-1)$. Since (λ, x) is an eigenpair of (f, g) with $\lambda \neq 0$ and $g(x) \neq 0$, the inclusion relation $0 \in \partial f(x) - \lambda\partial g(x)$ implies that $0 = \langle x, \partial f(x)\rangle - \lambda\langle x, \partial g(x)\rangle = pf(x) - \lambda qg(x)$, where we have applied Euler's identity for the p-homogeneous convex function f and the q-homogeneous convex function g to obtain that for any $u \in \partial f(x), v \in \partial g(x)$, $\langle x, u\rangle = pf(x)$ and $\langle x, v\rangle = qg(x)$. Thus, $f(x) > 0$ and $\lambda = pf(x)/qg(x) > 0$. Moreover, there exists $u \in \partial g(x)$ such that $\lambda u \in \partial f(x)$. which implies $u \in \mathrm{cone}(\partial f(x)) \cap \partial g(x) \neq \varnothing$. And for any $u \in \mathrm{cone}(\partial f(x)) \cap \partial g(x)$, there exists $v \in \partial f(x)$ and $\mu \ge 0$ such that $u = \mu v$. If follows from $\langle u, x\rangle = qg(x) \neq 0$ that $u \neq 0$ and hence $\mu > 0$. On the one hand, $\langle \lambda\mu v, x\rangle = \lambda\mu\langle v, x\rangle = \lambda\mu pf(x)$, and on the other hand, $\langle \lambda\mu v, x\rangle = \langle \lambda u, x\rangle = pf(x) \neq 0$. Thus, $\lambda\mu = 1$ and $v = \lambda u \in \partial f(x)$. Consequently, by the property of Legendre transform, $x \in \partial\mathcal{L}g(u)$ and $x \in \partial\mathcal{L}f(\lambda u) = \lambda^{p^*-1}\partial\mathcal{L}f(u)$. This implies

$$x \in \partial\mathcal{L}g(u) \cap \lambda^{p^*-1}\partial\mathcal{L}f(u) \neq \varnothing$$

which means that (λ^{p^*-1}, u) is an eigenpair of $(\mathcal{L}g, \mathcal{L}f)$. □

Theorem 5.2. *For $f \in CH_p^+(\mathbb{R}^n)$ and $g \in CH_q^+(\mathbb{R}^n)$, the nonzero eigenvalues of (f, g) and $(\mathcal{P}g, \mathcal{P}f)$ coincide up to a scaling factor. Precisely, for any eigenpair (λ, x) of (f, g) with $\lambda \neq 0$ and $g(x) \neq 0$, and for any $u \in \mathrm{cone}(\partial f(x)) \cap \partial g(x)$, $(\alpha\lambda, u)$ is an eigenpair of $(\mathcal{P}g, \mathcal{P}f)$, with $\alpha = (\frac{p}{q})^{p-2} \frac{(q-1)^{q-1}}{(p-1)^{p-1}}$.*

Proof. Let (λ, x) be an eigenpair of (f, g) with $\lambda \neq 0$ and $x \neq 0$. It is easy to see that $f(x) = 0$ $\Leftrightarrow g(x) = 0$, and in this case, we have $\mathcal{P}f(x) = 0$, $\mathcal{P}g(x) = 0$, and $0 \in \partial\mathcal{P}f(x) \cap \partial\mathcal{P}g(x)$ which implies $0 \in \partial\mathcal{P}g(x) - \lambda\partial\mathcal{P}f(x)$. Hence, (λ, x) is also an eigenpair of $(\mathcal{P}g, \mathcal{P}f)$. In fact, from this proof, we obtain that if $f^{-1}(0) \cap g^{-1}(0) \neq \{0\}$, then the spectra of (f, g) and $(\mathcal{P}g, \mathcal{P}f)$ are $\mathbb{R}$. Therefore, without loss of generality, we assume that $f^{-1}(0) \cap g^{-1}(0) = \{0\}$, $g(x) = 1$ and $f(x) = q\lambda/p \neq 0$. Thus, there exists $u \in \partial g(x)$ such that $\lambda u \in \partial f(x)$. Clearly, $u \neq 0$. It follows from the fact $\partial g(x) \subset (g^{-1}(0))^\perp = ((\mathcal{P}g)^{-1}(0))^\perp$ that $\mathcal{P}g(u) \neq 0$. Moreover, we have $\langle u, x\rangle = qg(x) = q$ by Euler's identity, and $\langle u, x'\rangle - q = \langle u, x' - x\rangle \leq g(x') - g(x) = g(x') - 1$, $\forall x' \in \mathbb{R}^n$ by the definition of the subgradient. So, $\langle u, x'\rangle \leq g(x') + q - 1$, and thus $\langle u, \frac{1}{q-1}x'\rangle \leq (q-1)^{q-1} g(\frac{1}{q-1}x') + 1$. Accordingly, $\mathcal{P}g(u) = (q-1)^{q-1}$, and for any $u' \in \mathbb{R}^n$,

$$\begin{aligned}
\langle u' - u, (q-1)^{q-1}x\rangle &= (q-1)^q \langle u' - u, (q-1)^{-1}x\rangle \\
&= (q-1)^q (\langle u', (q-1)^{-1}x\rangle - q(q-1)^{-1}) \\
&\leq (q-1)^q (\mathcal{P}g(u') g((q-1)^{-1}x) + 1 - q(q-1)^{-1}) \\
&= (q-1)^q ((q-1)^{-q}\mathcal{P}g(u') - (q-1)^{-1}) \\
&= \mathcal{P}g(u') - (q-1)^{q-1} = \mathcal{P}g(u') - \mathcal{P}g(u)
\end{aligned}$$

which implies that $(q-1)^{q-1}x \in \partial\mathcal{P}g(u)$. We can similarly derive that $\mathcal{P}f(u) = \frac{1}{\lambda}(\frac{q(p-1))}{p})^{p-1}$ and $(\frac{p}{q})^{2-p}(p-1)^{p-1}\frac{1}{\lambda}x \in \partial\mathcal{P}f(u)$. Therefore, $(\lambda(\frac{p}{q})^{p-2}\frac{(q-1)^{q-1}}{(p-1)^{p-1}}, u)$ is an eigenpair of $(\mathcal{P}g, \mathcal{P}f)$. □

Remark 5.3. *Although Artstein-Avidan and Rubinstein [2] introduce a polar subdifferential map which possesses very nice properties for the polarity transform, it is still necessary to use the usual subdifferential in Theorem 5.2.*

We would like to point out that for any $f \in CH_p^+(\mathbb{R}^n)$ with $p > 1$, the polar subdifferential $\partial^\circ f$ and the usual subdifferential ∂f satisfy

$$\partial^\circ f(x) = (p^* - 1)\frac{\partial f(x)}{f(x)}, \quad \text{when } x \text{ satisfies } f(x) \neq 0.$$

This property can be verified directly by [2, Corollary 3.4] and the basic properties of polarity transform in Rockafellar's celebrated book [45, Chapter 15].

Before presenting our main and final result of this section, we need a number of relevant preliminary observations and results. First, we show in the next proposition that both Legendre and polarity transforms are directly related to the norm-like transform of Section 3. Then, in Proposition 5.5 we show how the eigenpairs of (f, g) change when f and g are raised to some power. These two results will allow us to work on the spectral duality for Legendre and polarity transforms for p-homogeneous convex functions by means of the results previously shown for the case of norm-like duality for one-homogeneous convex functions.

Proposition 5.4. *Given $p \geq 1$ and $r \geq 1/p$, for any nonnegative p-homogeneous function f, f is convex if and only if f^r is convex. And, if f is nonnegative p-homogeneous and convex with $p > 1$, then*

$$\mathcal{L}f = \frac{p-1}{p^{p^*}}(\mathcal{N}f^{\frac{1}{p}})^{p^*} \text{ and } \mathcal{P}f = \frac{(p-1)^{p-1}}{p^p}(\mathcal{N}f^{\frac{1}{p}})^p. \tag{5.1}$$

If $f_p \in CH_p^+(\mathbb{R}^n)$ with $p > 1$ and $\{f_p\}$ Gamma-converges to f as p tends to 1, then we have

$$\mathcal{N}f = \lim_{p\to 1^+}(\mathcal{L}f_p)^{\frac{1}{p^*}} = \lim_{p\to 1^+}(\mathcal{P}f_p)^{\frac{1}{p}} = \lim_{p\to 1^+}\mathcal{P}f_p.$$

Proof. The first argument is equivalent to the statement that for any nonnegative one-homogeneous function f, and $p \geq 1$, f is convex $\Leftrightarrow$ f^p is convex. To show this property, first note that the direction that the convexity of f implies the convexity of f^p is easy since $t \mapsto t^p$ is increasing and convex for $t \in [0, \infty)$. We now show that the convexity of f^p implies the convexity of f. For any x, y with $f(x), f(y) > 0$, letting $C = tf(x) + (1-t)f(y)$, we have

$$\begin{aligned}\frac{f^p(tx+(1-t)y)}{C^p} &= f^p\Big(\frac{tf(x)}{C}\frac{x}{f(x)} + \frac{(1-t)f(y)}{C}\frac{y}{f(y)}\Big)\\ &\leq \frac{tf(x)}{C}f^p\Big(\frac{x}{f(x)}\Big) + \frac{(1-t)f(y)}{C}f^p\Big(\frac{y}{f(y)}\Big)\\ &= \frac{tf(x)}{C} + \frac{(1-t)f(y)}{C} = 1\end{aligned}$$

which yields $f(tx+(1-t)y) \leq C$. As the case of $f(x)f(y) = 0$ is straightforward, we obtain the convexity of f. The equalities shown in (5.1) are presented in [45, Chapter 15]. As for the final statement, note that if f_p Gamma-converges to f, then $f_p^{1/p}$ also Gamma-converges to f. And then, by the property of Gamma-convergence and uniform coercivity of f_p, $\mathcal{N}f_p^{1/p}(x)$ converges to $\mathcal{N}f$ as p tends to 1. Thus, by (5.1),

$$(\mathcal{L}f_p)^{\frac{1}{p^*}} = \left(\frac{p-1}{p^{p^*}}(\mathcal{N}f_p^{\frac{1}{p}})^{p^*}\right)^{\frac{1}{p^*}} = \frac{(p-1)^{\frac{1}{p^*}}}{p}\mathcal{N}f_p^{\frac{1}{p}} \to \mathcal{N}f$$

and

$$(\mathcal{P}f_p)^{\frac{1}{p}} = \left(\frac{(p-1)^{p-1}}{p^p}(\mathcal{N}f^{\frac{1}{p}})^p\right)^{\frac{1}{p}} = \frac{(p-1)^{\frac{p-1}{p}}}{p}\mathcal{N}f_p^{\frac{1}{p}} \to \mathcal{N}f$$

as p tends to 1. Clearly, $\mathcal{P}f_p \to \mathcal{N}f$, $p \to 1^+$. □

Proposition 5.5. *For $\lambda \neq 0$, (λ, x) is an eigenpair of (f, g) if and only if $(\frac{apf^{p-1}(x)}{bqg^{q-1}(x)}\lambda, x)$ is an eigenpair of (af^p, bg^q). Moreover, the eigenpairs of (f, g) and (af^p, bg^q) have a completely equivalent one-to-one correspondence.*

