SEPARATUM

ON THE EIGENVALUES OF TREES

by

L. LOVÁSZ and J. PELIKÁN (Budapest)

To the memory of A. RÉNYI

AKADÉMIAI KIADÓ, BUDAPEST
PUBLISHING HOUSE OF THE HUNGARIAN ACADEMY OF SCIENCES
VERLAG DER UNGARISCHEN AKADEMIE DER WISSENSCHAFTEN
MAISON D'EDITIONS DE L'ACADEMIE DES SCIENCES DE HONGRIE
ИЗДАТЕЛЬСТВО АКАДЕМИИ НАУК ВЕНГРИИ

Periodica Mathematica Hungarica Vol. 3 (1–2), (1973), pp. 175–182.

ON THE EIGENVALUES OF TREES

by

L. LOVÁSZ and J. PELIKÁN (Budapest)

To the memory of A. RÉNYI

Given a graph G (without loops and multiple edges) of n vertices labelled by $1, 2, \ldots, n$, we can form the adjacency matrix $A_G = (a_{ij})$ of G, defined by

$$a_{ij} = \begin{cases} 1 & \text{if the } i^{th} \text{ and } j^{th} \text{ vertices are joined by an edge,} \\ 0 & \text{otherwise.} \end{cases}$$

The adjacency matrix depends on the labelling of the vertices but its characteristic equation (and, consequently, its eigenvalues too) depend only on the graph G itself. As A_G is a symmetric matrix, these eigenvalues, called the eigenvalues of G, are real.

We denote by $f_G(\lambda)$ the characteristic polynomial $\det(\lambda I - A_G)$ of A_G and by $\Lambda(G)$ its largest root.

We shall begin with several general remarks on $f_G(\lambda)$ and $\Lambda(G)$, used in latter considerations. These propositions are special cases or easy consequences of general theorems on eigenvalues of non-negative matrices (see, e. g. [2] and [3]). Although they may be well-known for the reader, it may have some use to list them here.

Our main concern in this paper will be $f_G(\lambda)$ and $\Lambda(G)$ in the case when G is a tree (or more generally, a forest). We determine the maximal and minimal value of $\Lambda(G)$ among all trees of n vertices and give a method which enables us to determine the order of largest eigenvalues of two different trees in several cases.

NOTATIONS. $V(G)$ and $E(G)$ are the sets of vertices and edges of G, respectively. $G \cong G'$ means that G and G' are isomorphic. If G_1 and G_2 are arbitrary graphs then $G_1 + G_2$ is defined as follows: we consider a $G_1' \cong G_1$ and a $G_2' \cong G_2$ such that $V(G_1') \cap V(G_2') = \emptyset$ and let $V(G_1 + G_2) = V(G_1') + V(G_2')$, $E(G_1 + G_2) = E(G_1') + E(G_2')$. $G_1 + G_2$ is uniquely determined up to isomorphism. If $e \in E(G)$ and $x \in V(G)$ then $G - e$, $G - x$, $G - [e]$ denote the graphs arising from G by the removal of the edge e, of the vertex x and of the endpoints of e, respectively. If $e = (x,y)$ is a non-adjacent pair of vertices of G then $G \cup e$ denotes the graph obtained by adding the edge e to G. $G' \subseteq G$ means that $V(G') = V(G)$, $E(G') \subseteq E(G)$.

PROPOSITION 1. *If G has at least one edge then $\Lambda(G) > 0$ and there is an eigenvector belonging to $\Lambda(G)$ with non-negative coordinates. If G is connected then $\Lambda(G)$ has multiplicity 1 and a positive eigenvector.*

PROPOSITION 2. *If G' is a subgraph of G then $\Lambda(G') \leq \Lambda(G)$.*

PROPOSITION 3. *Let G_1, G_2 be two graphs on the same set of vertices. Then $\Lambda(G_1 \cup G_2) \leq \Lambda(G_1) + \Lambda(G_2)$.*

