Non-linear Generalizations of eigenvalues of Graph Laplacians and Algebraic Connectivity

E. M. Borba, S. Richter, U. Schwerdtfeger

Preprint 2017-6

Preprintreihe der Fakultät für Mathematik
ISSN 1614-8835

Non-linear Generalizations of eigenvalues of Graph Laplacians and Algebraic Connectivity

Elizandro Max Borba[1], Sebastian Richter[2], and Uwe Schwerdtfeger[2]

[1]Instituto de Matemática, Universidade Federal do Rio Grande do Sul,
CEP 91509-900, Porto Alegre, RS, Brazil,
elizandro.max@ufrgs.br
[2]Fakultät für Mathematik, Technische Universität Chemnitz,
D-09107 Chemnitz, Germany,
uwe.schwerdtfeger@mathematik.tu-chemnitz.de

December 20, 2017

Abstract

We generalize the notion of algebraic connectivity, the second smallest eigenvalue of the graph Laplacian matrix, by varying the norms in the Rayleigh-Ritz characterization. The so obtained parameters are shown to be well-defined and to be the least non-zero eigenvalues of the corresponding non-linear eigenvalue problem whenever the graph is connected. We provide combinatorial interpretations of several non-smooth cases.
Keywords: (signless) p-Laplacian, algebraic connectivity, isoperimetric number, Rayleigh quotient, graph partition, nonlinear eigenvalue problem
MSC 2010: 05C50, 05C40, 15A18, 47J10

1 Introduction

We begin with some notation. All graphs are simple and undirected without loops or multiple edges. For a graph $G = (V, E)$ with vertex set V the edge set E hence consists of two-element subsets of V and we shall frequently write ij rather than $\{i, j\}$ for an edge. For disjoint subsets $S, T \subseteq V$ define $E(S)$ as the edges $ij \in E$ spanned by S, $E(S, T)$ the edges with one vertex in S and the other in T, furthermore we define $n = |V|, m = |E|$, $e(S) = |E_G(S)|$, $e(S, T) = |E(S, T)|$ and $\mathrm{cut}_G(S) = |E_G(S, V \setminus S)|$. Furthermore $d_i, i \in V$ is the degree of vertex i and $\delta(G), \Delta(G)$ denote the minimum and maximum degree, respectively. We shall drop the subscripts if G is clear from the context. The volume of a set $X \subseteq V$ is $\mathrm{vol}(X) = \sum_{i\in X} d_i$. We find it convenient to index vectors and matrices associated with the graph by the vertex set V (resp. edge set E) and write therefore $x = (x_i, \in V) \in \mathbb{R}^V$ (resp. $x = (x_{ij}, \in E) \in \mathbb{R}^E$) rather than $x \in\ R^n$ (resp. $x \in \mathbb{R}^m$). Accordingly, we denote by $A \in \mathbb{R}^{V\times V}$ the adjacency matrix and by $D \in \mathbb{R}^{V\times V}$ the diagonal matrix of vertex degrees, incidence and gradient matrices are considered as elements of $\mathbb{R}^{E\times V}$. The incidence vector of a vertex set X is denoted by $\mathbb{1}_X \in \mathbb{R}^V$ and $\mathbb{1} \in \mathbb{R}^V$ denotes the all ones vector. Given an orientation $i \to j$ on every edge

$e = ij$ we define a gradient matrix $R \in \mathbb{R}^{E\times V}$ by

$$R_{ei} = \begin{cases} 1, \text{ if } i \in e \text{ and } e = i \to j \\ -1, \text{ if } i \in e \text{ and } e = j \to i \\ 0, \text{ else.} \end{cases}$$

The Laplacian matrix $R^\top R = L = D - A$ of the graph, its eigenvalues and -vectors and their applications have been widely studied and we refer the surveys [Moh91, Moh04] or the books [BLS07, CRS09] and references therein. The smallest eigenvalue is 0, afforded by $\mathbb{1}$, the second smallest eigenvalue λ_2 is non-zero if and only if G is connected and is therefore called the algebraic connectivity of G. By Rayleigh's principle we have

$$\lambda_2 = \min_{\substack{x\neq 0\\ \mathbb{1}^\top x=0}} \frac{x^\top L x}{x^\top x} = \min_{\substack{x\neq 0\\ \mathbb{1}^\top x=0}} \left(\frac{\|Rx\|_2}{\|x\|_2}\right)^2 \tag{1.1}$$

In this article we investigate a family of parameters obtained by replacing the Rayleigh quotient in this characterization by a quotient $\|Rx\|_p / \|x\|_q$ with possibly different p, q. In [Amg03, Amg06] eigenvalues of the discrete p-Laplacian (see Definition 2.11) are studied. The second smallest eigenvalue $\lambda_2^{(p)}$ satisfies a variational characterization as above, with the 2-norms replaced by p-norms ($p > 1$) and a suitable constraint. This idea is applied in a spectral clustering technique based on the p-Laplacian in [BH09] where the authors give the following bounds

$$\left(\frac{2}{\Delta}\right)^{p-1}\left(\frac{i(G)}{p}\right)^p \leq \lambda_2^{(p)} \leq 2^{p-1} i(G), \tag{1.2}$$

in terms of the isoperimetric number

$$i(G) = \min\left\{\frac{\mathrm{cut}(S)}{|S|},\ S \subseteq V,\ 0 < |S| \leq \frac{n}{2}\right\}$$

and use the corresponding eigenvectors to approximate $i(G)$. For $p = 2$ equation (1.2) yields a slightly weaker lower bound than Mohar's [Moh89]. For $p = 1$ the equality $\lambda_2^{(p)} = i(G)$ indeed holds, see Proposition 3.1.

The famous Motzkin-Straus Theorem serves as a motivation to also consider Rayleigh quotients $\|Rx\|_p / \|x\|_q$ with $p \neq q$. With $\omega(G)$ denoting the clique number it can be stated as

$$\max_{x\neq 0} \frac{x^\top A x}{\|x\|_1^2} = 1 - \frac{1}{\omega(G)}.$$

So the $\|\cdot\|_1$-adjacency spectral radius yields the clique number. Extremal problems on the $\|\cdot\|_p$- adjacency-spectral radius are discussed in a recent paper [KN14].

In order to formulate a proper generalization of equation (1.1) we remark that the constraint $\mathbb{1}^\top x = 0$ is equivalent to $\|x - m\mathbb{1}\|_2$ being minimal at $m = 0$. More generally, for any norm $\|\cdot\|$ on $\mathbb{R}^n$ we have

$$\|x - m\mathbb{1}\| \text{ is minimal for } m = 0 \Leftrightarrow \mathbb{1}^\top \partial\|x\| \ni 0 \tag{1.3}$$

where $\partial\|x\|$ denotes the subdifferential of $\|\cdot\|$ at the point x. For any two arbitrary norms $\|\cdot\|_E$ on $\mathbb{R}^E$ and $\|\cdot\|_V$ on $\mathbb{R}^V$ we define the Rayleigh quotient $\mathcal{R}_{\|\cdot\|_E,\|\cdot\|_V}(x)$ and the parameter

$a_{\|\cdot\|_E,\|\cdot\|_V}(G)$ by

$$\mathcal{R}(x) = \mathcal{R}_{\|\cdot\|_E,\|\cdot\|_V}(x) := \frac{\|Rx\|_E}{\|x\|_V} \text{ and } a(G) = a_{\|\cdot\|_E,\|\cdot\|_V}(G) = \min_{\substack{x\neq 0\\ 0\in \mathbb{1}^\top \partial\|x\|_V}} \mathcal{R}_{\|\cdot\|_E,\|\cdot\|_V}(x).$$

The eigenvalue problem for $\mathcal{R}$ asks for $(\lambda, x) \in \mathbb{R} \times (\mathbb{R}^V \setminus \{0\}$ with

$$\partial \left\|\cdot\right\|_E (Rx) \cap \lambda\partial \left\|\cdot\right\|_V (x) \neq \emptyset,$$

in which case $\lambda = \mathcal{R}(x)$ (see the next section). If both norms are Euclidean this is the usual eigenvalue problem. Clearly, the vector $x = \mathbb{1}$ affords the eigenvalue 0 which is the smallest critical value of $\mathcal{R}$. We will see below that any eigenvector x with $\lambda > 0$ satisfies (1.3) and that $a(G)$ is well-defined for any pair of norms and indeed the least non-zero eigenvalue of $\mathcal{R}$ whenever G is connected.