Proof. If (μ, x) is an eigenpair of (af^p, bg^q) where $\mu \neq 0$, then $f(x) > 0$ and $g(x) > 0$, and

$$\begin{aligned}0 &\in \partial af^p(x) - \mu\partial bg^q(x) = apf^{p-1}(x)\partial f(x) - \mu bqg^{q-1}(x)\partial g(x)\\ &= apf^{p-1}(x)\left(\partial f(x) - \frac{\mu bqg^{q-1}(x)}{apf^{p-1}(x)}\partial g(x)\right)\end{aligned}$$

which implies that $\frac{\mu b q g^{q-1}(x)}{apf^{p-1}(x)}$ is an eigenvalue of (f,g). Conversely, it is easy to see that if (λ,x) is an eigenpair of (f,g) with $\lambda\neq 0$, then $(\frac{apf^{p-1}(x)}{bqg^{q-1}(x)}\lambda,x)$ is an eigenpair of (af^p,bg^q).

In addition, it is clear that $(0,x)$ is an eigenpair of (f,g) if and only if $(0,x)$ is an eigenpair of (af^p,bg^q). □

Finally, we point out that we need to be careful with the case $p\neq q$ when dealing with multiplicities and variational eigenvalues. In that case, in fact, $r=f/g$ is not scale-invariant and the eigenvalues of (f,g) and their multiplicities have degenerate properties. Precisely,

Lemma 5.6. *Given $f\in CH_p^+(\mathbb{R}^n)$ and $g\in CH_q^+(\mathbb{R}^n)$ with $p,q\geq 1$, for any $\lambda\geq 0$ and $t>0$, there holds $x\in S_\lambda(f,g)$ if and only if $tx\in S_{t^{p-q}\lambda}(f,g)$. Moreover, if $p\neq q$, f and g are even, and (f,g) has a nonzero eigenvalue, then the function $\lambda\mapsto \mathrm{mult}_{f,g}(\lambda)$ is constant on $(0,+\infty)$.*

Proof. By $t>0$, the definition of the eigenspace S_λ, and the homogeneity of ∂f and ∂g, we have

$$\begin{aligned}x\in S_\lambda(f,g)&\iff 0\in\partial f(x)-\lambda\partial g(x)\iff 0\in t^{p-1}\partial f(x)-\lambda t^{p-1}\partial g(x)\\&\iff 0\in\partial f(tx)-\lambda t^{p-q}\partial g(tx)\iff tx\in S_{t^{p-q}\lambda}(f,g).\end{aligned}$$

If $p\neq q$ and (f,g) has a positive eigenvalue $\hat\lambda>0$, then for any $t>0$, $t^{p-q}\hat\lambda$ is also an eigenvalue of (f,g), that is, all positive numbers are eigenvalues of (f,g). Note that for any $t>0$, the map $\varphi_t:\mathbb{R}^n\to\mathbb{R}^n$ defined by $\varphi_t(x)=tx$ is an odd homeomorphism. Then, for any $\lambda>0$, it follows from $\varphi_t(S_\lambda(f,g))=S_{t^{p-q}\lambda}(f,g)$ and the homeomorphism-invariance of Krasnoselskii genus that

$$\mathrm{mult}_{f,g}(\lambda)=\mathrm{genus}(S_\lambda(f,g))=\mathrm{genus}(S_{t^{p-q}\lambda}(f,g))=\mathrm{mult}_{f,g}(t^{p-q}\lambda).$$

By the arbitrariness of $\lambda>0$ and $t>0$, the multiplicity function $\mathrm{mult}_{f,g}(\lambda)$ is independent of $\lambda>0$. □

Thus, when $p\neq q$, the (variational) eigenvalues of (f,g) change when the corresponding eigenvector is scaled, and their multiplicities are constant. To overcome this issue and have a meaningful definition of variational eigenvalues also for the $p\neq q$ case, it is useful to restrict the variational eigenvalues to suitable centrally symmetric convex surfaces. In particular, we note that for $p=q$ we have $f(x)/g(x)=f(x/g(x)^{1/p})$ and $g(x/g(x)^{1/p})=1$ for all $x\neq 0$. Thus, we can recast (2.2) as

$$\lambda_k(f,g)=\inf_{\substack{\mathrm{genus}(S)\geq k\\ S\subset g^{-1}(1)}}\sup_{x\in S}r(x)=\inf_{\substack{\mathrm{genus}(S)\geq k\\ S\subset g^{-1}(1)}}\sup_{x\in S}f(x)\tag{5.2}$$

i.e., for $p=q$ the k-th variational eigenvalue equals the k-th min-max critical value of f restricted to the centrally symmetric convex hypersurface $g^{-1}(1)$.

By constraining the eigenvalues to $g^{-1}(1)$, the next theorem provides the Legendre and polarity transforms' version of Theorem 4.1 and Corollary 4.2, i.e., it presents the overall spectral duality equivalence between Frenchel duality, polarity transform, and linear transformations.

Before going into details, it is useful to first consider in what sense we can claim that two eigenproblems are equivalent. Strictly speaking, given

$$f_j,g_j\in\bigcup_{p\geq 1}CH_p^+(\mathbb{R}^{n_j}),\ j=1,2,$$

the eigenproblems of the function pairs (f_1,g_1) and (f_2,g_2) are "*essentially equivalent*" if they have the following properties: There exist factors $a,b>0$ such that

1. for any $\lambda > 0$, λ is an eigenvalue of $(f_1, g_1) \Longleftrightarrow a\lambda^b$ is an eigenpair of (f_2, g_2);
2. the multiplicity of λ associated to (f_1, g_1) equals the multiplicity of $a\lambda^b$ associated to (f_2, g_2);
3. $\lambda_{n_1-j}(f_2, g_2) = a\big(\lambda_{n_2-j}(f_1, g_1)\big)^b$ for any $j = 0, 1, \cdots, \min\{n_1, n_2\} - 1$.

Theorem 5.7. *For any $f \in CH_p^+(\mathbb{R}^m)$, $g \in CH_q^+(\mathbb{R}^n)$, and any linear map $A : \mathbb{R}^n \to \mathbb{R}^m$, the eigenproblems of the following nine function pairs are essentially equivalent:*

- *$(\mathcal{M}_{A^\top} f, g)$, where $\mathcal{M}_{A^\top} f$ is defined as the composition of f and the linear map A, that is, $\mathcal{M}_{A^\top} f(x) = f(Ax)$, see Subsection 3.1*
- *$(f, \Pi_A g)$, where $\Pi_A g$ is defined by $\Pi_A g(x) = \inf_{y \in A^{-1}(\pi_A x)} g(\pi_A x)$ and π_A is the orthogonal projection onto the orthogonal complement of $\operatorname{Ker} A^\top$ introduced in Subsection 3.1*
- *$\mathcal{L}\Pi_A g, \mathcal{L}f)$, where $\mathcal{L}\Pi_A g$ indicates the Legendre dual of $\Pi_A g$*
- *$(\mathcal{P}\Pi_A g, \mathcal{P}f)$, where $\mathcal{P}\Pi_A g$ indicates the polarity dual of $\Pi_A g$*
- *$(\mathcal{L}g, \mathcal{L}\mathcal{M}_{A^\top} f)$, where $\mathcal{L}\mathcal{M}_{A^\top} f$ denotes the Legendre dual of $\mathcal{M}_{A^\top} f$*
- *$(\mathcal{P}g, \mathcal{P}\mathcal{M}_{A^\top} f)$, where $\mathcal{P}\mathcal{M}_{A^\top} f$ denotes the polarity dual of $\mathcal{M}_{A^\top} f$*
- *$(\mathcal{M}_{A^\top} f, \mathcal{M}_{A^\top}\Pi_A g)$, where $\mathcal{M}_{A^\top}\Pi_A g$ denotes the composition of $\Pi_A g$ and A*
- *$(\mathcal{L}\mathcal{M}_{A^\top}\Pi_A g, \mathcal{L}\mathcal{M}_{A^\top} f)$, where $\mathcal{L}\mathcal{M}_{A^\top}\Pi_A g$ is the Legendre dual of $\mathcal{M}_{A^\top}\Pi_A g$*
- *$(\mathcal{P}\mathcal{M}_{A^\top}\Pi_A g, \mathcal{P}\mathcal{M}_{A^\top} f)$, where $\mathcal{P}\mathcal{M}_{A^\top}\Pi_A g$ is the polarity dual of $\mathcal{M}_{A^\top}\Pi_A g$*

Proof. Note that "essential equivalence" is an equivalence relation. The proof is then divided into the following eight essential equivalences, displayed in the following diagram:

$$
\begin{array}{ccccc}
(\mathcal{L}\mathcal{M}_{A^\top}\Pi_A g, \mathcal{L}\mathcal{M}_{A^\top} f) & \Longleftrightarrow & (\mathcal{M}_{A^\top} f, \mathcal{M}_{A^\top}\Pi_A g) & \Longleftrightarrow & (\mathcal{P}\mathcal{M}_{A^\top}\Pi_A g, \mathcal{P}\mathcal{M}_{A^\top} f) \\
 & & \Updownarrow & & \\
(\mathcal{L}g, \mathcal{L}\mathcal{M}_{A^\top} f) & \Longleftrightarrow & (\mathcal{M}_{A^\top} f, g) & \Longleftrightarrow & (\mathcal{P}g, \mathcal{P}\mathcal{M}_{A^\top} f) \\
 & & \Updownarrow & & \\
(\mathcal{L}\Pi_A g, \mathcal{L}f) & \Longleftrightarrow & (f, \Pi_A g) & \Longleftrightarrow & (\mathcal{P}\Pi_A g, \mathcal{P}f)
\end{array}
$$

where the notation $(f, g) \Longleftrightarrow (f', g')$ indicates that the eigenproblems of (f, g) and (f', g') are essentially equivalent, as per the definition before Theorem 5.7.