PROPOSITION 4. *Let $\varphi(G)$, $\Phi(G)$ denote the minimum and maximum valency of G. Then*

$$\max\,(\varphi(G), \sqrt{\Phi(G)}) \leq \Lambda(G) \leq \Phi(G).$$

PROPOSITION 5. *A graph is bipartite iff its spectrum is symmetric to the origin.*

PROPOSITION 6. *A connected graph is bipartite iff $-\Lambda(G)$ is an eigenvalue of it.*

Our investigations will be based on the following

LEMMA 1. *If G is a forest and $e \in E(G)$ then*

$$f_G(\lambda) = f_{G-e}(\lambda) - f_{G-[e]}(\lambda).$$

PROOF. As G is a forest we can label its vertices in such a way that e joins the points k and $k+1$ and there is no other edges between a point i $(1 \leq i \leq k)$ and a point j $(k+1 \leq j \leq n)$. Now the Laplace expansion of the determinant $\det(\lambda I - A_G)$ by its first k columns gives the equality of the lemma.

THEOREM 1. *If G is a forest then*

$$f_G(\lambda) = \lambda^n - c_1\lambda^{n-2} + c_2\lambda^{n-4} \pm \ldots + (-1)^{\left[\frac{n}{2}\right]} c_{\left[\frac{n}{2}\right]} \lambda^{n-2\left[\frac{n}{2}\right]}$$

where c_k denotes the number of all k-element independent edge-systems in G.

PROOF. We proceed by induction on the number of edges of G. For the empty graph of n vertices the theorem is obvious.

Let now e be an edge of G. Then $c_k = c_k' + c_k''$ where c_k' and c_k'' are the numbers of k-element independent edge-systems not containing e and containing e, respectively. Note that thus c_k' is the number of k-element independent edge-systems in $G-e$ while c_k'' is the number of $(k-1)$-element independent edge-systems in $G-[e]$. Now by induction $(-1)^k c_k'$ is the coefficient of λ^{n-2k} in $f_{G-e}(\lambda)$ and $(-1)^{k-1}c_k''$ is the coefficient of λ^{n-2k} in $f_{G-[e]}(\lambda)$. By Lemma 1 this proves the theorem.

REMARK. This theorem implies but it is easy to see directly too that for a forest G,

$$|A_G| = \begin{cases} 1 & \text{if } G \text{ has a 1-factor,} \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, Theorem 1 follows from this observation. The coefficient of λ^k in $\det(\lambda I - A_G)$ is the sum of all symmetric subdeterminants of A_G of order $n-k$, i.e. the sum of all $n-k$-element spanned sub-forests of G. This gives the formula of the theorem.

We introduce notations for three important special forests: let E_n, S_n and P_n be the empty graph, the star and the path of n vertices, respectively.

Fig. 1

These graphs are extreme in the following sense: $\Lambda(E_n) = 0$ is the least among the largest eigenvalues of forests (this is trivial); $\Lambda(P_n) = 2\cos\dfrac{\pi}{n+1}$ is the least among the largest eigenvalues of trees with n points; finally, $\Lambda(S_n) = \sqrt{n-1}$ is the largest among the largest eigenvalues of forests (or trees) with n points.

To prove this we shall investigate a more general problem: We order all trees (or forests) by their largest eigenvalues; is it possible to describe this ordering by graph-theoretical means? A heuristic description of this ordering, suggested by examination of special cases, is the "density" of the tree: P_n is the less "dense", S_n is the most "dense" tree and in general the greater $\Lambda(G)$ the more dense G.

Let G and G' be forests of n vertices. Instead of the order of $\Lambda(G)$ and $\Lambda(G')$ we introduce the following more complicated but better applicable notion: let $G' \prec G$ iff $f_{G'}(\lambda) \geq f_G(\lambda)$ for every $\lambda \geq \Lambda(G)$. Obviously, $G' \prec G$ implies $\Lambda(G') \leq \Lambda(G)$. Conversely this is not true even for trees, as shown by the graphs of Fig. 1. However, it is easy to see that $G' \prec G$ is a partial ordering.