The paper is organized as follows. In the next section we define the eigenvalue problem for an arbitrary pair $\|\cdot\|_E$ and $\|\cdot\|_V$ and show that a_G is the least non-zero eigenvalue whenever G is connected. In the subsequent sections we consider $a(G)$ for various pairs We shall specialize to weighted p-norms $\|x\| = (\sum m_i x_i^p)^{1/p}$, $m_i > 0$, $1 \leq p \leq \infty$ which also cover the case of weighted and normalized Laplacians. and write $\mathcal{R}_{p,q} = \mathcal{R}_{\|\cdot\|_p,\|\cdot\|_q}$ and $a_{p,q}(G)a_{\|\cdot\|_p,\|\cdot\|_q}(G)$ for short. Notice that with our notation $\lambda_2^{(p)}$ as considered in [Amg03, BH09] reads $a_{p,p}(G)^p$

2 Critical points, eigenvalues and eigenvectors

The subdifferential $\partial f(x)$ of a locally Lipschitz function $f\colon \mathbb{R}^n \to \mathbb{R}$ is defined as in [Cla90, Chapter 2]. We shall sometimes write ∂_x to emphasize that the subdifferential is taken with respect to x. We recall the most important facts for our purposes. If f is differentiable then $\partial f(x) = \{\nabla f(x)\}$ (gradient) and if f is convex then

$$\partial f(x) = \left\{ s \in \mathbb{R}^n \colon \forall y \in \mathbb{R}^n \colon f(y) \geq f(x) + s^\top (y - x) \right\}. \tag{2.1}$$

Furthermore, for an affine function $g(x) = Ax + b$ with $A \in \mathbb{R}^{n\times m}$ and $b \in \mathbb{R}^n$ and convex f we have the chain rule

$$\partial(f \circ g)(x) = A^\top \partial f(Ax + b) \tag{2.2}$$

and for $f_1, f_2 \colon \mathbb{R}^n \to \mathbb{R}$ locally Lipschitz with $f_2(x) > 0$ the quotient rule

$$\partial\left(\frac{f_1}{f_2}\right)(x) \subseteq \frac{f_2(x)\partial f_1(x) - f_1(x)\partial f_2(x)}{f_2(x)^2}. \tag{2.3}$$

The point x is called a critical point of f if $0 \in \partial f(x)$. If f has a local extremum at x then x is necessarily critical. If f is convex then this is also sufficient. The following Lagrange multiplier rule [Cla90, Theorem 6.1.1] holds in this context.

Proposition 2.1. *Let f, h_i, $i = 1, \ldots, k$ be locally Lipschitz. If x is an optimal solution to*

$$\min f(x) \text{ s.t. } h_i(x) = 0\, \forall i = 1, \ldots, k$$

then there are multipliers $(\phi, \lambda) \in \mathbb{R} \times \mathbb{R}^k \setminus \{(0, 0)\}$ such that

$$0 \in \phi\partial f(x) + \sum_{i=1}^{k} \lambda_i \partial h_i(x).$$

We are particularly interested in the cases when f is a norm in which case we use the short hand notation

$$\partial \|x\| := \partial \|\cdot\| (x).$$

For a norm $\|\cdot\|$ its dual norm is defined as

$$\|z\|^* = \max_{\|x\|=1} z^\top x$$

Furthermore we define the function

$$d(x) := d_{\|\cdot\|}(x) = \min_{m\in\mathbb{R}} \|x - m\mathbb{1}\| . \tag{2.4}$$

The function $x \mapsto d(x)$ is the distance of x to the (convex and closed) subspace spanned by $\mathbb{1}$ and therefore a convex function.

We collect some facts for later reference in the following lemma.

Lemma 2.2. *Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ and $x \neq 0$ and $f\colon \mathbb{R}^n \to \mathbb{R}$ convex.*

1. *If f is positively homogeneous of degree p, i.e. $f(\lambda x) = \lambda^p f(x)$ for $\lambda \geq 0$ and $s \in \partial f(x)$ then $pf(x) = s^\top x$ and $\partial f(\lambda x) = \lambda^{p-1}\partial f(x)$. In particular, if $s \in \partial\|\cdot\|(x)$ then $s^\top x = \|x\|$ and $\|s\|^* = 1$.*
2. *$\|\cdot\|^{**} = \|\cdot\|$ and if $\|s\|^* = 1 = \|x\|$ then it holds that $s \in \partial \|x\| \Leftrightarrow x \in \partial \|s\|^*$.*
3. *$\|x - m\mathbb{1}\|$ is minimal for $m = 0$ if and only if $0 \in \mathbb{1}^\top \partial \|x\|$.*
4. *For $\alpha > 0$ and $\beta \in \mathbb{R}$ we have $d(\alpha x + \beta\mathbb{1}) = \alpha d(x)$.*
5. *Let $d(x) = \|x - m_x\mathbb{1}\|$. Then $\partial d(x) = \partial \|x - m_x\mathbb{1}\| \cap \mathbb{1}^\perp$.*

Proof. For 1, compute $\partial_\lambda f(\lambda x)$ in two ways. The chain rule yields $\partial_\lambda f(\lambda x) = x^\top \partial_x f(\lambda x)$. On the other hand, use $f(\lambda x) = \lambda^p f(x)$ to get $\partial_\lambda f(\lambda x) = p\lambda^{p-1} f(x)$ and $\lambda = 1$ yields the first equality. The equality $\partial f(\lambda x) = \lambda^{p-1}\partial f(x)$ follows easily from (2.1). Finally, we have $\|x\| = s^\top x \leq \|x\| \|s\|^*$, thus $\|s\|^* \geq 1$. For the converse inequality observe that $\forall y \in \mathbb{R}^n \setminus \{0\}\colon\ \|y\| \geq \|x\| + s^\top(y - x) = s^\top y$ and hence $\|s\|^* \leq 1$.
For 2, let $s \in \mathbb{R}^n$ with $\|s\|^* = 1$ and $\|x\|^{**} = s^\top x$ and $t \in \partial \|x\|$. Then $\|x\|^{**} = s^\top x \leq \|s\|^* \|x\| = \|x\|$ and, by 1, $\|x\| = t^\top x \leq \|t\|^* \|x\|^{**} = \|x\|^{**}$. Now assume that $\|s\|^* = 1 = \|x\|$ and $s \in \partial \|x\|$. Then for any t we have $\|s\|^* + x^\top(t - s) = \|s\|^* - \|x\| + x^\top t \leq \|x\| \|t\|^* = \|t\|^*$ and hence $x \in \partial \|s\|^*$. The converse inclusion follows analogously.
Assertion 3 follows from the chain rule and convexity of $m \mapsto \|x - m\mathbb{1}\|$.
For assertion 4 let m_x denote any $m \in \mathbb{R}$ for which $d(x) = \|x - m_x\mathbb{1}\|$. Clearly $d(x) = d(x + \beta\mathbb{1})$. Furthermore,

$$\begin{aligned} d(\alpha x) = \|\alpha x - m_{\alpha x}\mathbb{1}\| &= \alpha \left\|x - \alpha^{-1} m_{\alpha x}\mathbb{1}\right\| \geq \alpha d(x) \\ &= \alpha \|x - m_x\mathbb{1}\| = \|\alpha x - \alpha m_x \mathbb{1}\| \geq d(\alpha x). \end{aligned}$$

For assertion 5 let $s \in \partial d(x)$. Then $d(x) = d(x \pm \mathbb{1}) \geq d(x) \pm s^\top \mathbb{1}$ and so $s^\top \mathbb{1} = 0$. Furthermore, for any $y \in \mathbb{R}^n : \|y\| \geq d(y) \geq d(x) + s^\top(y - x) = \|x - m_x\mathbb{1}\| + s^\top(y - (x - m_x\mathbb{1}))$ and hence $s \in \partial \|x - m_x\mathbb{1}\| \cap \mathbb{1}^\perp$. For the converse inclusion first observe that there is $s \in \partial \|x - m_x\mathbb{1}\| \cap \mathbb{1}^\perp$ by 3. So for any such s and any $y \in \mathbb{R}^n$ we have $d(y) = \|y - m_y\mathbb{1}\| \geq \|x - m_x\mathbb{1}\| + s^\top(y - x + (m_x - m_y)\mathbb{1}) = d(x) + s^\top(y - x)$. ∎

2.1 $a(G)$ is well-defined and an eigenvalue

Since for $c > 0$ we have $\mathcal{R}(cx) = \mathcal{R}(x)$ we can restrict the feasible region in the definition of $a(G)$ to the $\|\cdot\|_V$-unit sphere and therefore define

$$\mathcal{S} = \{x \in \mathbb{R}^V : \|x\|_V = 1\} \text{ and } \mathcal{X} = \{x \in \mathcal{S} : 0 \in \mathbb{1}^\top \partial \|x\|_V\}.$$

Lemma 2.3. *$\mathcal{X}$ is compact. In particular, $a(G) = \min_{x\in\mathcal{X}} \mathcal{R}(x)$ is well defined.*

Proof. We only have to prove that $\mathcal{X}$ is closed. Let $x_n \to x$ be a convergent sequence in $\mathcal{X}$ and $s_n \in \partial \|x_n\|_V$ with $\mathbb{1}^\top s_n = 0$. Since $\|s_n\|_V^* = 1$ and $\mathbb{1}^\perp$ is closed there is a subsequence $s_{n_k} \to s \in \mathbb{1}^\perp$. Then for any $y \in \mathbb{R}^V$ we have $\|y\|_V \geq \|x_{n_k}\|_V + s_{n_k}^\top (y - x_{n_k})$ and by continuity $\|y\|_V \geq \|x\|_V + s^\top (y - x)$, hence $s \in \partial \|x\|_V$ and $x \in \mathcal{X}$. $\square$

As in the linear case we have

Lemma 2.4. *$a(G) = 0 \Leftrightarrow G$ is disconnected.*

Proof. Recall that $\ker R = \operatorname{span}\{\mathbb{1}_C : C \text{ component of } G\}$ and observe that $0 \notin \mathbb{1}^\top \partial \|\mathbb{1}\|$. So, if $a(G) = 0$ then there is a non-zero vector in $\ker R$ which is not a scalar multiple of $\mathbb{1}$ and hence G is not connected. Conversely, for disconnected G choose any nonzero $\tilde{x} \in \ker R \setminus \operatorname{span}\{\mathbb{1}\}$ and take $x = \tilde{x} + m\mathbb{1}$ with a suitable m which minimizes $\|\tilde{x} + m\mathbb{1}\|_V$. $\square$

Next we define what an eigenvalue in this general setting is and show that $a(G)$ is indeed the least non-zero eigenvalue whenever G is connected.