Theorems 5.1 and 5.2 imply that the nonzero spectra of (f, g), $(\mathcal{L}g, \mathcal{L}f)$ and $(\mathcal{P}g, \mathcal{P}f)$ coincide up to some scaling or power factors. For any $f \in CH_p^+(\mathbb{R}^n)$, let $\tilde{f} = f^{\frac{1}{p}}$. Then, $\tilde{f} \in CH_1^+(\mathbb{R}^n)$ and (5.1) in Proposition 5.4 implies that

$$\mathcal{L}f = l_p(\mathcal{N}\tilde{f})^{p^*} \text{ and } \mathcal{P}f = a_p(\mathcal{N}\tilde{f})^p$$

where $l_p = \frac{p-1}{p^{p^*}}$ and $a_p = \frac{(p-1)^{p-1}}{p^p}$ are constants. Then, by Proposition 5.5, the eigenvalue problems of (f, g), $(\mathcal{L}g, \mathcal{L}f)$ and $(\mathcal{P}g, \mathcal{P}f)$ can be equivalently reduced to that of $(\tilde{f}, \tilde{g})$ and

$(\mathcal{N}\tilde{g}, \mathcal{N}\tilde{f})$ up to some scaling factors. It follows from Theorem 4.1 that the spectra of $(\tilde{f}, \tilde{g})$ and $(\mathcal{N}\tilde{g}, \mathcal{N}\tilde{f})$ coincide exactly, and hence the eigenvalue problems of (f, g), $(\mathcal{L}g, \mathcal{L}f)$ and $(\mathcal{P}g, \mathcal{P}f)$ are strongly equivalent.

Moreover, according to Theorem 4.1 and Corollary 4.2, we have the following strong equivalences regarding norm-like duality:

$$\begin{array}{ccc}
(\mathcal{M}_{A^\top}\tilde{f}, \mathcal{M}_{A^\top}\Pi_A\tilde{g}) & \xLeftrightarrow{\text{norm-like dual}} & (\mathcal{N}\mathcal{M}_{A^\top}\Pi_A\tilde{g}, \mathcal{N}\mathcal{M}_{A^\top}\tilde{f}) \\
\Updownarrow & & \\
(\mathcal{M}_{A^\top}\tilde{f}, \tilde{g}) & \xLeftrightarrow{\text{norm-like dual}} & (\mathcal{N}\tilde{g}, \mathcal{N}\mathcal{M}_{A^\top}\tilde{f}) \\
\Updownarrow & & \\
(\tilde{f}, \Pi_A\tilde{g}) & \xLeftrightarrow{\text{norm-like dual}} & (\mathcal{N}\Pi_A\tilde{g}, \mathcal{N}\tilde{f})
\end{array}$$

Therefore, by Propositions 5.4 and 5.5, the strong equivalences among the nine eigenvalue problems shown in the diagram in the statement are established. □

Remark 5.8. *According to Proposition 3.3, we can replace* Π_A *by* $\mathcal{N}\mathcal{M}_A\mathcal{N}$ *in Theorem 5.7, if we have the additional assumption that* $\operatorname{Ker} A \supset f^{-1}(0)$.

6 Spectral duality for standard duality transforms

While in many applications (see also next Section 8) it is useful to consider eigenvalue problems with function pairs that have a linear kernel and whose dual is not infinity, in the field of convex analysis or convex geometry it is frequent to use the standard version of the definitions of duality and infimal postcomposition $\hat{\Pi}_A f(x) := A \triangleright f(x)$. Note that if we use the latter in place of Π_A, we can for example remove the condition $f^{-1}(0) \subset \operatorname{Ker}(A)$ in Proposition 3.3, that is, for any $f \in CH_1^+(\mathbb{R}^n)$, $\forall x \in \mathbb{R}^n$, we have $\hat{\Pi}_A f(x) = \hat{\mathcal{N}}\mathcal{M}_A\hat{\mathcal{N}} f(x)$, where $\hat{\mathcal{N}}$ denotes the standard norm dual operator $\hat{\mathcal{N}} f(x) = f^*(x) = \sup_{f(y)\le 1}\langle x, y\rangle$.

Let $CVH_p^+(\mathbb{R}^n)$ be the collection of all convex, positively p-homogeneous functions from $\mathbb{R}^n$ to $[0, +\infty]$. Clearly, $CH_p^+(\mathbb{R}^n) \subsetneqq CVH_p^+(\mathbb{R}^n)$. It is known that $\hat{\mathcal{N}} : CVH_1^+(\mathbb{R}^n) \to CVH_1^+(\mathbb{R}^n)$ and $\hat{\mathcal{P}} : CVH_p^+(\mathbb{R}^n) \to CVH_p^+(\mathbb{R}^n)$ are bijections, whereas $\hat{\mathcal{L}} : CVH_p^+(\mathbb{R}^n) \to CVH_{p^*}^+(\mathbb{R}^n)$ is a bijection when $p > 1$. A straightforward modification of the proofs of Theorems 4.1, 5.1 and 5.2, leads to the following results.

Theorem 6.1. *For any nonconstant* $f, g \in CVH_1^+(\mathbb{R}^n)$*, the nonzero eigenvalues of* (f, g) *and* $(\hat{\mathcal{N}}g, \hat{\mathcal{N}}f)$ *coincide. Precisely, for any eigenpair* (λ, x) *of* (f, g) *with* $\lambda \ne 0$ *and* $g(x) \ne 0$, $\forall u \in \operatorname{cone}(\partial f(x)) \cap \operatorname{cone}(\partial g(x))$, (λ, u) *is an eigenpair of* $(\hat{\mathcal{N}}g, \hat{\mathcal{N}}f)$*. Moreover, if* f *and* g *are even functions, then the variational eigenvalues* (2.2) *of* (f, g) *and* $(\hat{\mathcal{N}}g, \hat{\mathcal{N}}f)$ *as well as their multiplicities coincide exactly.*

Theorem 6.2. *Given* $p, q > 1$*, for any functions* $f \in CVH_p^+(\mathbb{R}^n)$ *and* $g \in CVH_q^+(\mathbb{R}^n)$*, for any eigenpair* (λ, x) *of* (f, g) *with* $\lambda \ne 0$ *and* $g(x) \ne 0$*, and for any* $u \in \operatorname{cone}(\partial f(x)) \cap \partial g(x)$, (λ^{p^*-1}, u) *is an eigenpair of* $(\hat{\mathcal{L}}g, \hat{\mathcal{L}}f)$.

Theorem 6.3. *For any functions* $f \in CVH_p^+(\mathbb{R}^n)$ *and* $g \in CVH_q^+(\mathbb{R}^n)$*, for any eigenpair* (λ, x) *of* (f, g) *with* $\lambda \ne 0$ *and* $g(x) \ne 0$*, and for any* $u \in \operatorname{cone}(\partial f(x)) \cap \partial g(x)$, $((\frac{p}{q})^{p-2}\frac{(q-1)^{q-1}}{(p-1)^{p-1}}\lambda, u)$ *is an eigenpair of* $(\hat{\mathcal{P}}g, \hat{\mathcal{P}}f)$.

7 Infinite dimensional setting

Even though the focus of this work is the real finite-dimensional case, we show in this section that some of the main results presented can be directly extended to real infinite-dimensional spaces. In fact, we believe it is possible to extend most of the results of this paper to the infinite-dimensional case, but that would exceed the scope of this work and is left to future developments.

Let X be a reflexive Banach space over $\mathbb{R}$ and X^* be its dual space, and let $f, g : X \to [0, +\infty)$ be convex one-homogeneous functions in $CH_1^+(X)$. We assume without loss of generality that $\ker(f) = \ker(g) = 0$, which can always be achieved by replacing X with the quotient space $X/(\ker(f) + \ker(g))$. For this setting, the following extension of Theorem 4.1 (P1) holds

Theorem 7.1. *The nonzero eigenvalues of (f, g) and $(\mathcal{N}g, \mathcal{N}f)$ coincide. Furthermore, for any eigenpair (λ, x) of (f, g) with $\lambda \neq 0$, and for any $u \in \operatorname{cone}(\partial f(x)) \cap \operatorname{cone}(\partial g(x))$, (λ, u) is an eigenpair of $(\mathcal{N}g, \mathcal{N}f)$.*

Proof. Since (λ, x) is an eigenpair of (f, g) and $\lambda \neq 0$, there exists $u \in \partial g(x)$ such that $\lambda u \in \partial f(x)$. Note that ∂g and ∂f possess the zero-homogeneity, i.e., $\partial f(tx) = \partial f(x)$ and $\partial g(tx) = \partial g(x)$, $\forall t > 0$, $\forall x \in X$. Hence, without loss of generality, we may assume that the eigenpair (λ, x) further satisfies $g(x) = 1$. Then, it is easy to see $\langle u, x\rangle = g(x) = 1$ and $\lambda = \langle \lambda u, x\rangle = f(x)$.

By definition of subgradient, for any $y \in X$,

$$\langle u, y\rangle - 1 = \langle u, y\rangle - \langle u, x\rangle = \langle u, y - x\rangle \leq g(y) - g(x) = g(y) - 1$$

which yields $\langle u, y\rangle \leq g(y)$, $\forall y \in X$. Thus, it can be checked that

$$\mathcal{N}g(u) = \sup_{y \in X : g(y) = 1} \langle u, y\rangle = 1.$$

Let $J : X \to X^{**}$ be the canonical map, i.e., for each $z \in X$, $\langle Jz, v\rangle = \langle v, z\rangle$, $\forall v \in X^*$. Then, for any $v \in X^*$,

$$\langle Jx, v - u\rangle = \langle v - u, x\rangle = \langle v, x\rangle - 1 \leq \mathcal{N}g(v) - 1 = \mathcal{N}g(v) - \mathcal{N}g(u)$$

which implies that $Jx \in \partial\mathcal{N}g(u)$. For convenience, we identify Jx with x, and simply rewrite $x \in \partial\mathcal{N}g(u)$.