LEMMA 2. *If $G' \subseteq G$ then $G' \prec G$.*

PROOF. We may assume $G' = G - e$. Let $\lambda \geq \Lambda(G)$. By Proposition 2, $\lambda \geq \Lambda(G - [e])$, thus $f_{G-[e]}(\lambda) \geq 0$, hence by Lemma 1,

$$f_G(\lambda) = f_{G'}(\lambda) - f_{G-[e]}(\lambda) \leq f_{G'}(\lambda).$$

LEMMA 3. *Let G, G' be forests of n points, $e \in E(G)$, $e' \in E(G')$ and assume that*

$$G' - e' \prec G - e, \qquad G' - [e'] \succ G - [e].$$

Then $G' \prec G$.

PROOF. Let $\lambda \geq \Lambda(G)$. Then, by Proposition 2, $\lambda \geq \Lambda(G - e)$ and thus by the assumption,

$$f_{G-e}(\lambda) \leq f_{G'-e'}(\lambda).$$

Again by Proposition 2, $\lambda \geq \Lambda(G - e) \geq \Lambda(G' - e') \geq \Lambda(G' - [e'])$, hence

$$f_{G-[e]}(\lambda) \geq f_{G'-[e']}(\lambda).$$

By Lemma 1,

$$f_G(\lambda) = f_{G-e}(\lambda) - f_{G-[e]}(\lambda) \leq f_{G'-e'}(\lambda) - f_{G'-[e']}(\lambda) = f_{G'}(\lambda).$$

THEOREM 2. *If G is a forest of n vertices then $E_n \prec G \prec S_n$.*

PROOF. $E_n \subseteq G$, hence by Lemma 2 $E_n \prec G$. On the other hand, we prove $G \prec S_n$ by induction on n. For $n = 1$ it is trivial, similarly for $G = E_n$. Let x be a vertex of G of valency 1 and let e denote the edge incident with it. Let g be an edge of S_n. Then

$$f_{G-e}(\lambda) = \lambda f_{G-x}(\lambda), \qquad f_{S_n-g}(\lambda) = \lambda f_{S_{n-1}}(\lambda)$$

and since by induction $G - x \prec S_{n-1}$, this implies

$$G - e \prec S_n - g.$$

Since $S_n - [g] = E_{n-2}$, we have obviously

$$G - [e] \succ S_n - [g].$$

By Lemma 3, this implies $G \prec S_n$.

THEOREM 3. *If G is a tree of n vertices then $P_n \prec G$.*

PROOF. Consider a tree G such that there is no other tree G' such that $G' \prec G$; we have to prove $G = P_n$. Assume indirectly that there exist vertices of valency ≥ 3 in G. Let x be a point of valency ≥ 3 such that a certain component of $G - x$ does not contain further points having valency ≥ 3 in G. This component is a path $(a_1, \ldots, a_k)$, a_1 being joined to x. Let $e = (x, b)$ be another edge incident with x and put $e' = (a_k, b)$, $G' = G - e \cup e'$. It is easy to see that G has more endpoints than G', hence $G \not\cong G'$. Furthermore, $G - e = G' - e'$ and $G - [e]$ is isomorphic to a subgraph of $G' - [e']$. Hence by Lemmata 2 and 3 we obtain $G' \prec G$, a contradiction.

To complete Theorems 2 and 3 we have to determine the largest eigenvalues of S_n and P_n.

THEOREM 4.

$$\Lambda(S_n) = \sqrt{n-1}, \quad \Lambda(P_n) = 2\cos\frac{\pi}{n+1}.$$

PROOF. The first proposition follows easily from Theorem 1, since by this theorem

$$f_{S_n}(\lambda) = \lambda^n - (n-1)\,\lambda^{n-2}.$$

The largest root of this polynomial is $\sqrt{n-1}$ indeed.