Definition 2.5. *We call $(\lambda, x) \in \mathbb{R} \times \mathbb{R}^V$ an eigenpair for $\mathcal{R}$ consisting of an eigenvalue λ and an eigenvector x if there is $s \in \partial \|Rx\|_E$ and $t \in \partial \|x\|_V$ such that the eigenequation $R^\top s = \lambda t$ holds.*

As in the linear case we have

Lemma 2.6.
1. *If $x \in \mathbb{R}^V \setminus \{0\}$ is a critical point of $\mathcal{R}$, i.e. $0 \in \partial\mathcal{R}(x)$, then x is an eigenvector with corresponding eigenvalue $\lambda = \mathcal{R}(x)$.*
2. *If (λ, x) is an eigenpair then $\lambda = \mathcal{R}(x)$.*
3. *If (λ, x) is an eigenpair with $\lambda \neq 0$ then for the corresponding $t \in \partial \|x\|_V$ it holds that $\mathbb{1}^\top t = 0$, in particular $0 \in \mathbb{1}^\top \partial \|x\|_V$.*

Proof. The first statement follows from the chain and quotient rules (2.2) and (2.3) applied to $\mathcal{R}$

$$\partial\mathcal{R}(x) \subseteq \frac{\|x\|_V R^\top \partial \|Rx\|_E - \|Rx\|_E \partial \|x\|_V}{\|x\|_V^2}.$$

For the second statement multiply the eigenequation from the left with $x^\top$ and apply Lemma 2.2. For the third statement multiply the eigenequation by $\mathbb{1}^\top$. $\square$

Remark: The converse of 1 is not true in general. A counterexample is given in [CSZ15].

Lemma 2.7. *Assuming that G is connected we have*

$$a(G) = \min_{d(x)=1} \|Rx\|_E = \left(\max_{\|Rx\|_E=1} d(x) \right)^{-1}$$

and for $x \in \mathcal{X}$ with $\mathcal{R}(x) = a(G)$ there are $t \in \partial \|x\|_V$ and $s \in \partial \|Rx\|_E$ with $R^\top s = a(G)t$ and $a(G)$ is the least non-zero eigenvalue of $\mathcal{R}$.

Proof. First observe that $d(x) = \|x\|_V \Leftrightarrow 0 \in \mathbb{1}^\top \partial \|x\|_V$ and $d(x) = 0 \Leftrightarrow x = m\mathbb{1}$ for some m. By Lemma 2.2 item 4 we have

$$a(G) = \min_{\substack{x \neq 0 \\ 0 \in \mathbb{1}^\top \partial \|x\|_V}} \mathcal{R}(x) = \min_{\substack{\|x\|_V > 0 \\ \|x\|_V = d(x)}} \frac{\|Rx\|_E}{\|x\|_V} = \min_{d(x)>0} \frac{\|Rx\|_E}{d(x)} = \min_{d(x)=1} \|Rx\|_E \,.$$

If y with $d(y) = \|y - m_y\mathbb{1}\| = 1$ is a minimizer then so is $x = y - m_y\mathbb{1} \in \mathcal{X}$. By Proposition 2.1 a necessary condition for a minimum at x is the existence of a Lagrange multiplier λ and subgradients $s \in \partial \|Rx\|_E$, $t \in \partial d(x)$ such that $R^\top s - \lambda t = 0$. By Lemma 2.2 item 5 we have that $t \in \partial \|x\| \cap \mathbb{1}^\perp$ and multiplication by $x^\top$ yields $\lambda = a(G)$. Therefore we have established that $(a(G), x)$ is indeed an eigenpair for $\mathcal{R}$. If (μ, z) is another eigenpair for $\mathcal{R}$ with $\mu > 0$ and $\|z\|_V = 1$ then by Lemma 2.6 we have $z \in \mathcal{X}$ and thus $\mu = \mathcal{R}(z) \geq a(G)$.

The assertion $a(G) = \left(\max_{\|Rx\|_E=1} d(x)\right)^{-1}$ follows from Lemma 2.2 item 4 because G is assumed connected and therefore $\|Rx\|_E = 0 \Leftrightarrow x = m\mathbb{1} \Leftrightarrow d(x) = 0$. □

2.2 $a(G)$ is a critical value of $\mathcal{R}$

We have already remarked that an eigenvalue of $\mathcal{R}$ is not necessarily critical. The eigenvalue $a(G)$, however, is always critical. To show this we follow [HT, Cha16] and define a sequence of critical values of $\mathcal{R}$ in the vein of the Rayleigh-Ritz characterisation of the k-th largest eigenvalue of a symmetric matrix. The notion of dimension is generalized in the following definition.

Definition 2.8. *Let $A \subseteq \mathcal{S}$ be a closed and symmetric ($A = -A$) subset of $\mathcal{S}$. The Krasnoselskii genus (genus, for short) $\gamma(A)$ of A is defined as*

$$\gamma(A) = \begin{cases} 0 \text{ if } A = \emptyset, \\ \inf\{m \colon \exists h \colon A \to \mathbb{R}^m \setminus \{0\},\ h \text{ continuous and } h(-u) = -h(u)\}, \\ \infty \text{ if } \{\cdots\} = \emptyset. \end{cases}$$

Furthermore, let $\mathcal{F}_k = \{A \colon A \subseteq \mathcal{S},\ A = -A,\ \gamma(A) \geq k\}$.

The restriction $\hat{\mathcal{R}} = \mathcal{R}|_{\mathcal{S}}$ satisfies the Palais-Smale condition [Cha81, Definition 2] and hence the following sequence is well-defined

$$\lambda_k = \inf_{A \in \mathcal{F}_k} \max_{x \in A} \hat{\mathcal{R}}(x),\ k = 1, \ldots, n.$$

We remark that inf can actually be replaced by min. In [HT, Lemma 2.3] it is shown that the minimizing set contains a critical point x of $\mathcal{R}$ ($0 \in \partial\mathcal{R}(x)$) with $\mathcal{R}(x) = \lambda_k$. The proof uses the deformation theorem [Cha81, Theorem 3.1, Remark 3.3] and and adapts to our framework almost verbatim. We omit it and only show the following proposition.

Proposition 2.9. *For any graph G we have $\lambda_2 = a(G)$.*

Proof. Let $x \in \mathcal{X}$ be such that $\mathcal{R}(x) = a(G)$ and let $B = \mathcal{S} \cap \operatorname{span}(\mathbb{1}, x)$. Then for every $y = \alpha\mathbb{1} + \beta x$ we have

$$\mathcal{R}(y) = \frac{\|\alpha R\mathbb{1} + \beta Rx\|_E}{\|\alpha\mathbb{1} + \beta x\|_V} = \frac{\|Rx\|_E}{\left\|\frac{\alpha}{\beta}\mathbb{1} + x\right\|_V} \leq \frac{\|Rx\|_E}{\|x\|_V} = a(G).$$

Because $\gamma(B) = 2$ we have

$$a(G) = \max_{x\in B} \mathcal{R}(x) \geq \min_{A\in\mathcal{F}_2} \max_{x\in A} \mathcal{R}(x) = \lambda_2.$$

For the converse inequality, if $\lambda_2 > 0$ then any normalized critical y with $\mathcal{R}(y) = \lambda_2$ is contained in $\mathcal{X}$ by Lemma 2.6 and therefore $\lambda_2 = \mathcal{R}(y) \geq a(G)$. If $\lambda_2 = 0$ and A with $\gamma(A) \geq 2$ is a minimizing set then $\mathcal{R}(y) = 0$ for all $y \in A$ and thus $A \subseteq \mathcal{S} \cap \ker R$. This implies $\gamma(\mathcal{S} \cap \ker R) = \dim\ker R \geq \gamma(A) \geq 2$ and thus G is disconnected and hence $a(G) = 0$. $\square$

2.3 Subdifferentials of some norms

To conclude this section we list the subdifferentials and the sets $\mathcal{X}$ for some norms. We need the set-valued function

$$\operatorname{sign}(x) = \begin{cases} \{1\} \text{ if } x > 0, \\ \{-1\} \text{ if } x < 0, \\ [-1,1] \text{ if } x = 0. \end{cases}$$

which describes the subdifferential of $|x|$ at x. Furthermore, for $1 < p < \infty$ and an index set J (usually V or E) we define the function Φ_p

$$\mathbb{R}^J \ni x = (x_i,\ i \in J) \mapsto \Phi_p(x) = \big(\operatorname{sign}(x_i)\,|x_i|^{p-1},\ i \in J\big)$$

and for $p = 1$ the set-valued function

$$\mathbb{R}^J \ni x \mapsto \Phi_1(x) = \{s \in \mathbb{R}^J\colon s_i \in \operatorname{sign}(x_i),\ i \in J\} =: (\operatorname{sign}(x_i),\ i \in J).$$

We shall frequently abuse notation by identifying a singleton set with its element. For a vector $x \in \mathbb{R}^J$ let P_x, N_x and Z_x denote the subsets of J on which x is positive, negative and zero respectively, and let p_x, n_x and z_x be there respective cardinalities.