Repeating the above process for f, it follows from $f(\frac{x}{\lambda}) = 1$ and $\lambda u \in \partial f(x) = \partial f(\frac{x}{\lambda})$ that

$$\frac{x}{\lambda} \in \partial\mathcal{N}f(\lambda u) = \partial\mathcal{N}f(u),$$

where we used the zero-homogeneity of ∂f and $\partial\mathcal{N}f$. Accordingly, $x \in \lambda\partial\mathcal{N}f(u)$. In consequence, $x \in \mathcal{N}g(u) \cap \lambda\partial\mathcal{N}f(u)$ meaning that (λ, u) is an eigenpair of $(\mathcal{N}g, \mathcal{N}f)$.

Note that in the above proof, we only require $u \in \partial g(x) \cap \frac{1}{\lambda}\partial f(x)$. By the zero-homogeneity of ∂g and ∂f, we can relax this condition to $u \in \operatorname{cone}(\partial g(x)) \cap \operatorname{cone}(\partial f(x))$. The proof is then completed. ☐

By a similar argument, following Theorem 4.2, we can prove the following property for the Legendre transform in the infinite-dimensional case. We omit the proof for brevity.

Theorem 7.2. *For any $f \in CH_p^+(X)$ and $g \in CH_q^+(X)$ with some $p, q > 1$, the nonzero eigenvalues of (f, g) and $(\mathcal{L}g, \mathcal{L}f)$ coincide up to a power factor. Precisely, for any eigenpair (λ, x) of (f, g) with $\lambda \neq 0$ and $g(x) \neq 0$, and for any $u \in \mathrm{cone}(\partial f(x)) \cap \partial g(x)$, (λ^{p^*-1}, u) is an eigenpair of $(\mathcal{L}g, \mathcal{L}f)$.*

We point out that this infinite-dimensional setting yields quite a significant extension of Proposition 1.8 in the recent work by Bungert and Korolev [10]. In fact, their result shows that if f is a seminorm and $g = \|\cdot\|$ is the prescribed norm of a Banach space X, then the first nontrivial eigenvalue of (f, g) coincides with that of $(\mathcal{N}g, \mathcal{N}f)$. With the help of such result, Bungert and Korolev provide an excellent characterisation of the L^∞ eigenvalue problem [10]. On the other hand, our Theorem 7.1 above shows that, for the more general choice of functions $f, g \in CH_1^+(X)$, all the nontrivial eigenvalues of (f, g) coincide with that of $(\mathcal{N}g, \mathcal{N}f)$, exactly.

8 Examples and applications

We devote this final section to discussing a number of problems where discrete nonlinear eigenvalue problems and the nonlinear spectral duality properties developed in the previous sections can be used in application settings from graph and hypergraph optimization and convex geometry.

8.1 Nonlinear Laplacians on graphs

Let $A \in \mathbb{R}^{n\times m}$ and consider the functions pair $(\|Ax\|_a, \|x\|_b)$, where $\|\cdot\|_a$ and $\|\cdot\|_b$ are vector norms. Nonlinear eigenvalue problems for this type of convex one-homogeneous functions are among the best studied problems in nonlinear spectral theory and arise in a broad range of application settings, including inverse problems in imaging [11,21,28], graph clustering, unsupervised and supervised learning [8,31,38,50], community and core-periphery detection in networks [17,51,54], graph and hypergraph matching [43]. Note that these type of eigenvalue problems are directly connected with generalized operator matrix norms, which coincide with the largest (variational) eigenvalue $\lambda_m(f, g) = \max_{x\neq 0} \|Ax\|_a/\|x\|_b =: \|A\|_{a,b}$.

Note that, if $f = \|\cdot\|_a$ and $g = \|\cdot\|_b$, then this type of eigenvalue problem coincides with the eigenvalue problem for the function pair $(\mathcal{M}_{A^\top} f, g)$. Based on this observation, in this subsection, we review several example eigenvalue problems with a direct application to combinatorial optimization problems on finite graphs and discuss what the various corresponding dual forms are. In particular, we will show that several famous nonlinear graph eigenvalue equations can be recast in various forms, which have the potential to unleash a variety of new results both from the theoretical and computational points of view. In fact, established algorithms for the solution of these eigenvalue problems, such as the inverse iteration [34,35], the family of ratioDCA methods [31,54], or the continuous flow approaches [9,22], can be directly transferred to their dual versions and may exhibit improved convergence properties. Moreover, new relations between the graph structure and the nonlinear eigenpair may be shown, see also [20]. Some of the graph theoretic results presented next are known and properly referenced; others are new and are accompanied by proofs and additional details.

Before proceeding, we briefly recall some useful graph notation and terminology. A finite undirected graph $G = (V, E, w)$ is the pair of vertex (or node) set $V = \{1, \dots, n\}$ and edge set $E \subseteq V \times V$, which we equip with a positive weight function $E \ni ij \mapsto w_{ij} > 0$. Any such graph is uniquely represented by the incidence matrix $K = (\kappa_{e,u}) \in \mathbb{R}^{E\times V}$, which maps any $x \in \mathbb{R}^V$ into the vector with entries $(Kx)_e = \sum_u \kappa_{e,u}x_u = \pm w_{ij}(x_i - x_j)$, where $e = ij \in E$ is the edge connecting nodes i and j. Note that the choice of the sign in $(Kx)_e$ is arbitrary but fixed. Different norms of Kx correspond to different energies on G. For example, $\|Kx\|_1 = \sum_{ij\in E} w_{ij}|x_i - x_j|$ is the graph total variation, $\|Kx\|_2^2 = \sum_{ij\in E} w_{ij}(x_i - x_j)^2$ the electric potential, $\|Kx\|_\infty = \max_{ij\in E} w_{ij}|x_i - x_j|$ the graph node-wise variation.

8.1.1 $(1,1)$-Laplacian: Cheeger constant

Let $G = (V, E, w)$ be a weighted graph and consider the nonlinear eigenvalue problem

$$0 \in \partial \sum_{ij\in E} w_{ij}|x_i - x_j| - \lambda\partial \sum_{i\in V} |x_i|. \tag{8.1}$$

Note that, if $A = K$ is the incidence matrix of G, then (8.1) coincides with the eigenvalue problem for the functions pair $(\mathcal{M}_{A^\top} f, g)$, where $f = \|\cdot\|_1$ and $g = \|\cdot\|_1$ are the standard l^1-norms on $\mathbb{R}^E$ and $\mathbb{R}^V$, respectively. Thus, by Corollary 4.2 it follows that (8.1) is equivalent to the following alternative eigenvalue problems

$$0 \in \partial\|x\|_\infty - \lambda\,\partial \inf_{\sum_{ij\in E:j<i} w_{ij}y_{ij} - \sum_{ij\in E:j>i} w_{ij}y_{ij} = x_i} \|y\|_\infty \tag{8.2}$$

$$0 \in \partial \max_{i\in V} \Big| \sum_{ij\in E:j<i} w_{ij}y_{ij} - \sum_{ij\in E:j>i} w_{ij}y_{ij} \Big| - \lambda\,\partial\|y\|_\infty \tag{8.3}$$

$$0 \in \partial\|y\|_1 - \lambda\,\partial \inf_{x:Kx=y} \|x\|_1 \tag{8.4}$$

$$0 \in \partial \sum_{ij\in E} w_{ij}|x_i - x_j| - \lambda\partial \inf_{t\in\mathbb{R}} \sum_{i\in V} |x_i - t| \tag{8.5}$$

which correspond to the eigenproblems of the pairs $(\mathcal{N}g, \mathcal{N}\mathcal{M}_{A^\top} f)$, $(\mathcal{M}_A\mathcal{N}g, \mathcal{N}f)$, $(f, \Pi_A g)$ and $(\mathcal{M}_{A^\top} f, \mathcal{M}_{A^\top}\Pi_A g)$, respectively. All the above nonlinear eigenvalue problems have the same nonzero eigenvalues (with the same corresponding multiplicities).

The eigenvalue problem (8.1) is known as the 1-Laplacian eigenvalue problem on G. This is one of the key objects of nonlinear spectral graph theory, and many useful properties of the 1-Laplacian are known. For example, when the graph is connected, the smallest positive eigenvalue of (8.1) coincides with the Cheeger isoperimetric constant of G [8, 14]. Moreover, when the graph is a tree, each variational eigenvalue of (8.1) coincides with the k-th Cheeger constant [18, 19, 50]. Precisely, let

$$h_k(G) := \min_{\text{disjoint subsets } V_1,\ldots,V_k\subset V} \max_{1\le i\le k} \frac{\mathrm{vol}(\mathrm{cut}(V_i))}{\mathrm{vol}(V_i)}, \tag{8.6}$$

where $\mathrm{vol}(V_i) = |V_i|$ and $\mathrm{vol}(\mathrm{cut}(V_i)) = \sum_{u\in V_i, v\notin V_i} w_{uv}$ are the (weighted) volumes of V_i and its cut set, respectively. Then, if λ_k is the k-th variational eigenvalue of the 1-Laplacian (8.1), it holds $\lambda_2 = h_2(G)$ and $\lambda_k = h_k(G)$ for $k > 2$ if G is a tree. More in general, we have $\lambda_m \le h_k(G) \le \lambda_k$ for a generic graph G, where m is the largest number of nodal domains of any eigenvector of λ_k [50]. By Theorem 4.1 and Corollary 4.2, the same fundamental graph theoretic properties hold for the variational eigenvalues of each of the nonlinear eigenvalue problems (8.2)–(8.5).