In the case of P_n the formula of Theorem 1 is too complicated, therefore we deduce another formula for $f_{P_n}(\lambda)$. Lemma 1 gives

$$f_{P_n}(\lambda) = \lambda f_{P_{n-1}}(\lambda) - f_{P_{n-2}}(\lambda)$$

which can be considered to be a recursive definition of $f_{P_n}(\lambda)$. It is known that a sequence $\{x_n\}$ defined by the recursion

$$x_n = ax_{n-1} + bx_{n-2}, \quad a^2 + 4b \neq 0$$

is of the form

$$x_n = c_1 y_1^n + c_2 y_2^n$$

where $y_{1,2} = \frac{a \pm \sqrt{a^2+4b}}{2}$ and c_1, c_2 are determined by the initial values of the sequence $\{x_n\}$. In our case we obtain

$$f_{P_n}(\lambda) = \frac{1}{\sqrt{\lambda^2-4}}\left[\left(\frac{\lambda+\sqrt{\lambda^2-4}}{2}\right)^{n+1} - \left(\frac{\lambda-\sqrt{\lambda^2-4}}{2}\right)^{n+1}\right].$$

Thus, the eigenvalues of P_n satisfy

$$\frac{\lambda+\sqrt{\lambda^2-4}}{2} = \varepsilon\,\frac{\lambda-\sqrt{\lambda^2-4}}{2} \tag{2}$$

where $\varepsilon^{n+1} = 1$. λ being real we obtain

$$\lambda^2 = \frac{(1+\varepsilon)^2}{\varepsilon} = |1+\varepsilon|^2.$$

This means that every eigenvalue of P_n is of the form

$$\lambda = \pm|1+\varepsilon| = \pm\sqrt{2+2\cos\frac{2k\pi}{n+1}} = \pm 2\cos\frac{k\pi}{n+1}$$

where $\varepsilon = e^{\frac{2k\pi}{n+1}}$, $k = 0, \ldots, n$. One easily checks that all these numbers satisfy (2), hence if $\varepsilon \neq 1$ they satisfy $f_{P_n}(\lambda) = 0$. There are just n different numbers of the form $2\cos\frac{k\pi}{n+1}$ $k = 1, \ldots, n$, therefore these and only these numbers are the eigenvalues of P_n. This proves the theorem.

Remark 1. It is easy to see that

$$f_{P_n}(\lambda) = \sum_{k=0}^{[\frac{n}{2}]} (-1)^k \binom{n-k}{k} \lambda^{n-2k},$$

but this form of f_{P_n} does not lead to the proof of Theorem 6.

Remark 2. We see that the sequence $\Lambda(P_n)$ is monotone increasing and tends to 2. The first of these statements is a consequence of Proposition 2 while the fact that $\Lambda(P_n) \leq 2$ follows from a theorem of Hoffman which states that the least eigenvalue of a line-graph is ≥ -2 (see [1]).

We determine the most "dense" tree after S_n and the less "dense" tree after P_n. These will be the trees S'_n, P'_n defined by $V(S'_n) = \{1, \ldots, n\}$, $E(S'_n) = \{(1, 2), (2, n), (3, n), \ldots, (n-1, n)\}$; $V(P'_n) = \{1, \ldots, n\}$, $E(P'_n) = \{(1, 3), (2, 3), (3, 4), \ldots, (n-1, n)\}$. We need

Lemma 4. *If* $k + l = k' + l'$, $k < k' \leq l'$ *then* $P_k + P_l \prec P_{k'} + P_{l'}$.

Proof. We use induction on $k + l$. It is enough to deal with the case $k' = k + 1$. Let a, b, a', b' be endpoints of, in order, P_k, P_l, P_{k+1} P_{l-1} and let e, g be the edges incident with b and a', respectively. Now

$$P_k + P_l - e \simeq P_{k+1} + P_{l-1} - g$$

and

$$P_k + P_l - [e] = P_k + P_{l-2},\ P_{k+1} + P_{l-1} - [g] = P_{k-1} + P_{l-1}.$$

By induction $P_{k-1} + P_{l-1} \prec P_k + P_{l-2}$ (since $k - 1 < k \leq l - 2$), this implies by Lemma 3 that $P_k + P_l \prec P_{k+1} + P_{l-1}$.