Lemma 2.10. *1. For $1 \leq p \leq \infty$ we have $\|\cdot\|_p^* = \|\cdot\|_q$, where $\frac{1}{p} + \frac{1}{q} = 1$, in particular $\|\cdot\|_1^* = \|\cdot\|_\infty$*

2. If $1 < p < \infty$ then $\partial\|x\|_p = \nabla\|x\|_p = \|x\|_p^{1-p}\,\Phi_p(x)$ and

$$\mathcal{X}_p = \left\{x\colon\ \|x\|_p = 1,\ \sum_{i\in V} \operatorname{sign}(x_i)\,|x_i|^{p-1} = 0\right\}$$

3. $\partial\|x\|_1 = \Phi_1(x) = \{t \in \mathbb{R}^n\colon t_i \in \operatorname{sign}(x_i)\}$ and $\mathcal{X}_1 = \{x\colon\ \|x\|_1 = 1,\ |p_x - n_x| \leq z_x\}$

4. If $\|x\|_V = \|Dx\|_1 = \sum_{i\in V} d_i|x_i|$ then $\mathcal{X} = \{x\colon\ \|Dx\|_1 = 1,\ |\operatorname{vol}(P_x) - \operatorname{vol}(N_x)| \leq \operatorname{vol}(Z_x)\}$.

5. $\partial \|x\|_\infty = \operatorname{conv}(\operatorname{sign}(x_i)\, e_i \colon i \in V \text{ and } |x_i| = \|x\|_\infty)$ *and* $\mathcal{X}_\infty = \{x \in \mathbb{R}^V \colon \max_{i\in V} x_i = 1 = -\min_{i\in V} x_i\}$.

Proof. We only show the claims about $\mathcal{X}$. For 3 observe that $x \in \mathcal{X}_1$ if and only if

$$0 \in \mathbb{1}^\top \partial \|x\|_1 = p_x - n_x + \sum_{i\colon x_i = 0} [-1, 1] = p_x - n_x + [-z_x, z_x].$$

The proof of 4 is analogous. For 5 let $\|x\|_\infty = 1$ with $x_i = 1$ and $x_j = -1$ then $s = (e_i - e_j)/2 \in \partial \|x\|_\infty$ is a subgradient with $\mathbb{1}^\top s = 0$. Conversely, if $\|x\|_\infty = 1$ with $x_i > -1$ for all $i \in V$ then any subgradient has non-negative entries only and $\mathbb{1}^\top s > 0$. $\square$

When $\|\cdot\|_E$ is a (weighted) p-norm on $\mathbb{R}^E$ the eigenequation can be reformulated in terms of the p-Laplace operator.

Definition 2.11. *Let $1 < p < \infty$ and $w \in \mathbb{R}^E$ be a vector of non-negative edge weights and $W = Diag(w_{ij}^{1/p})$. Define the (weighted) p-Laplace operator $L_p \colon \mathbb{R}^V \to \mathbb{R}^V$ by $x \mapsto L_p(x) = R^\top W \Phi_p(WRx)$,* i.e. *for $i \in V$*

$$L_p(x)_i = \sum_{j\colon ij \in E} w_{ij} \operatorname{sign}(x_i - x_j)\, |x_i - x_j|^{p-1}.$$

Equivalently, $L_p(x) = \nabla F_p(x)$ for $F_p(x) = p^{-1} \sum_{ij\in E} w_{ij} |x_i - x_j|^p = p^{-1} \|Rx\|_p^p$.

So if $\|\cdot\|_E = \|\cdot\|_p$ then (λ, x) is an eigenpair for $\mathcal{R}$ if and only if there is a subgradient $t \in \partial \|x\|_V$ with

$$L_p(x) = \lambda^p \|x\|_V^{p-1}\, t \tag{2.5}$$

3 The case $\|\cdot\|_E = \|\cdot\|_1$

The eigenvalues of $\mathcal{R}_{1,1}$ and $\mathcal{R}_{1,\infty}$ all have a rather special form as the next proposition shows. The results on $\mathcal{R}_{1,1}$ with the degree weighted 1-norm have been stated in [CSZ15].

Proposition 3.1. 1. *If $s \in \partial \|Rx\|_1$ then also $s \in \partial \left\|Rx^{(i)}\right\|_1$ for $i \in \{1, \infty\}$ where*

$$x^{(1)} = \mathbb{1}_{P_x} \text{ and } x^{(\infty)} = \mathbb{1}_{P_x \cup Z_x} - \mathbb{1}_{N_x}.$$

2. *If (λ, x) is an eigenpair for $\mathcal{R}_{1,1}$ with $\lambda \neq 0$ and (without loss of generality) $P_x \neq \emptyset$ then $(\lambda, x^{(1)})$ is also an eigenpair for $\mathcal{R}_{1,1}$ and hence $\lambda = \operatorname{cut}(P_x)/|P_x|$. In particular, $a_{1,1}(G) = i(G)$.*

3. *If (λ, x) is an eigenpair for $\mathcal{R}_{1,\infty}$ with $\lambda \neq 0$ then $(\lambda, x^{(\infty)})$ is also an eigenpair for $\mathcal{R}_{1,\infty}$ and hence $\lambda = 2\operatorname{cut}(P_x)$. In particular, $a_{1,\infty}(G) = 2\operatorname{mincut}(G)$.*

Proof. 1. When passing from x to $y \in \{x^{(1)}, x^{(\infty)}\}$ we have for every $ij \in E$ that $s_{ij} \in \operatorname{sign}(x_i - x_j) \subseteq \operatorname{sign}(y_i - y_j)$ and so $s \in \partial \|Ry\|_1$.

For the second and third assertions we consider the eigenequation $R^\top s = \lambda t$.
For assertion 2 we have $t \in \partial \|x\|_1 \subseteq \partial \|\mathbb{1}_{P_x}\|_1$. By 1 we have that $(\lambda, \mathbb{1}_{P_x})$ is an eigenpair and hence $\lambda = \mathcal{R}_{1,1}(\mathbb{1}_{P_x}) = \frac{\operatorname{cut}(P_x)}{|P_x|}$. Since $x \in \mathcal{X}_1$ we have $p_x = |P_x| \leq |V|/2$ and hence $\lambda \geq i(G)$.

Conversely, $\frac{1}{|C|}\mathbb{1}_C \in \mathcal{X}_1$ for every $C \subset V$, $1 \le |C| \le |V|/2$, in particular a set C which yields $i(G)$. So $a_{1,1}(G) \le i(G)$ follows.

For assertion 3 we can assume that x is normalized, hence $x \in \mathcal{X}_\infty$ and $P_x \neq \emptyset \neq N_x$. Furthermore we have $t \in \partial \|x\|_\infty \subseteq \partial \left\|x^{(\infty)}\right\|_\infty$ and thus $(\lambda, x^{(\infty)})$ is another eigenpair for $\mathcal{R}_{1,\infty}$. The rest of the discussion is analogous. $\square$

4 The case $\|\cdot\|_E = \|\cdot\|_2$

We consider the case when $\|y\|_E = \left(\sum_{e\in E} w_e y_e^2\right)^{1/2} = \left\|W^{1/2}y\right\|_2$ with edge weights $w_e > 0$, $W = \mathrm{diag}(w_e, e \in E)$ and $\|\cdot\|_V = \|\cdot\|$ is an arbitrary norm on $\mathbb{R}^V$. Then we have $\|Rx\|_E^2 = x^\top Lx$ with a weighted Laplacian matrix $L = R^\top WR$. We give a reformulation of the minimization problem of $a(G)$ as a maximization problem involving the Moore-Penrose-pseudoinverse $L^\dagger$ of L and the resistance matrix T of G which are defined as follows. If L has eigenvalues (in the linear algebra sense) $0 = \nu_1 < \nu_2 \le \ldots \le \nu_n$ and spectral decomposition $L = \sum_{i=2}^n \nu_i x_i x_i^\top$ then $L^\dagger = \sum_{i=2}^n \nu_i^{-1} x_i x_i^\top$. In particular,

$$LL^\dagger = L^\dagger L = I - \mathbb{1}\mathbb{1}^\top/n.$$

The resistance distance T_{ij} of two vertices $i, j \in V$ is the quantity

$$T_{ij} = L^\dagger_{ii} + L^\dagger_{jj} - 2L^\dagger_{ij} = (e_i - e_j)^\top L^\dagger (e_i - e_j)$$

and we define the resistance matrix as $T = (T_{ij}, i, j \in V)$. With $L^\dagger_d = \mathrm{diag}(L^\dagger_{ii}, i \in V)$ we have

$$T = \mathbb{1}\mathbb{1}^\top L^\dagger_d + L^\dagger_d \mathbb{1}\mathbb{1}^\top - 2L^\dagger. \tag{4.1}$$

It is well-known that this quantity defines a metric on the vertices [KR93, Gur10]. The interpretation is as follows: an edge e of G is considered a w_e^{-1}-Ohm-resistor and T_{ij} is the total resistance between nodes i and j. The quantities $KI(G) = \frac{1}{2}\sum_{i,j\in V} T_{ij}$ and $KI'(G) = \frac{1}{2}\sum_{i,j\in V} d_i d_j T_{ij}$ are called the Kirchhoff index and the degree Kirchhoff index of G, recpectively. The degree d_i of vertex i is understood as the sum of the weights of the edges incident with the vertex i.