8.1.2 (∞,∞)-Laplacian: graph's diameter

Suppose $G=(V,E,w)$ is a weighted graph and consider the so-called ∞-Laplacian eigenvalue problem:

$$0\in\partial\max_{ij\in E}w_{ij}|x_i-x_j|-\lambda\partial\max_{i\in V}|x_i|. \tag{8.7}$$

Let $A=K=(\kappa_{e,i})$ be the incidence matrix of the graph, and let $f:=\|\cdot\|_\infty$ and $g:=\|\cdot\|_\infty$ be the standard unweighted l^∞-norms on $\mathbb{R}^E$ and $\mathbb{R}^V$, respectively. Then, (8.7) coincides with the nonlinear eigenvalue problem $0\in\partial_x\|Ax\|_\infty-\lambda\partial\|x\|_\infty$, i.e., the eigenvalue problem for the functions pair $(\mathcal{M}_{A^\top}f,g)$. By Corollary 4.2, we obtain several new eigenvalue problems equivalent to the graph ∞-Laplacian:

$$0\in\partial\|x\|_1-\lambda\partial\inf_{y:K^\top y=x}\|y\|_1 \tag{8.8}$$

$$0\in\partial\sum_{i\in V}\Big|\sum_{e\in E}\kappa_{e,i}y_e\Big|-\lambda\partial\|y\|_1 \tag{8.9}$$

$$0\in\partial\|y\|_\infty-\lambda\partial\inf_{x:Kx=y}\|x\|_\infty \tag{8.10}$$

$$0\in\partial\max_{ij\in E}w_{ij}|x_i-x_j|-\lambda\partial\Big\|x-\frac{\max_i x_i+\min_i x_i}{2}\mathbf{1}\Big\|_\infty \tag{8.11}$$

where $\mathbf{1}$ denotes the vector of all ones. We emphasize that the formulation in (8.9) corresponds to a form of 1-Laplacian eigenvalue problem on the dual graph, i.e., the eigenvalue problem for the functions pair $f(x)=\|K^\top y\|_1$ and $g(x)=\|y\|_1$.

When the graph is connected, the variational eigenvalues of (8.7) are related to the graph diameter. More precisely, define a ball $B=B_r(v)\subseteq V$ centered in $v\in V$ and of radius $r=\mathrm{radius}(B)$ as the set $B_r(v)=\{u\in V:\mathrm{dist}(u,v)\le r\}$ where dist is the shortest path distance on G. Two such balls $B=B_r(v)$ and $B'=B_{r'}(v')$ are disjoint if $\mathrm{dist}(v,v')\ge r+r'$. With this notation, it holds

$$\lambda_k\le\min_{\text{disjoint balls }B_1,\dots,B_k\subset V}\ \max_{1\le i\le k}\ \frac{1}{\mathrm{radius}(B_i)}$$

where λ_k is the k-th variational eigenvalue of (8.7). In particular, note that the smallest nonzero variational eigenvalue coincides with $2/\mathrm{diam}(G)$, where $\mathrm{diam}(G):=\max_{i,j\in V}\mathrm{dist}(i,j)$, and dist represents the shortest path distance on G. More precisely, if G has k connected components, $G_1,\cdots,G_k$, then the smallest positive variational eigenvalue coincides with

$$\min_{i=1,\cdots,k}\frac{2}{\mathrm{diam}(G_i)}=\frac{2}{\max\limits_{i=1,\cdots,k}\mathrm{diam}(G_i)}.$$

By Corollary 4.2, all the above properties transfer directly to the nonlinear spectrum of any of the eigenvalue problems (8.8)–(8.11).

Remark 8.1. *The cycle graph C_n is the only graph which is dual to itself, i.e., is such that $K=K^\top$. If we work on a cycle graph, the 1-Laplacian eigenvalue problem* (8.1) *is equivalent to the ∞-Laplacian eigenvalue problem* (8.7)*, via the spectral duality equivalence shown in* (8.9)*. In particular, their k-th variational eigenvalues coincide, and they are bounded by the k-th Cheeger* *constant $h_k(G)$ which is consistent with the reciprocal of the largest radius of any ball in any set of k pairwise disjoint balls inside the cycle graph.*

It is interesting to note that in a Euclidean space, a ball B of radius r satisfies $\frac{\mathrm{vol}(\partial B)}{\mathrm{vol}(B)} \sim \frac{1}{r}$, where ∂B is the boundary of B and $\sim$ denotes here that the two quantities are proportional. As cut *is the graph analogue of the boundary, an interesting open question is whether or not $h_k(G) \sim \frac{1}{r_k(G)}$ for a generic graph G, where $r_k(G)$ denotes the largest radius of any ball in any set of k pairwise disjoint balls in the graph.*

8.1.3 $(1,\infty)$-Laplacian: maxcut and mincut

Suppose that $G = (V, E, w)$ is a weighted graph. Consider the eigenvalue problem for the functions pair $(\mathcal{M}_{A^\top} f, g)$ with $f(x) = \|x\|_1$, $g(x) = \|x\|_\infty$ and $A = K$, namely

$$0 \in \partial \sum_{ij \in E} w_{ij}|x_i - x_j| - \lambda \partial \|x\|_\infty \,. \tag{8.12}$$

By the spectral duality principle in Theorem 4.1 and Corollary 4.2, (8.12) is equivalent to

$$0 \in \partial \|x\|_1 - \lambda \partial \inf_{\sum_{j<i} w_{ij}y_{ij} - \sum_{j>i} w_{ij}y_{ij} = x_i} \|y\|_\infty \tag{8.13}$$

$$0 \in \partial \sum_{i \in V} \Big| \sum_{j<i} w_{ij}y_{ij} - \sum_{j>i} w_{ij}y_{ij} \Big| - \lambda \partial \|y\|_\infty \tag{8.14}$$

$$0 \in \partial \|y\|_1 - \lambda \partial \inf_{x: Kx = y} \|x\|_\infty \tag{8.15}$$

$$0 \in \partial \sum_{ij \in E} w_{ij}|x_i - x_j| - \lambda \partial \Big\| x - \frac{\max_i x_i + \min_i x_i}{2} \mathbf{1} \Big\|_\infty \,. \tag{8.16}$$

It is not difficulty to see that the smallest nonzero variational eigenvalue and the largest variational eigenvalue of (8.12) coincide with the mincut and the maxcut values of G, respectively defined as

$$\mathrm{mincut}(G) = \min_{S \subset V} \mathrm{vol}(\mathrm{cut}(S)) \quad \text{and} \quad \mathrm{maxcut}(G) = \max_{S \subset V} \mathrm{vol}(\mathrm{cut}(S)) \,.$$

We also remark that (a) $\mathrm{maxcut}(G)$ is actually equivalent to the largest eigenvalue for the pair $(\|Kx\|_p, \|x\|_\infty)$, for any $1 \le p < \infty$; and (b) when w_{ij} in (8.12) is replaced by the modularity weights $m_{ij} := d_i d_j / (\sum_k d_k) - w_{ij}$, $d_i = \sum_j w_{ij}$, the largest eigenvalue of (8.12) corresponds to the leading community in G, see Theorem 3.7 in [54]. Due to the nonlinear spectral duality principle, the same properties hold for each of the nonlinear eigenvalue problems in (8.13)–(8.16). Moreover, the following relation holds for their k-th variational eigenvalue.

Theorem 8.2. *Let λ_k be the k-th variatonal eigenvalue of the eigenvalue problem* (8.12). *Then*

$$\lambda_k \le \min_{V_1, \cdots, V_k \text{ form a partition of } V} \mathrm{maxcut}(G[V_1, \cdots, V_k])$$

where $\mathrm{maxcut}(G[V_1, \cdots, V_k]) := 2 \max_{S \subset \{1, \cdots, k\}} \sum_{i \in S, j \in V \setminus S} w_{V_i, V_j}$ *denotes the maxcut value of the graph $G[V_1, \cdots, V_k]$, formed by k vertices corresponding to the k sets $V_1, \cdots, V_k$, with edge weights*

$$w_{V_i, V_j} = \sum_{a \in V_i, b \in V_j} w_{ab}, \qquad V_i \ne V_j \,.$$

Proof. For any partition $(V_1,\cdots,V_k)$ of V, denote by 1_{V_i} the indicator vector of V_i. Then 1_{V_1}, $\cdots$, 1_{V_k} are linearly independent. Thus $\mathrm{genus}(\mathrm{span}(1_{V_1},\cdots,1_{V_k}))=k$ and we have

$$\begin{aligned}\lambda_k(\mathcal{M}_{A^\top}f,g) &\le \max_{x\in\mathrm{span}(1_{V_1},\cdots,1_{V_k})}\frac{\sum\limits_{\{i,j\}\in E} w_{ij}|x_i-x_j|}{\|x\|_\infty}\\ &= \max_{(t_1,\cdots,t_k)\in\mathbb{R}^k\setminus\{0\}}\frac{\sum\limits_{1\le i<j\le k} w_{V_i,V_j}|t_i-t_j|}{\max\limits_{i=1,\cdots,k}|t_i|} = 2\max_{S\subset\{1,\cdots,k\}}\sum_{i\in S,j\in V\setminus S} w_{V_i,V_j}.\end{aligned}$$

where the last equality is elementary. □

8.1.4 $(\infty,1)$-Laplacian: spherical packing

Let $G=(V,E,w)$ be a weighted graph. Consider the eigenvalue problem for the functions pair $f(x)=\|Kx\|_\infty$ and $g(x)=\|x\|_1$, namely

$$0\in\partial\max_{ij\in E} w_{ij}|x_i-x_j|-\lambda\partial\|x\|_1\,. \tag{8.17}$$

Then, by Theorem 4.1 and Corollary 4.2, (8.17) is equivalent to

$$0\in\partial\|x\|_\infty-\lambda\partial\inf_{\sum_{j<i} w_{ij}y_{ij}-\sum_{j>i} w_{ij}y_{ij}=x_i,\forall i}\|y\|_1 \tag{8.18}$$

$$0\in\partial\max_{i\in V}\Big|\sum_{j<i} w_{ij}y_{ij}-\sum_{j>i} w_{ij}y_{ij}\Big|-\lambda\partial\|y\|_1 \tag{8.19}$$

$$0\in\partial\|y\|_\infty-\lambda\partial\inf_{x:Kx=y}\|x\|_1 \tag{8.20}$$

$$0\in\partial\max_{ij\in E} w_{ij}|x_i-x_j|-\lambda\partial\Big\|x-\frac{\max_i x_i+\min_i x_i}{2}\mathbf{1}\Big\|_1 \tag{8.21}$$

Moreover, the following result holds for the variational eigenvalue of all the above eigenvalue problems.

Theorem 8.3. *Let λ_k be the k-th variational eigenvalue of the eigenvalue equation* (8.17)*. It holds*

$$\lambda_k\le\min_{\text{disjoint balls } B_1,\cdots,B_k\subset V}\ \max_{1\le i\le k}\frac{1}{\mathrm{size}(B_i)} \tag{8.22}$$

where, for $B=B_r(v)$ we let $\mathrm{size}(B)=\sum_{i=0}^r (r-i)|\{u\in V:\mathrm{dist}(u,v)=i\}|$.