Theorem 5. *Any tree G of n vertices different from P_n satisfies $P'_n \prec G$.*

Proof. Consider a tree $G \not\simeq P_n$ such that $G' \prec G$, $G' \not\simeq G$ holds only for $G' = P_n$. We show that $G = P'_n$. The argument followed in the proof of Theorem 3 gives that G has three endpoints, i.e. G consists of three paths having one common endpoint x. Let $(a_0, a_1, \ldots, a_k)$ and $(b_0, b_1, \ldots, b_l)$ be the two shorter paths, $x = a_0 = b_0$. Put $e = (x, a_1)$, $e' = (b_{l-1}, a_1)$, $G' = G - e \cup e'$. Then $G - e = G' - e'$, and by Lemma 4, $G - [e] \prec G' - [e']$. Hence by Lemma 3 we have $G' \prec G$, which shows by the minimality of G that G and G' are isomorphic, i.e. $l = 1$. Similarly $k = 1$, which proves the theorem.

THEOREM 6. *Any forest G of n vertices different from S_n satisfies $G \prec S'_n$.*

PROOF. We use induction on n. For $n \leq 3$ or $G = P_n$ the statement is trivial. We may assume that G is a tree. Let e be an edge of G such that an endpoint x of e is an endpoint of the tree and $G - x$ is not a star. Let $e' = (3, n) \in E(S'_n)$. By induction $G - e \prec S'_n - e'$, on the other hand obviously $G - [e] \succ S'_n - [e']$, thus by Lemma 3 $G \prec S'_n$.

REMARK 1. One could prove the relation $\prec$ among many other pairs of trees. Thus e. g. the tree obtained by joining $k - 1$ new points to an endpoint of a path is "minimal" among all trees having a point of valency $\geq k$.

REMARK 2. There is a different proof of Theorems 2 and 6 based on Theorem 1. We outline the proof of the essential part of Theorem 2, namely that, for any tree G, $G \prec S_n$.

A simple calculation shows that

$$c_{k+1} \leq \frac{c_k(n-1-k)}{k+1} \leq (n-1)\, c_k \qquad (k = 0, 1, \ldots;\ c_0 = 1).$$

Hence

$$f_G(\lambda) = \sum_k (c_{2k}\lambda^{n-4k} - c_{2k+1}\lambda^{n-4k-2}) \geq \sum_k c_{2k}\big(\lambda^{n-4k} - (n-1)\,\lambda^{n-4k-2}\big) =$$

$$= f_{S_n}(\lambda) \sum_k \frac{c_{2k}}{\lambda^{4k}}$$

which shows that, for $\lambda \geq \Lambda(S_n)$, we have indeed

$$f_G(\lambda) \geq f_{S_n}(\lambda).$$

REMARK 3. The largest eigenvalues of P'_n and S'_n can be calculated similarly to the proof of Theorem 4; we give here the results:

$$\Lambda(P'_n) = 2\cos\frac{\pi}{2n-2}; \quad \Lambda(S'_n) = \sqrt{\frac{n-1+\sqrt{n^2-6n+13}}{2}}.$$

The authors are thankful to P. GÁCS for his valuable suggestions.

Added in proof: In a recent paper of A. MOWSHOWITZ (*J. Combinatorial Theory Ser. B* **12** (1972), 177—193) Lemma 1, Theorem 1 and Remark 1 after Theorem 4 have been proved.

REFERENCES

[1] A. J. Hoffman, Some recent results on spectral properties of graphs, *Beiträge zur Graphentheorie (Kolloquium, Manebach, 1967)*, Leipzig, 1968, 75—80.
[2] F. R. Gantmacher, *Matrizenrechnung* II, Berlin, 1959.
[3] M. Doob, Graphs with a small number of distinct eigenvalues, *Ann. New York Acad. Sci.* **175** (1970), 104—110.

(Received October 27, 1971)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
GEOMETRIAI TANSZÉK
BUDAPEST, VIII., MÚZEUM KRT. 6–8.

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
ALGEBRA ÉS SZÁMELMÉLET TANSZÉK
BUDAPEST, VIII., MÚZEUM KRT. 6–8.
HUNGARY