Theorem 4.1. *With the above notations we have $a(G) = \mu^{-1/2}$ where*

$$\mu = \max_{\substack{\|s\|^*=1 \\ \mathbb{1}^\top s=0}} s^\top L^\dagger s = -\frac{1}{2} \min_{\substack{\|s\|^*=1 \\ \mathbb{1}^\top s=0}} s^\top T s$$

Proof. If x is a normalized eigenvector for $\lambda = a(G) \neq 0$ (hence $\lambda = \left\|W^{1/2}Rx\right\|_2$) then the eigenequation reads

$$(W^{1/2}R)^\top \frac{W^{1/2}Rx}{\left\|W^{1/2}Rx\right\|_2} = \lambda s \Leftrightarrow Lx = \lambda^2 s.$$

for a suitable $s \in \partial \|x\|$ with $\mathbb{1}^\top s = 0$. Upon multiplying both sides by $s^\top L^\dagger$ we obtain

$$\lambda^2 s^\top L^\dagger s = s^\top L^\dagger L x = s^\top (x + m\mathbb{1}) = s^\top x + m s^\top \mathbb{1} = \|x\| + 0 = 1$$

for some $m \in \mathbb{R}$. It follows that

$$\lambda^{-2} \leq \max_{\substack{\|s\|^*=1 \\ \mathbb{1}^\top s=0}} s^\top L^\dagger s =: \mu \tag{4.2}$$

For the converse inequality, observe that a necessary optimality condition for t with $\|t\|^* = 1$ and $\mathbb{1}^\top t = 0$ is the existence of Lagrange multipliers ϕ, μ and ν, $(\phi, \mu, \nu) \neq (0, 0, 0)$, and a subgradient $y \in \partial \|t\|^*$ such that

$$\phi L^\dagger t - \mu y - \nu \mathbb{1} = 0,$$

see Proposition 2.1. Then $\phi \neq 0$ for if ϕ were 0 then multiplication by $t^\top$ yields $\mu \|t\|^* = \mu = 0$ and hence also $\nu = 0$. So we can assume that $\phi = 1$. Multiplication by $t^\top$ from the left gives $\mu = t^\top L^\dagger t$. Since $\mathbb{1}^\top t = 0$ we have $LL^\dagger t = t$ and hence

$$\mu^{-1} t = Ly$$

Since $1 = (\|y\|^*)^* = \|y\|$, $t \in \partial \|y\|$ and $\mathbb{1}^\top t = 0$ we get $\mu^{-1} = y^\top Ly \geq \lambda^2$ and therefore equality holds in (4.2).

Lastly, the quadratic forms of $L^\dagger$ and $-T/2$ coincide on $\mathbb{1}^\perp$ by (4.1). □

Corollary 4.2. *1. If $\|\cdot\|_V = \|\cdot\|_\infty$ then μ in Theorem 4.1 equals $\frac{1}{4} \max_{\substack{i,j \in V \\ i \neq j}} T_{ij}$.*

2. If $\|x\|_V = \|Dx\|_1$ then μ in Theorem 4.1 equals

$$2 \max \left\{ \sum_{i \in P, j \in N} d_i d_j T_{ij} + \frac{|E| - \mathrm{vol}(N)}{d_z} \sum_{i \in P} d_i d_z T_{iz} + \frac{|E| - \mathrm{vol}(P)}{d_z} \sum_{i \in N} d_i d_z T_{iz} \right\} - KI'(G)$$

where the maximum is taken over all partitions of the vertex set V which are of one of the following types:

(a) $V = P \dot{\cup} N$ with $\mathrm{vol}(P) = \mathrm{vol}(N) = |E|$ or

(b) $V = P \dot{\cup} N \dot{\cup} \{z\}$ with $\mathrm{vol}(P), \mathrm{vol}(N) < |E|$.

Observe that for $V = P \cup N$ of the first type the terms in the formula which depend on z are zero.

3. Let $\|\cdot\|_V = \|\cdot\|_1$. If $|V|$ is even then μ in Theorem 4.1 equals

$$2 \max \left\{ \sum_{i \in P, j \in N} T_{ij} \colon P, N \subseteq V,\ P \cap N = \emptyset,\ |P| = |N| = \frac{|V|}{2} \right\} - KI(G)$$

If $|V|$ is odd then μ equals

$$2 \max \left\{ \sum_{i \in P, j \in N} T_{ij} + \frac{1}{2} \sum_{i \in P} T_{iz} + \frac{1}{2} \sum_{i \in N} T_{iz} \right\} - KI(G)$$

where the maximum is taken over all partitions $V = P \dot{\cup} N \dot{\cup} \{z\}$ with $|P| = |N| = \frac{|V|-1}{2}$.

Proof. Since $L^\dagger$ is positive semidefinite and therefore $s \mapsto s^\top L^\dagger s$ is convex we can replace the constraint $\|s\|^* = 1$ by $\|s\|^* \le 1$ in the optimization problem in Theorem 4.1 because the maximum is taken on the boundary. In all three cases the feasible set is a convex polytope Q and hence the convex hull of its vertices. By the convexity of the objective function the maximum is attained at some vertex of Q.

1. $Q = \{s\colon \|s\|_\infty^* = \|s\|_1 \le 1,\ \mathbb{1}^\top s = 0\}$. The vertices of Q are precisely the points $(e_i - e_j)/2$, $i \ne j$. For if s is a point with two non-zero entries of like sign, say $s_i, s_j > 0$, then define $s^\pm = s \pm \epsilon(e_i - e_j)$ $(s_j \ge \epsilon > 0,)$ $\|s^\pm\|_1 = \|s\|_1$ and $s = (s^+ + s^-)/2$, hence s is not a vertex of P.

2. $Q = \{s\colon \|Ds\|_1^* = \left\|D^{-1}s\right\|_\infty \le 1,\ \mathbb{1}^\top s = 0\}$. If s is a point with at least two entries s_i, s_j with $|s_i| < d_i, |s_j| < d_j$ then we can write $s = (s^+ + s^-)/2$ as in the first part. So, if s is a vertex then there is at most one component z with $|s_z| < d_z$ and we let $P = \{i \in V\colon s_i = d_i\}$ and $N = \{i \in V\colon s_i = -d_i\}$. If $V \setminus (P \cup N) = \emptyset$ then $\mathrm{vol}(P) + \mathrm{vol}(N) = 2|E|$ and from $0 = \mathbb{1}^\top s = \mathrm{vol}(P) - \mathrm{vol}(N)$ and we get $\mathrm{vol}(P) = |E| = \mathrm{vol}(N)$, so $P \cup N$ is of type (a). If $V \setminus (P \cup N) = \{z\}$ then $\mathrm{vol}(P) + \mathrm{vol}(N) + d_z = 2|E|$ and from $0 = \mathbb{1}^\top s = \mathrm{vol}(P) - \mathrm{vol}(N) + s_z$ we get $s_z = \lambda d_z$ with $\lambda = d_z^{-1}(\mathrm{vol}(N) - \mathrm{vol}(P))$. $V = P \dot\cup N \dot\cup \{z\}$ is of type (b), because if, say, $\mathrm{vol}(P) \ge |E|$ then $s_z = \mathrm{vol}(N) - \mathrm{vol}(P) = 2(|E| - \mathrm{vol}(P)) - d_z \le -d_z$, a contradiction. Conversely, for every partition of types (a) or (b) we get a feasible vector by setting $s_i = d_i$ on P, $s_i = -d_i$ on N and $s_z = \mathrm{vol}(N) - \mathrm{vol}(P)$.