Proof. For a $x\in\mathbb{R}^n$ and a ball B with radius r and centered at the vertex v define $x^B\in\mathbb{R}^n$ by $(x^B)_i=\max\{r-\mathrm{dist}(v,i),0\}$. Then, for any k disjoint balls $B_1,\cdots,B_k\subset V$, $x^{B_1},\cdots,x^{B_k}$ are linearly independent. Thus

$$\mathrm{genus}(\mathrm{span}(x^{B_1},\cdots,x^{B_k}))\ge k$$

and we have

$$\lambda_k\le\max_{x\in\mathrm{span}(x^{B_1},\cdots,x^{B_k})}\frac{\max\limits_{\{i,j\}\in E}|x_i-x_j|}{\|x\|_1}$$

$$\leq \max_{1\leq s\leq k} \frac{\max\limits_{\{i,j\}\in E} |x_i^{B_s} - x_j^{B_s}|}{\|x^{B_s}\|_1} = \max_{1\leq s\leq k} \frac{1}{\mathrm{size}(B_s)}.$$

where the second inequality follows from the fact that the x^{B_j} have disjoint support. By taking the minimum over all possible disjoint balls we obtain (8.22). □

Note that, as a consequence of the above theorem we obtain that the smallest eigenvalue is at most the reciprocal of the size of the largest ball inscribed in the graph.

8.1.5 Hypergraph p-Laplacian

We now extend the concepts and considerations on p-Laplacians to general (oriented) hypergraphs, which cover the entire range $1 \leq p \leq \infty$, that is, also includes the limiting cases $p = 1$ and $p = \infty$, which had been given special treatment above. While some aspects are straightforward, some new phenomena and difficulties arise.

The first issue is the definition of the p-Laplacian, and in fact, several notions of p-Laplacians on hypergraphs have been proposed and investigated in the literature. For instance, the hypergraph p-Laplacians introduced in [32] are related to the Lovász extension. The corresponding functionals for the p-Laplacians in [32] are of the form

$$x \mapsto \sum_{e\in E} \max_{v,v'\in e} |x_v - x_{v'}|^p ,$$

and they are non-smooth in general, even for $p \in (1, +\infty)$.

In this section, we consider the p-Laplacian on oriented hypergraphs as a natural combinatorial realization of a given incidence matrix K. That is, we think of a hypergraph $H = (V, E)$ as an incidence operator K linking the vertex set V and the hyperedge set E, e.g. (for simplicity):

- $V = \{1, \cdots, n\}$ is the set of vertices
- $E = \{e_1, \cdots, e_m\}$ is the set of hyperedges
- an incidence matrix $K \in \{-1, 0, 1\}^{E\times V}$ (i.e., an incidence operator $K : E\times V \to \{-1, 0, 1\}$) induces an orientation on H, and it satisfies $K_{e,v} \in \{-1, 1\}$ iff $v \in e$.

Example 8.4 (signed graphs)**.** *Suppose $E \subset \mathcal{P}(V)$ and $|e| = 2$, $\forall e \in E$. Let the signature $s : E \to \mathbb{R}$ be defined by $s(e) = \prod_{v\in e} K_{e,v}$, $\forall e \in E$. Then we reconstruct the signed graph (V, E, s) from the incidence matrix K.*

Example 8.5 (simplicial complexes)**.** *Let H_k be a hypergraph with $V = \Sigma_k$, $E = \Sigma_{k+1}$ and $K_{\sigma,\tau} = \mathrm{sgn}([\tau], \partial[\sigma])$ represents the relative signature of the incidence pair (σ, τ), $\forall \tau \in \Sigma_k, \sigma \in \Sigma_{k+1}$. Then the orientation on the simplicial complex induces an orientation on the hypergraph $H_k = (\Sigma_k, \Sigma_{k+1})$.*

The incidence-style p-Laplacian on a hypergraph considered in this section is simply defined as $L_p = K^\top \phi_p K$, i.e,

$$L_p x := K^\top(\phi_p(Kx))$$

where $\phi_p(y) := \partial\|y\|_p$ for any vector y. The eigenvalue problem is to find λ and x such that

$$0 \in L_p x - \lambda\phi_p(x).$$

Given a hypergraph $H = (V, E)$, for any $v \in V$, let $\mathbf{star}(v) = \{e \in E : v \in e\}$ be its star neighborhood.

The hypergraph $H^* = (E, \mathbf{V})$ is called the *dual hypergraph* of H, where $\mathcal{V} = \{\mathbf{star}(v) : v \in V\}$. Moreover, we take $K^* := K^\top$ as the incidence matrix of H^*, i.e., $K^*_{\mathbf{star}(v),e} = K^\top_{v,e} = K_{e,v}$, $\forall v \in V$, $\forall e \in E$.

Clearly, the map $v \mapsto \mathbf{star}(v)$ induces an isomorphism between V and $\mathbf{V}$.

Example 8.6. *We can think of a graph $G = (V, E)$ as a hypergraph. Then, the edge p-Laplacian on G is actually defined to be the p-Laplacian on the dual hypergraph of G.*

Now we are ready to state our main theorem on spectral duality for hypergraph p-Laplacians:

Theorem 8.7. *Given a hypergraph $H = (V, E)$ and its dual hypergraph H^*, for any $p \in (1, +\infty)$, we have*

$$\{\lambda^{\frac{1}{p}} : \lambda \textit{ is a positive eigenvalue of } L_p \textit{ on } H\}$$
$$= \{\lambda^{\frac{1}{p^*}} : \lambda \textit{ is a positive eigenvalue of } L_{p^*} \textit{ on } H^*\}$$

where $1/p + 1/p^* = 1$. In the case of $p = 1$ (or $p = +\infty$), we simply have:

$$\{\textit{nonzero eigenvalues of } L_1 \textit{ on } H\} = \{\textit{nonzero eigenvalues of } L_\infty \textit{ on } H^*\}.$$

Moreover, the corresponding multiplicities coincide exactly.

Suppose $|V| = n$ and $|E| = m$. Then, for $i = 0, 1, \cdots, \min\{n, m\} - 1$, and $1 < p < +\infty$,

$$\lambda_{n-i}^{\frac{1}{p}}(L_p(H)) = \lambda_{m-i}^{\frac{1}{p^*}}(L_{p^*}(H^*)),$$

$$\lambda_{n-i}(L_1(H)) = \lambda_{m-i}(L_\infty(H^*)) \textit{ and } \lambda_{n-i}(L_\infty(H)) = \lambda_{m-i}(L_1(H^*)).$$

For convenience, we call $\lambda^{\frac{1}{p}}$ a *normalized eigenvalue* of the p-Laplacian L_p if λ is an eigenvalue of L_p, where $1 \le p < \infty$. And the eigenvalues of the infinite Laplacian L_∞ are also called the *normalized eigenvalues* of L_∞. Hence, the normalized eigenvalues of L_1 (resp., L_∞) agree with the eigenvalues of L_1 (resp., L_∞). Theorem 8.7 indeed shows that the nonzero normalized eigenvalues of L_p on H coincide with the nonzero normalized eigenvalues of L_{p^*} on H^*.

Next, we discuss several consequences and applications of Theorem 8.7.

Applying the above spectral duality theorem (i.e., Theorem 8.7) to the edge p-Laplacian (see Example 8.6), we have that for $1 < p < +\infty$,

$$\{\lambda^{\frac{1}{p}} : \lambda \text{ is a positive eigenvalue of } L_p\} = \{\lambda^{\frac{1}{p^*}} : \lambda \text{ is a positive eigenvalue of } L^E_{p^*}\},$$

and

$$\lambda_{n-i}^{\frac{1}{p}}(L_p) = \lambda_{m-i}^{\frac{1}{p^*}}(L^E_{p^*}), \quad i = 0, 1, \cdots, \min\{n, m\} - 1,$$

in which $L^E_{p^*}$ indicates the edge p^*-Laplacian. For the degenerate cases $p = 1$ and $p = +\infty$,

$$\text{spec}_{\neq 0}(L^E_1) = \text{spec}_{\neq 0}(L_\infty) \text{ and } \text{spec}_{\neq 0}(L^E_\infty) = \text{spec}_{\neq 0}(L_1)$$

where $\text{spec}_{\neq 0}(T)$ stands for the nonzero spectrum of a given operator T.

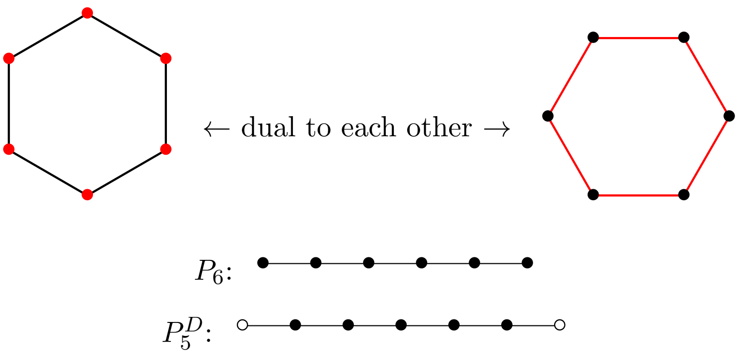

Figure 1: Top: A cycle graph on $n = 6$ nodes is dual to itself. Bottom: P_n and P_{n-1}^D path graphs for $n = 6$, with $\bullet$ representing interior vertices and $\circ$ denoting boundary vertices.

Based on Example 8.5, the incidence-style p-Laplacians on H_k and on its dual H_k^* can be directly used to define the k-th up p-Laplacian and the $(k+1)$-th down p-Laplacian on a simplicial complex Σ. It is then easy to get a nonlinear extension of the identity relating the nonzero spectra of up- and down- Laplacians on Σ, and for simplicity, we omit the specific formulation.

Clearly, Theorem 8.7 can be viewed as a new tool for computing the spectrum of some particular graphs. For example, a cycle graph is dual to itself (see Figure 1). Thus, one obtains that for $1 < p < +\infty$, it holds

$$\{\lambda^{\frac{1}{p}} : \lambda \in \mathrm{spec}(L_p)\} = \{\lambda^{\frac{1}{p^*}} : \lambda \in \mathrm{spec}(L_{p^*})\}, \;\; \lambda_k^{\frac{1}{p}}(L_p) = \lambda_k^{\frac{1}{p^*}}(L_{p^*}), \; k = 1, \cdots, n.$$

Similarly, for path graphs, we can establish a precise relationship between the p-Laplacian on path graphs (used in unsupervised learning) and the p^*-Laplacian on path graphs under homogeneous Dirichlet boundary condition (used in semi-supervised learning). Let P_n be the path graph on n vertices, and let P_{n-1}^D be the path graph with $(n-1)$ interior vertices and 2 boundary vertices (see Figure 1). Then, we can see P_n as a hypergraph, and there is a canonical way to identify P_{n-1}^D with the dual hypergraph of P_n. Hence, for any $n \geq 2$ and $1 < p < +\infty$, we have

$$\{\lambda^{\frac{1}{p}} : \lambda \in \mathrm{spec}(L_p(P_n)) \setminus \{0\}\} = \{\lambda^{\frac{1}{p^*}} : \lambda \in \mathrm{spec}(L_{p^*}(P_{n-1}^D))\},$$

$$\lambda_k^{\frac{1}{p}}(L_p(P_n)) = \lambda_{k-1}^{\frac{1}{p^*}}(L_{p^*}(P_{n-1}^D)), \; k = 2, \cdots, n,$$

$$\mathrm{spec}(L_1(P_n)) \setminus \{0\} = \mathrm{spec}(L_\infty(P_{n-1}^D)), \; \mathrm{spec}(L_\infty(P_n)) \setminus \{0\} = \mathrm{spec}(L_1(P_{n-1}^D))$$

and for $k = 2, \cdots, n$,

$$\lambda_k(L_1(P_n)) = \lambda_{k-1}(L_\infty(P_{n-1}^D)) \text{ and } \lambda_k(L_\infty(P_n)) = \lambda_{k-1}(L_1(P_{n-1}^D)),$$

where $\mathrm{spec}(T)$ denotes the spectrum of a given operator T.

8.1.6 Hypergraphs and core-periphery detection

Consider a hypergraph $H = (V, E, w)$ made by a set of vertices $V = \{1, \dots, n\}$, hyperedges $E = \{e_1, \cdots, e_m\}$ and the weight function $w : E \to \mathbb{R}_+$. Here, unlike the graph case, each $e \in E$ contains an arbitrary number of nodes. The core-periphery detection problem consists of identifying the optimal subdivision of V into a core set highly connected with the rest of H and a periphery set, connected only (or mostly) to the core.

It is shown in [53] that this combinatorial problem on H boils down to the norm-constrained optimization problem,

$$\max \sum_{e \in E} w_e \|x|_e\|_q \quad \text{s.t.} \quad \|x\|_p = 1\,. \tag{8.23}$$

Clearly, if $f(x) = \sum_{e\in E} w_e \|x|_e\|_q$ and $g(x) = \|x\|_p$, the above problem coincides with the largest eigenvalue of the nonlinear eigenvalue problem for the functions pair (f, g). Now, we shall write down the dual eigenvalue problem, i.e., the eigenvalue problem for the function pair $(\mathcal{N}g, \mathcal{N}f)$.

For g we have $\mathcal{N}g(x) = \|x\|_{p^*}$, where $1/p + 1/p^* = 1$. As for f, note that

$$f(x) = \|(\|x|_{e_1}\|_q, \dots, \|x|_{e_m}\|_q)\|_{1,w}\,,$$

where $\|\cdot\|_{1,w}$ indicates the weighted l^1-norm on $\mathbb{R}^E$. Then, by Proposition 3.5, we have

$$\mathcal{N}f(x) = \inf_{\substack{\sum_{e\in E} y_e = x \\ \mathrm{supp}(y_e)\subset e}} \left\| \Big(\frac{\|y_{e_1}\|_{q^*}}{w_{e_1}}, \dots, \frac{\|y_{e_m}\|_{q^*}}{w_{e_m}} \Big) \right\|_\infty = \inf_{\substack{\sum_{e\in E} y_e = x \\ \mathrm{supp}(y_e)\subset e}} \max_{e\in E} \frac{\|y_e\|_{q^*}}{w_e}$$

where y_e denotes a vector in $\mathbb{R}^V$ with the support in e.

Moreover, using Corollary 4.2 we can obtain additional equivalent formulations. For $e \in E$, let $A_e : \mathbb{R}^V \to \mathbb{R}^V$ be a matrix defined as $(A_e x)_i = x_i$ if $i \in e$ and $(A_e x)_i = 0$ if $i \notin e$. Clearly, $\|A_e x\|_q = \|x|_e\|_q$. Thus, we can write f as $f = \tilde{f} \circ A$, i.e., $f(x) = \tilde{f}(Ax)$, where $\tilde{f} : \mathbb{R}^{nm} \to [0, +\infty)$ is the norm defined as

$$\tilde{f}(y_1, \dots, y_m) = \|(\|y_1\|_q, \dots, \|y_m\|_q)\|_{1,w}$$

with $y_j \in \mathbb{R}^n$ and $A : \mathbb{R}^n \to \mathbb{R}^{nm}$ defined as $Ax = (A_{e_1}x, \dots, A_{e_m}x)$. Thus, we immediately see that

$$\mathcal{N}\tilde{f}(y_1, \dots, y_m) = \|(\|y_1\|_{q^*}/w_{e_1}, \dots, \|y_m\|_{q^*}/w_{e_m})\|_\infty$$

for any vector $(y_1, \dots, y_m)$ of dimension $n \times m$. By Corollary 4.2, the largest eigenvalue of the dual eigenvalue problem $(\mathcal{N}g \circ A^\top, \mathcal{N}\tilde{f})$, i.e.,

$$0 \in \partial \Big\| \sum_{i=1}^m A_{e_i}^\top y_i \Big\|_{p^*} - \lambda \partial \left\| \Big(\frac{\|y_1\|_{q^*}}{w_{e_1}}, \dots, \frac{\|y_m\|_{q^*}}{w_{e_m}} \Big) \right\|_\infty$$

coincides with the core-periphery eigenvalue problem (8.23) for (f, g).

8.2 Distance between convex bodies

The Banach-Mazur distance is a key quantity in convex geometry and functional analysis, which has led to noteworthy progress in both those areas, see e.g. [36]. Here, we focus on the Banach-Mazur distance in its multiplicative form between two centrally symmetric convex bodies K and L, centered at the origin point 0 in $\mathbb{R}^n$. This distance is defined as

$$d(K, L) = \inf\{r \geq 1 : \exists A \in GL(\mathbb{R}^n) \text{ s.t. } L \subset AK \subset rL\}, \tag{8.24}$$

where $GL(\mathbb{R}^n)$ is the general linear group. By translating our spectral duality properties into the language of convex geometry we can immediately obtain properties about this distance between convex bodies via the eigenvalue problem for function pairs.

For two convex bodies K and L containing the origin as an interior point, there exists some scaling constant $\lambda > 0$ such that the two convex surfaces ∂K and $\lambda\partial L := \{\lambda y : y \in \partial L\}$ are tangent to each other at some point, where ∂K and ∂L are the boundary surfaces of the bodies K and L, respectively. Here, we say that two convex surfaces are tangent at a if they have a common supporting hyperplane at a. Let

$$\mathcal{ST}(K, L) = \{\lambda > 0 : \partial K \text{ and } \lambda\partial L \text{ are tangent at some point}\}.$$

A first key observation is that $\mathcal{ST}(K, L)$ is a compact subset of $(0, +\infty)$ and it coincides with the set of all the nonzero eigenvalues of the function pair $(\|\cdot\|_K, \|\cdot\|_L)$, where $\|\cdot\|_K$ is the Minkowski functional norm of K, i.e., the norm such that $K = \{x \in \mathbb{R}^n : \|x\|_K \leq 1\}$. Also, it is easy to see that $\mathcal{ST}(L, K) = \{\lambda^{-1} : \lambda \in \mathcal{ST}(K, L)\}$. For such a pair of convex bodies K and L, we can still use (8.24) to define their simple Banach-Mazur distance and use our spectral duality to introduce new distances and observe new identities.

Preciesely, let $\lambda_{\max}(K, L) = \max\{\lambda : \lambda \in \mathcal{ST}(K, L)\}$ and $\lambda_{\min}(K, L) = \min\{\lambda : \lambda \in \mathcal{ST}(K, L)\}$. Corollary 4.2 implies that $\mathcal{ST}(A^\top L^*, K^*) = \mathcal{ST}(AK, L)$ for any $n \times n$ invertible matrix A. Thus, we have the following new representation of the Banach-Mazur distance

$$\begin{aligned} d(K, L) &= \inf_{A \in GL(\mathbb{R}^n)} \frac{\lambda_{\max}(AK, L)}{\lambda_{\min}(AK, L)} = \min_{A \in SL(\mathbb{R}^n)} \frac{\lambda_{\max}(AK, L)}{\lambda_{\min}(AK, L)} \\ &= \min_{A \in SL(\mathbb{R}^n)} \lambda_{\max}(AK, L)\lambda_{\max}(A^\top L^*, K^*) \end{aligned}$$

where $SL(\mathbb{R}^n)$ indicates the special linear group, i.e., the set of matrices with determinants equal to one.

From this formulation, we immediately obtain $d(K, L) = d(K^*, L^*)$, which generalizes the known equality for symmetric convex bodies to the nonsymmetric case. In fact, by Corollary 4.2, $\lambda_{\max}(A^\top L^*, K^*) = \lambda_{\max}(AK, L)$ and $\lambda_{\min}(A^\top L^*, K^*) = \lambda_{\min}(AK, L)$, and thus,

$$d(L^*, K^*) = \inf_{A^\top \in GL(\mathbb{R}^n)} \frac{\lambda_{\max}(A^\top L^*, K^*)}{\lambda_{\min}(A^\top L^*, K^*)} = \inf_{A \in GL(\mathbb{R}^n)} \frac{\lambda_{\max}(AK, L)}{\lambda_{\min}(AK, L)} = d(K, L).$$

Using a similar argument, we can obtain a similar result for other distances. In particular, consider the distance defined by

$$\widehat{d}(K, L) = \inf\left\{r \geq 1 : \frac{1}{r}L \subset K \subset rL\right\}.$$

This distance is used for studying floating and illumination bodies [42], and is equivalent to the Goldman-Iwahori metric introduced for Bruhat-Tits buildings [30]. We have

$$\widehat{d}(K, L) = \max\{\lambda_{\max}(K, L), \frac{1}{\lambda_{\min}(K, L)}\} = \max\{\lambda_{\max}(K, L), \lambda_{\max}(L, K)\}$$

and from this new representation, we easily obtain the duality identity $\widehat{d}(K^*, L^*) = \widehat{d}(K, L)$ via Theorem 4.1 and the discussion above.