In either case $s = D(\mathbb{1}_P - \mathbb{1}_N + \lambda e_z) = D(t_P - t_N)$ where we let $t_P = \mathbb{1}_P + \frac{|E| - \mathrm{vol}(P)}{d_z} e_z$ and $t_N = \mathbb{1}_N + \frac{|E| - \mathrm{vol}(N)}{d_z} e_z$. Observe that $t_P + t_N = \mathbb{1}$. Then we have

$$\begin{aligned}
s^\top L^\dagger s &= -\frac{1}{2}(t_P - t_N)^\top DTD(t_P - t_N) \\
&= -\frac{1}{2}\left((t_P + t_N)^\top DTD(t_P + t_N) - 4s_P^\top DTDt_N\right) \\
&= 2t_P^\top DTDt_N - KI'(G)
\end{aligned}$$

which yields the desired formula.

3. The proof is very similar to the proof of 2. A vertex s is seen to be of the form $\mathbb{1}_P - \mathbb{1}_N$ with P, N as in the statement. In the case when $|V|$ is odd write $s = t_P - t_N$ with $t_P = \mathbb{1}_P + e_z/2$ and $t_N = \mathbb{1}_N + e_z/2$. Then $s^\top L^\dagger s = 2t_p^\top T t_N - KI(G)$ yields the formula. $\square$

A result analogous to Theorem 4.1 holds for the largest eigenvalue, namely

$$\lambda_n^{-2} = \left(\max_{x \in \mathcal{X}} \mathcal{R}(x)\right)^{-2} = \min_{\substack{\|s\|^* = 1 \\ \mathbb{1}^\top s = 0}} s^\top L^\dagger s = -\frac{1}{2} \max_{\substack{\|s\|^* = 1 \\ \mathbb{1}^\top s = 0}} s^\top T s$$

Especially when $\|\cdot\|_V = \|\cdot\|_\infty$ we have that $\max_{x \in \mathcal{X}_\infty} \mathcal{R}(x)^2 = \max_{\|x\|_\infty \le 1} x^\top L x = 4\mathrm{maxcut(G)}$ because by convexity the maximum is attained at some vertex of $\mathcal{S}_\infty$, *i.e.* some $x \in \{\pm 1\}^V$. This gives the following relation between maxcut and resistance matrix.

Corollary 4.3.

$$\min_{\substack{\|s\|_1 = 1 \\ \mathbb{1}^\top s = 0}} s^\top L^\dagger s = -\frac{1}{2} \max_{\substack{\|s\|_1 = 1 \\ \mathbb{1}^\top s = 0}} s^\top T s = \frac{1}{4\mathrm{maxcut}(G)}.$$

5 The case $\|\cdot\|_V = \|\cdot\|_\infty$, $\|\cdot\|_E = \|\cdot\|_p$, $1 \le p \le \infty$

It turns out that the parameters $a_{p,\infty}(G) = a_{\|\cdot\|_p,\|\cdot\|_\infty}(G)$ can be viewed as the inverse diameter of the graph with respect to a variant of the resistance distance.

For $k \neq l \in V$ define the quantities

$$c_{k,l} = \min\left\{ f(x) = \|Rx\|_p = \left(\sum_{ij\in E} |x_i - x_j|^p \right)^{1/p} : x \in \mathbb{R}^V \text{ and } x_k = 1, x_l = -1 \right\}. \tag{5.1}$$

These quantities are well-defined and positive because G is connected and clearly $c_{k,l} = c_{l,k}$. The problem of determining $a_{p,\infty}(G)$ reduces to solving the subproblems (5.1) as the following proposition shows:

Proposition 5.1. *For $p < \infty$ an optimal solution x to the optimization problem in 5.1 satisfies $|x_i| \le 1$ and for $p = \infty$ there exists an optimal x with this property. In particular $a_{p,\infty}(G) = \min\{c_{k,l} \colon k, l \in V,\ k \neq l\}$.*

Proof. If, say, $\max_{i\in V} x_i = c > 1$ then consider the set $C = \{i \in V \colon x_i > 1\}$ and the vector y with $y_i = 1$ if $i \in C$ and $y_i = x_i$ otherwise. If $ij \in E(C)$ then $0 = |y_i - y_j|^p \le |x_i - x_j|^p$ and if $ij \in E(C, V \setminus C)$ with $i \in C, j \in V \setminus C$ then

$$|y_i - y_j|^p = |1 - x_j|^p < |x_i - x_j|^p.$$

The set $E(C, V \setminus C)$ is non-empty because G is connected and $V \setminus C$ contains at least the vertices k and l. Hence y yields a strictly smaller objective value, a contradiction. If $p = \infty$, then replacing x by y does not increase the objective value. The argument for $\min x_i < -1$ is similar. $\square$

Notice that we can equivalently write $c_{k,l} = \min\{f(x) \colon x_k - x_l = 2\}$. This is a convex minimization problem, hence, for a necessary and sufficient optimality condition for x is the existence of a multiplier λ (and $s \in \partial \|Rx\|_p$, if $p \in \{1, \infty\}$) such that

$$L_p(x) = \lambda^p \frac{e_k - e_l}{2} \qquad \left(\text{respectively, } R^\top s = \lambda \frac{e_k - e_l}{2} \right) \tag{5.2}$$

For $p = 2$ this yields $c_{k,l}^{-2} = \frac{1}{4}(e_k - e_l)^\top L^\dagger (e_k - e_l)$ in accordance with Corollary 4.2. Observe further that if we choose an optimal x with $x_k = 1 = -x_l$ then $(e_k - e_l)/2 \in \partial \|x\|_\infty$. In summary we have

Corollary 5.2. *For $k, l \in V$, $k \neq l$, $c_{k,l}$ is an eigenvalue of $\mathcal{R}_{p,\infty}$.*

Equation (5.2) is studied in [Gur10] and it turns out that the numbers $c_{k,l}$ can be used to define a metric on the vertex set.

Proposition 5.3 (Theorem 1 in [Gur10]). *Let $1 < p < \infty$ and define $\rho_{k,l} = (2/c_{k,l})^p$ and $\rho_{k,k} = 0$ for $k \in V$. Then for any triplet $k, l, z \in V$ we have*

$$\rho_{kl}^{1/(p-1)} \le \rho_{kz}^{1/(p-1)} + \rho_{zl}^{1/(p-1)}$$

with equality if and only if z lies on any k-l-path in G. In particular, if G is a tree then $\rho_{kl}^{1/(p-1)}$ is the usual graph distance.

Finally, if $p \le 2$ then also $\rho_{kl} \le \rho_{kz} + \rho_{zl}$ holds and for $p = 2$ we have that ρ_{kl} equals the resistance distance $T_{k,l}$.

Remark: In [Gur10] the limits of $c_{k,l}$ for $p \to 1$ and $p \to \infty$ are shown to exist and the combinatorial interpretations given below are stated. However, the non-smooth cases $p = 1$ and $p = \infty$ are not treated explicitly.

Proposition 5.4. *1. Let $p = 1$ and $\rho_{k,l} = 2/c_{k,l}$. We have that $c_{k,l} = 2\mathrm{mincut}(G, k, l)$ where $\mathrm{mincut}(G, k, l)$ denotes the size of a minimum k-l-cut in G (or, equivalently, a maximum k-l-flow). The numbers $\rho_{k,l}$ thus define an ultrametric on V (with the convention that $\mathrm{mincut}(G, k, k) = \inf \emptyset = \infty$ and $1/\infty = 0$),* i.e. *the strong triangle inequality $\rho_{k,l} \le \max\{\rho_{k,z}, \rho_{l,z}\}$ holds for all $k, l, z \in V$.*

2. Let $p = \infty$. Then $\rho_{k,l} = 2/c_{k,l}$ is the length of a shortest k-l-path in G.

Proof. 1. Any y of the form $y = \mathbb{1}_X - \mathbb{1}_{V\setminus X}$ for some vertex set $X \ni k, l \notin X$ yields a feasible solution with objective value $2\mathrm{cut}(X)$, so $c_{k,l} = 2/\rho_{k,l} \le 2\mathrm{mincut}(G, k, l)$. To show that equality holds we consider an arbitrary optimal solution y. Necessarily, there exists $\lambda \in \mathbb{R}$ and $s \in \partial \|Ry\|_1 = (\mathrm{sign}(y_i - y_j), ij \in E)$ with $R^\top s = \lambda(e_k - e_l)/2$ (compare (5.2)) and hence $\lambda = c_{k,l}$. Now $x = \mathbb{1}_{P_y \cup Z_y} - \mathbb{1}_{N_y}$ satisfies $x_k - x_l = 2$ and we also have $s \in \partial \|Rx\|_1$ (see Proposition 3.1). Thus x is also optimal and $c_{k,l} = 2\mathrm{cut}(N_y) \le 2\mathrm{mincut}(G, k, l)$.

It remains to prove the strong triangle inequality

$$c_{k,l}^{-1} \le \max\left\{c_{k,z}^{-1}, c_{l,z}^{-1}\right\} \Leftrightarrow 1 \le \max\{c_{k,l}/c_{k,z}, c_{k,l}/c_{l,z}\}.$$

To that end let $c_{k,l} = 2\mathrm{mincut}(G, k, l) = 2\mathrm{cut}(X)$ with $k \in X$ and $l \in V \setminus X$. If $z \in V \setminus X$ then $\mathrm{mincut}(G, k, l) \ge \mathrm{mincut}(G, k, z)$ and $c_{k,l}/c_{k,z} \ge 1$. Otherwise, if $z \in X$ we get $c_{k,l}/c_{l,z} \ge 1$.
2. The optimum value $c_{k,l}$ is positive so we can reformulate the original problem

$$\begin{aligned}
c_{k,l} &= \min_{x_k=1, x_l=-1} \max_{ij\in E} |x_i - x_j| \\
&= \min\{\lambda \colon x_k = 1, x_l = -1,\ |x_i - x_j| \le \lambda\ \forall ij \in E\} \\
&= \max\left\{\frac{1}{\lambda} \colon \frac{x_k}{\lambda} = \frac{1}{\lambda}, \frac{x_l}{\lambda} = -\frac{1}{\lambda},\ \left|\frac{x_i}{\lambda} - \frac{x_j}{\lambda}\right| \le 1\ \forall ij \in E, \lambda > 0\right\}^{-1} \\
&= \max\{y_k \colon y_l = -y_k,\ -1 \le y_i - y_j \le 1\ \forall ij \in E\}^{-1}
\end{aligned}$$

For $a \in V$ denote by l_{ka} the length of a shortest k-a-path in G. Adding the inequalities corresponding to the edges in a shortest k-l-path shows that $|y_k - y_l| \le l_{kl}$ and thus $y_k \le l_{kl}/2$ by the first constraint. Equality is attained for $y_k = l_{kl}/2$ and $y_a = y_k - l_{ka}$ for $a \in V \setminus \{k\}$. $\square$

6 The case $\|\cdot\|_E = \|\cdot\|_\infty$

Assuming that G is connected we write according to Lemma 2.7

$$a_{\|\cdot\|_\infty, \|\cdot\|_V}(G)^{-1} = \max_{\|Rx\|_\infty = 1} d(x) = \max\left\{d(x) \colon x \in \mathbb{R}^V,\ -1 \le x_i - x_j \le 1\ \forall ij \in E\right\} \tag{6.1}$$

where the "=" in the constraint can be replaced by "≤" by positive homogeneity of both the constraint and the objective functions. By the invariance under translations by scalar multiples of $\mathbb{1}$ we can reformulate (6.1) as a maximum norm graph embedding problem

$$a_{\|\cdot\|_\infty, \|\cdot\|_V}(G)^{-1} = \max\left\{\|x\|_V \colon x \in \mathbb{R}^V,\ 0 \in \mathbb{1}^\top \partial \|x\|_V,\ -1 \le x_i - x_j \le 1\ \forall ij \in E\right\}. \tag{6.2}$$

Clearly, any optimal solution to the latter problem solves the former one; conversely, if x solves (6.1) then $x^* = x - m_x \mathbb{1}$ solves (6.2) where $d(x) = \|x^*\|_V$.

Given a feasible point x of (6.1) (resp. (6.2)) we define the *active subgraph* as $G_x = (V, E_x = \{ij \in E \colon |x_i - x_j| = 1\})$. Clearly, G_x is always bipartite. It can also be chosen connected

Proposition 6.1. *1. There is an optimal solution x to* (6.1) *(resp.* (6.2)*) for which G_x is connected. Furthermore, x can be chosen integral. If $x^* = x - m^* \mathbb{1}$ such that $d(x) = \|x^*\|_V$ then x^* solves* (6.2)*.*

2. Let x be optimal with connected G_x, $V_{=c} = \{i \in V \colon x_i = c\}$ and analogously $V_{<c}$ and $V_{>c}$. If $\min_{i \in V} < c < \max_{i \in V}$ and $V_c \neq \emptyset$ then V_c separates $V_{<c}$ and $V_{>c}$.

3. If $\|\cdot\|_V = \|\cdot\|_p$ with $1 < p < \infty$ then G_x is connected for any optimal x.

Proof. 1. Both D and the feasible region of (6.1) are invariant under translations by scalar multiples of $\mathbb{1}$ so we can restrict D to the polytope $\mathcal{P} = \{x \in \mathbb{R}^V \colon \mathbb{1}^\top x = 0,\ -\mathbb{1}_E \leq Rx \leq \mathbb{1}_E\}$ without changing the optimum. By convexity of D the optimum is attained on at least one vertex of $\mathcal{P}$. Every vertex of $\mathcal{P}$ corresponds to a choice of $|V| - 1$ linearly independent rows of R and therefore some spanning tree of G. Hence x can be chosen such that G_x contains an active spanning tree.

2. By the constrains there cannot be an edge between $V_{<c} \subset \{\alpha \colon \alpha \leq c - 1\}$ and $V_{>c} \subset \{\alpha \colon \alpha \geq c + 1\}$.

3. Let x be an optimal solution to (6.2) and assume that G_x is disconnected, say $V = C_1 \dot\cup C_2$ with $E(C_1, C_2) \cap E_x = \emptyset$. Let x_1 and x_2 be the restrictions of x to C_1 and C_2, respectively. Then $\|x\|_p^p = \|x_1\|_p^p + \|x_2\|_p^p$ and $\nabla \|x\|_p^p = p\Phi_p(x) = p(\Phi_p(x_1) + \Phi_p(x_2))$. By the constraints we have $\mathbb{1}_{C_1}^\top \Phi_p(x_1) + \mathbb{1}_{C_2}^\top \Phi_p(x_2) = 0$ and without loss of generality we assume that $g_1 := \mathbb{1}_{C_1}^\top \Phi_p(x_1) \geq 0 \geq \mathbb{1}_{C_2}^\top \Phi_p(x_2) =: g_2$. Recall that g_1 is the derivative of $m \mapsto \|x_1 + m\mathbb{1}_{C_1}\|_p^p$ at $m = 0$, analogously for g_2. Thus, if $\epsilon, \delta > 0$ then $\|x_1 + \epsilon \mathbb{1}_{C_1}\|_p > \|x_1\|_p$ and $\|x_2 - \delta \mathbb{1}_{C_2}\|_p > \|x_2\|_p$ by the strict convexity of $\|\cdot\|_p$. Clearly, ϵ and δ can be chosen in such a way that $x + \epsilon \mathbb{1}_C - \delta \mathbb{1}_D$ is feasible and we have a solution with a larger objective value, a contradiction. ☐

We now take a closer look at the 1- and 2-norm cases.

6.1 $\|\cdot\|_E = \|\cdot\|_\infty$, $\|\cdot\|_V = \|\cdot\|_1$

The maximum norm embedding problem reads (see Lemma 2.10)

$$\text{maximize} \sum_{i \in V} |x_i| \quad \text{s.t.} \begin{cases} x \in \mathbb{R}^V, \\ |p_x - n_x| \leq z_x, \\ |x_i - x_j| \leq 1 \ (ij \in E). \end{cases} \tag{6.3}$$

We have already seen that there exists an optimal x for which G_x is connected, i.e. x lives on a lattice. Furthermore such x can be assumed integer (and must be if $|V|$ is odd) because if $x \notin \mathbb{Z}^V$ then $p_x = n_x = |V|/2$ and $z_x = 0$ and we can replace x by $x - m\mathbb{1}$ where $m \in \{\max N_x, \min P_x\}$, that is, m is a median of x.

So assume that $x \in \mathbb{Z}^V$ with G_x connected. Every $i \in P_x$ has a neighbour embedded at $x_i - 1$ because otherwise i could be embedded at $x_i + 1$ without affecting feasibility and thereby increase the objective value, similarly for $x \in N_x$. Thus every vertex is embedded at shortest path distance from Z_x and $\|x\|_1$ is the sum of the graph distances of the vertices from Z_x. Recall that Z_x separates P_x and N_x (assuming that both are non-empty). We thus have the following formulation as a combinatorial optimization problem

Proposition 6.2. *Denote by $\delta(i,j)$ the usual shortest path graph distance, for $Z \subseteq V$ define $\delta(i,Z) = \min\{\delta(i,z)\colon z \in Z\}$ and $\delta(Z) = \sum_{i \in V} \delta(i,Z)$. Then we have that*

$$a_{\infty,1}(G)^{-1} = \max \delta(Z) \text{ s.t. } \begin{cases} V = P \dot\cup Z \dot\cup N, \\ E(P,N) = \emptyset, \\ ||P| - |N|| \le |Z|. \end{cases}$$

Notice that neither P nor N must be non-empty.

For parallelization purposes, for example in the computation of large sparse positive definite systems, one seeks a small vertex separator which separates two subgraphs of approximately the same size. In [PSL90] a spectral technique is devised which computes an edge separator from a second eigenvector x of the Laplacian. An edge cut is obtained by partitioning the vertices according to x_i lying above or below a median of x. A vertex separator is obtained via bipartite matching (the corresponding vertex cover).

In this regard an eigenvector corresponding to $a_{\infty,1}(G)$ could be of interest for one may hope that for sparse graphs the maximization of $\delta(Z)$ in Proposition 6.2 forces Z to be a small set and thus P and N of approximately same size. It would be interesting to know wether this intuition is justified or if $a_{\infty,1}(G)$ is computable in polynomial time.