References

[1] S. Artstein-Avidan and V. D. Milman. Hidden structures in the class of convex functions and a new duality transform. *J. Eur. Math. Soc.*, 13:975–1004, 2011.

[2] S. Artstein-Avidan and Y. A. Rubinstein. Differential analysis of polarity: Polar Hamilton-Jacobi, conservation laws, and Monge-Ampère equations. *Journal d'Analyse Mathématique*, 132:133–156, 2017.

[3] W. Ballmann. Der satz von lyusternik und schnirelmann. *Bonn. Math. Schr.*, 102, 1978.

[4] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer-Verlag, 2011.

[5] S. Boyd. *Convex Optimization*. Cambridge University Press, 2004.

[6] Z. M. Boyd, E. Bae, X.-C. Tai, and A. L. Bertozzi. Simplified energy landscape for modularity using total variation. *SIAM J. Appl. Math.*, 78(5):2439–2464, 2018.

[7] X. Bresson, X.-C. Tai, T. F. Chan, and A. Szlam. Multi-class transductive learning based on ℓ^1 relaxations of Cheeger cut and Mumford-Shah-Potts model. *J. Math. Imaging Vis.*, 49(1):191–201, 2014.

[8] T. Bühler and M. Hein. Spectral clustering based on the graph p-Laplacian. In *International Conference on Machine Learning*, page 81–88, 2009.

[9] L. Bungert and M. Burger. Gradient flows and nonlinear power methods for the computation of nonlinear eigenfunctions. *Handbook of Numerical Analysis, Numerical Control: Part A*, 23, 2022.

[10] L. Bungert and Y. Korolev. Eigenvalue problems in l^∞: Optimality conditions, duality, and relations with optimal transport. *Commun. Am. Math. Soc.*, 2:345–373, 2022.

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

[12] Y. Cai, L.-H. Zhang, Z. Bai, and R.-C. Li. On an eigenvector-dependent nonlinear eigenvalue problem. *SIAM J. Matrix Anal. Appl.*, 39(3):1360–1382, 2018.

[13] J. Calder. The game theoretic p-Laplacian and semi-supervised learning with few labels. *Nonlinearity*, 32(1):301–330, 2018.

[14] F. R. Chung. *Spectral graph theory*, volume 92. American Mathematical Soc., 1997.

[15] F. H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley New York, 1983.

[16] O. Cornea, G. Lupton, J. Oprea, and D. Tanré. Lusternik-schnirelmann category. *Mathematical Surveys and Monographs*, 103, 2003.

[17] A. Cristofari, F. Rinaldi, and F. Tudisco. Total variation based community detection using a nonlinear optimization approach. *SIAM J. Appl. Math.*, 80(3):1392–1419, 2020.

[18] P. Deidda, M. Burger, M. Putti, and F. Tudisco. The graph ∞-laplacian eigenvalue problem. *SIAM J. Math. Anal.*, 2026.

[19] P. Deidda, M. Putti, and F. Tudisco. Nodal domain count for the generalized graph p-Laplacian. *Appl. Comput. Harmon. Anal.*, 64:1–32, 2023.

[20] P. Deidda, F. Tudisco, and D. Zhang. Nonlinear spectral graph theory. *SIAM Rev.*, to appear.

[21] A. Elmoataz, M. Toutain, and D. Tenbrinck. On the p-Laplacian and ∞-Laplacian on graphs with applications in image and data processing. *SIAM J. Imaging Sci.*, 8:2412–2451, 11 2015.

[22] T. Feld, J.-F. Aujol, G. Gilboa, and N. Papadakis. Rayleigh quotient minimization for absolutely one-homogeneous functionals. *Inverse Problems*, 35, 2019.

[23] D. Fernández-Ternero, E. Macías-Virgós, and J. A. Vilches. Lusternik–Schnirelmann category of simplicial complexes and finite spaces. *Topology Appl.*, 194:37–50, 2015.

[24] M. Flores, J. Calder, and G. Lerman. Analysis and algorithms for ℓ_p-based semi-supervised learning on graphs. *Appl. Comput. Harmon. Anal.*, 60:77–122, 2022.

[25] A. Gautier, M. Hein, and F. Tudisco. The global convergence of the nonlinear power method for mixed-subordinate matrix norms. *J. Sci. Comput.*, 88:21, 2021.

[26] A. Gautier, Q. N. Nguyen, and M. Hein. Globally optimal training of generalized polynomial neural networks with nonlinear spectral methods. *Advances in Neural Information Processing Systems*, 29, 2016.

[27] A. Gautier, F. Tudisco, and M. Hein. A unifying Perron–Frobenius theorem for nonnegative tensors via multihomogeneous maps. *SIAM J. Matrix Anal. Appl.*, 40(3):1206–1231, 2019.

[28] G. Gilboa. *Nonlinear eigenproblems in image processing and computer vision.* Advances in Computer Vision and Pattern Recognition. Springer, Cham, 2018.

[29] M. Gromov. Volume and bounded cohomology. *Inst Hautes Études Sci. Publ. Math.*, 1982(56):5–99, 1983.

[30] T. Haettel. Injective metrics on buildings and symmetric spaces. *Bull. Lond. Math. Soc.*, 54:2297–2313, 2022.

[31] M. Hein and S. Setzer. Beyond spectral clustering - tight relaxations of balanced graph cuts. In *Advances in Neural Information Processing Systems*, volume 24, 2011.

[32] M. Hein, S. Setzer, L. Jost, and S. Rangapuram. The total variation on hypergraphs-learning on hypergraphs revisited. In *Proceedings of the 26th International Conference on Neural Information Processing Systems*, pages 2427–2435, 2013.

[33] H. Hu, T. Laurent, M. A. Porter, and A. L. Bertozzi. A method based on total variation for network modularity optimization using the mbo scheme. *SIAM J. Appl. Math.*, 73(6):2224–2246, 2013.

[34] R. Hynd and E. Lindgren. Approximation of the least Rayleigh quotient for degree p homogeneous functionals. *J. Funct. Anal.*, 272:4873–4918, 2017.

[35] E. Jarlebring, S. Kvaal, and W. Michiels. An inverse iteration method for eigenvalue problems with eigenvector nonlinearities. *SIAM J. Sci. Comput.*, 36(4):A1978–A2001, 2014.

[36] W. B. Johnson and G. Schechtman. The number of closed ideals in l(lp). *Acta Math.*, 2021.

[37] B. Kawohl and M. Novaga. The p-Laplace eigenvalue problem as p→1 and Cheeger sets in a Finsler metric. *J. Convex Anal.*, 15(3):623, 2008.

[38] V. Khrulkov and I. Oseledets. Art of singular vectors and universal adversarial perturbations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8562–8570, 2018.

[39] M. A. Krasnosel'ski. *Topological methods in the theory of nonlinear integral equations.* MacMillan, 1964.

[40] P. Lindqvist. A nonlinear eigenvalue problem. *Topics in mathematical analysis*, 3:175–203, 2008.

[41] L. Lusternik and L. Schnirelmann. Méthodes topologiques dans les problémes variationnels,. *Hermann, Paris*, 1934.

[42] O. Mordhorst and E. M. Werner. Floating and illumination bodies for polytopes: Duality results. *Discrete Anal.*, 11, 2019.

[43] Q. Nguyen, F. Tudisco, A. Gautier, and M. Hein. An efficient multilinear optimization framework for hypergraph matching. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 39(6):1054–1075, 2017.

[44] K. Prokopchik, A. R. Benson, and F. Tudisco. Nonlinear feature diffusion on hypergraphs. In *International Conference on Machine Learning*, 2022.

[45] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

[46] Y. Saad, J. R. Chelikowsky, and S. M. Shontz. Numerical methods for electronic structure calculations of materials. *SIAM Rev.*, 52(1):3–54, 2010.

[47] M. Sion. On general minimax theorems. *Pacific J. Math.*, 8:171–176, 1958.

[48] D. Slepčev and M. Thorpe. Analysis of p-Laplacian regularization in semisupervised learning. *SIAM J. Math. Anal.*, 51(3):2085–2120, 2019.

[49] M. Struwe. *Variational methods*, volume 34 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics].* Springer-Verlag, Berlin, fourth edition, 2008. Applications to nonlinear partial differential equations and Hamiltonian systems.

[50] F. Tudisco and M. Hein. A nodal domain theorem and a higher-order Cheeger inequality for the graph p-Laplacian. *J. Spectr. Theory*, 8:883–908, 2018.

[51] F. Tudisco and D. J. Higham. A nonlinear spectral method for core-periphery detection in networks. *SIAM J. Math. Data Sci.*, 1:269–292, 2019.

[52] F. Tudisco and D. J. Higham. Node and edge eigenvector centrality for hypergraphs. *Communications Physics*, 4(201), 2021.

[53] F. Tudisco and D. J. Higham. Core-periphery detection in hypergraphs. *SIAM J. Math. Data Sci.*, 5(1):1–21, 2023.

[54] F. Tudisco, P. Mercado, and M. Hein. Community detection in networks via nonlinear modularity eigenvectors. *SIAM J. Appl. Math.*, 78:2393–2419, 2018.

[55] P. Upadhyaya, E. Jarlebring, and E. H. Rubensson. A density matrix approach to the convergence of the self-consistent field iteration. *Numer. Algebra Control Optim.*, 11(1):99–115, 2021.

[56] K. Yosida. *Functional Analysis.* Springer Berlin Heidelberg, 1974.

[57] E. Zeidler. *Nonlinear functional analysis and its applications III: Variational methods and optimization.* Springer, 2nd edition, 2013.

[58] G. Zhou, L. Caccetta, K. L. Teo, and S.-Y. Wu. Nonnegative polynomial optimization over unit spheres and convex programming relaxations. *SIAM J. Optim.*, 22(3):987–1008, 2012.

- Francesco Tudisco
  School of Mathematics and Maxwell Institute, The University of Edinburgh, EH93FD Edinburgh (UK)
  f.tudisco@ed.ac.uk

- Dong Zhang
  School of Mathematical Sciences, Peking University, 100871 Beijing (China)
  dongzhang@math.pku.edu.cn