Similar considerations hold if we replace $\|\cdot\|_1$ by the degree weighted one norm $\|x\|_V = \sum_{i \in V} d_i |x_i|$. In this case the optimization problem of Proposition 6.2 asks for a decomposition $V = P \dot\cup Z \dot\cup N$ with $E(P,N) = \emptyset$ and $\mathrm{vol}(Z) \ge |\mathrm{vol}(P) - \mathrm{vol}(N)|$ such that $\delta(Z) = \sum_{i \in V} d_i \delta(i,Z)$ is maximized. A small volume separator Z would yield P and N of approximately equal volume.

6.2 $\|\cdot\|_E = \|\cdot\|_\infty$, $\|\cdot\|_V = \|\cdot\|_2$

In this case we have that $a_{\infty,2}(G)^{-2}$ is the optimum of

$$\text{maximize} \sum_{i \in V} x_i^2 \quad \text{s.t.} \begin{cases} x \in \mathbb{R}^V, \\ \sum_{i \in V} x_i = 0, \\ |x_i - x_j| \le 1 \ (ij \in E). \end{cases} \tag{6.4}$$

A relaxation of this problem is obtained when we allow embeddings into some arbitrary $\mathbb{R}^k$ rather than the line,

$$\text{maximize} \sum_{i \in V} \|x_i\|_2^2 \quad \text{s.t.} \begin{cases} x_i \in \mathbb{R}^k \ (i \in V), k \in \mathbb{N} \\ \sum_{i \in V} x_i = 0 \\ \|x_i - x_j\|_2^2 \le 1 \ (ij \in E) \end{cases} \tag{6.5}$$

Clearly whenever (6.5) admits a one dimensional optimal solution the optima of (6.5) and (6.4) coincide. In [GHW08] (6.5) is shown to be dual problem of a semidefinite formulation of determining the *absolute algebraic connectivity* $\widehat{a}(G)$ of a graph G. To explain that notion, let $L_w = R^\top \mathrm{Diag}(w_{ij}, ij \in E)R$ be a weighted Laplacian matrix and $\lambda_2(L_w)$ its second smallest eigenvalue. Then

$$\widehat{a}(G) = \max\left\{\lambda_2(L_w)\colon w_{ij} \geq 0,\ \sum_{ij\in E} w_{ij} = 1\right\} \tag{6.6}$$

The optimum of (6.5) is $\widehat{a}(G)^{-1}$. If $X = (x_i, i \in V) \in \mathbb{R}^{k\times V}$ is any optimal solution to (6.5) then the (non-zero) rows of X are eigenvectors of any optimally weighted Laplacian of (6.6), $L_w X^\top = \lambda_2(L_w)X^\top = \widehat{a}(G)X^\top$ (KKT complementarity). To summarize

Proposition 6.3. *We have 1 ⇒ 2 ⇔ 3 where*

1. $\widehat{a}(G)$ *is a simple eigenvalue of the optimally weighted* L_w.
2. (6.5) *admits a rank one solution.*
3. *The optimal values of* (6.5) *and* (6.4) *coincide, i.e.*

$$a_{\infty,2}(G) = \widehat{a}(G)^{1/2}.$$

The converse implication 2 ⇒ 1 is wrong: $K_{1,4}$ is a counterexample.

7 Conclusion and outlook

We have introduced an eigenvalue problem generalizing the graph Laplacian eigenvalues and studied the smallest non-zero eigenvalue. We looked at several non-smooth cases which yield possibly interesting combinatorial graph parameters, for example $a_{2,1}(G)$ or $a_{\infty,1}(G)$. In particular the latter raises the question if it is of use for graph partitioning when one is interested in finding a small vertex separator which separates sets of approximately equal size.

Regarding the eigenvalue problem for $\mathcal{R}_{1,1}$ and $\mathcal{R}_{1,\infty}$ it would be interesting to know which graph cuts yield in fact eigenvectors and to study the sequence of critical values of Section 2.2 in more detail.

Another interesting question is for nodal domain theorems. In [HT] they prove such a theorem for $\mathcal{R}_{p,p}$. We are currently investigating how this generalizes in our setting. As a partial result in this direction we can show that their theorem holds for $\mathcal{R}_{p,q}$ if $p < q$. We don't expect it to hold for $p > q$ in general. Consider $a_{\infty,1}(G)$ for the tree on 7 vertices with edges $\{12, 23, 14, 45, 16, 67\}$. It is optimally embedded in $(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = (-1, -2, -3, 0, 1, 0, 1)$ and hence we have three weak nodal domains not two. Similarly, $x + 4/7\mathbb{1}$ is an eigenvector for $a_{\infty,2}(G)$.

A similar generalization of the eigenvalue problem for the signless Laplacian $Q = D + A$ is straight forward: recall that for the (unsigned) incidence matrix $B \in \{0,1\}^{E\times V}$ we have $x^\top Qx = \|Bx\|_2^2$ and hence one considers a Rayleigh quotient $\|Bx\|_E \,/\, \|x\|_V$. The smallest eigenvalue $b(G) = b_{\|\cdot\|_E,\|\cdot\|_V}(G)$ is much easier to handle than $a(G)$ because criticality comes for free. The smallest (generalized) eigenvalue is 0 if and only if G has a bipartite component. The case where both norms are p-norms ($1 \leq p \leq \infty$) has been considered in [BS16]. We have not considered the case of "mixed" norms yet.

Acknowledgments This work was partially supported by CAPES Grant PROBRAL 408/13 - Brazil and DAAD PROBRAL Grant 56267227 - Germany.

References

[Amg03] S. Amghibech. Eigenvalues of the discrete p-Laplacian for graphs. *Ars Comb.*, 67:283–302, 2003.

[Amg06] S. Amghibech. Bounds for the largest p-Laplacian eigenvalue for graphs. *Discrete Mathematics*, 306(21):2762–2771, 2006.

[BH09] T. Bühler and M. Hein. Spectral clustering based on the graph p-Laplacian. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pages 81–88, New York, NY, USA, 2009. ACM.

[BLS07] T. Bıyıkoğlu, J. Leydold, and P. F. Stadler. *Laplacian eigenvectors of graphs: Perron-Frobenius and Faber-Krahn type theorems*, volume 1915 of *Lecture Notes in Mathematics*. Springer Verlag, Berlin Heidelberg, 2007.

[BS16] E.M. Borba and U. Schwerdtfeger. Eigenvalue bounds for the signless p-Laplacian. arXiv:1610.01357, 2016.

[Cha81] Kung-Ching Chang. Variational methods for non-differentiable functionals and their applications to partial differential equations. *Journal of Mathematical Analysis and Applications*, 80(1):102 – 129, 1981.

[Cha16] K. C. Chang. Spectrum of the 1-Laplacian and Cheeger's constant on graphs. *Journal of Graph Theory*, 81(2):167–207, 2016.

[Cla90] F. Clarke. *Optimization and Nonsmooth Analysis*. Society for Industrial and Applied Mathematics, 1990.

[CRS09] D. Cvetković, P. Rowlinson, and S. Simić. *An Introduction to the Theory of Graph Spectra*. Cambridge University Press, Cambridge, 2009.

[CSZ15] K.C. Chang, Sihong Shao, and Dong Zhang. The 1-Laplacian Cheeger cut: Theory and algorithms. *Journal of Computational Mathematics*, 33(5):443–467, 2015.

[GHW08] Frank Göring, Christoph Helmberg, and Markus Wappler. Embedded in the shadow of the separator. *SIAM Journal on Optimization*, 19(1):472–501, 2008.

[Gur10] Vladimir Gurvich. Metric and ultrametric spaces of resistances. *Discrete Applied Mathematics*, 158(14):1496 – 1505, 2010.

[HT] M. Hein and F. Tudisco. A nodal domain theorem and a higher-order Cheeger inequality for the graph p-Laplacian. arXiv:1602.05567, to appear in Journal of spectral theory.

[KN14] L. Kang and V. Nikiforov. Extremal problems for the p-spectral radius of graphs. arXiv:1402.3239, 2014.

[KR93] D. J. Klein and M. Randić. Resistance distance. *Journal of Mathematical Chemistry*, 12(1):81–95, Dec 1993.

[Moh89] Bojan Mohar. Isoperimetric numbers of graphs. *Journal of Combinatorial Theory, Series B*, 47(3):274 – 291, 1989.

[Moh91] B. Mohar. The Laplacian spectrum of graphs. In Y. Alavi, G. Chartrand, O. R. Oellermann, and A. J. Schwenk, editors, *Graph Theory, Combinatorics, and Applications*, volume 2, pages 871–898. Wiley, 1991.

[Moh04] B. Mohar. Graph Laplacians. In L.W. Beineke and R.J. Wilson, editors, *Topics in Algebraic Graph Theory*, volume 102 of *Encyclopedia of Mathematics and Its Applications*, pages 113–136. Cambridge University Press, Cambridge, 2004.

[PSL90] Alex Pothen, Horst D. Simon, and Kang-Pu Liou. Partitioning sparse matrices with eigenvectors of graphs. *SIAM Journal on Matrix Analysis and Applications*, 11(3):430–452, 1990.