BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 59, Number 1, January 2022, Pages 31–58
https://doi.org/10.1090/bull/1722
Article electronically published on February 18, 2021

EIGENVECTORS FROM EIGENVALUES: A SURVEY OF A BASIC IDENTITY IN LINEAR ALGEBRA

PETER B. DENTON, STEPHEN J. PARKE, TERENCE TAO, AND XINING ZHANG

ABSTRACT. If A is an $n \times n$ Hermitian matrix with eigenvalues $\lambda_1(A), \dots, \lambda_n(A)$ and $i, j = 1, \dots, n$, then the jth component $v_{i,j}$ of a unit eigenvector v_i associated to the eigenvalue $\lambda_i(A)$ is related to the eigenvalues $\lambda_1(M_j), \dots, \lambda_{n-1}(M_j)$ of the minor M_j of A formed by removing the jth row and column by the formula

$$|v_{i,j}|^2 \prod_{k=1;k\neq i}^{n} (\lambda_i(A) - \lambda_k(A)) = \prod_{k=1}^{n-1} (\lambda_i(A) - \lambda_k(M_j)).$$

We refer to this identity as the *eigenvector-eigenvalue identity* and show how this identity can also be used to extract the relative phases between the components of any given eigenvector. Despite the simple nature of this identity and the extremely mature state of development of linear algebra, this identity was not widely known until very recently. In this survey we describe the many times that this identity, or variants thereof, have been discovered and rediscovered in the literature (with the earliest precursor we know of appearing in 1834). We also provide a number of proofs and generalizations of the identity.

1. INTRODUCTION

If A is an $n \times n$ Hermitian matrix, we denote its n real eigenvalues by $\lambda_1(A), \dots, \lambda_n(A)$. The ordering of the eigenvalues will not be of importance in this survey, but for sake of concreteness let us adopt the convention of nondecreasing eigenvalues:

$$\lambda_1(A) \leq \cdots \leq \lambda_n(A).$$

If $1 \leq j \leq n$, let M_j denote the $n-1 \times n-1$ minor formed from A by deleting the jth row and column from A. This is again a Hermitian matrix, and thus has $n-1$ real eigenvalues $\lambda_1(M_j), \dots, \lambda_{n-1}(M_j)$, which for sake of concreteness we again arrange in nondecreasing order. In particular we have the well known *Cauchy interlacing inequalities* (see, e.g., [Wil1963, pp. 103–104])

$$\lambda_i(A) \leq \lambda_i(M_j) \leq \lambda_{i+1}(A) \tag{1}$$

for $i = 1, \dots, n-1$.

Received by the editors March 4, 2020.
2020 *Mathematics Subject Classification.* Primary 15A18; Secondary 15A42, 15B57.
The first author acknowledges the United States Department of Energy under Grant Contract desc0012704 and the Fermilab Neutrino Physics Center.
This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. FERMILAB-PUB-19-377-T.
The third author was supported by a Simons Investigator grant, the James and Carol Collins Chair, the Mathematical Analysis & Application Research Fund Endowment, and by NSF grant DMS-1764034.

©2021 American Mathematical Society

By the spectral theorem, we can find an orthonormal basis of eigenvectors $v_1, \dots, v_n$, where the v_i are in $\mathbb{C}_n$ of A associated to the eigenvalues $\lambda_1(A), \dots, \lambda_n(A)$ respectively. For any $i, j = 1, \dots, n$, let $v_{i,j}$ denote the jth component of v_i. This survey paper is devoted to the following elegant relation, which we will call the *eigenvector-eigenvalue identity*, relating this eigenvector component to the eigenvalues of A and M_j:

Theorem 1 (Eigenvector-eigenvalue identity)**.** *With the notation as above, we have*

$$|v_{i,j}|^2 \prod_{k=1;k\neq i}^{n} (\lambda_i(A) - \lambda_k(A)) = \prod_{k=1}^{n-1} (\lambda_i(A) - \lambda_k(M_j)) \,. \tag{2}$$

If one lets $p_A : \mathbb{C} \to \mathbb{C}$ denote the characteristic polynomial of A,

$$p_A(\lambda) \coloneqq \det(\lambda I_n - A) = \prod_{k=1}^{n} (\lambda - \lambda_k(A)), \tag{3}$$

where I_n denotes the $n \times n$ identity matrix, and similarly let $p_{M_j} : \mathbb{C} \to \mathbb{C}$ denote the characteristic polynomial of M_j,

$$p_{M_j}(\lambda) \coloneqq \det(\lambda I_{n-1} - M_j) = \prod_{k=1}^{n-1} (\lambda - \lambda_k(M_j)),$$

then the derivative $p'_A(\lambda_i(A))$ of p_A at $\lambda = \lambda_i(A)$ is equal to

$$p'_A(\lambda_i(A)) = \prod_{k=1;k\neq i}^{n} (\lambda_i(A) - \lambda_k(A)) \,,$$

and so (2) can be equivalently written in the characteristic polynomial form

$$|v_{i,j}|^2 p'_A(\lambda_i(A)) = p_{M_j}(\lambda_i(A)). \tag{4}$$

Example 2. If we set $n = 3$ and

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{pmatrix},$$

then the eigenvectors and eigenvalues are

$$v_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}; \quad \lambda_1(A) = 0$$

$$v_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}; \quad \lambda_2(A) = 3$$

$$v_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}; \quad \lambda_3(A) = 4$$

with minors M_j and eigenvalues $\lambda_i(M_j)$ given by

$$M_1 = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}; \qquad \lambda_{1,2}(M_1) = 2,\ 4$$

$$M_2 = \begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix}; \qquad \lambda_{1,2}(M_2) = 2 \mp \sqrt{2} \approx 0.59,\ 3.4$$

$$M_3 = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}; \qquad \lambda_{1,2}(M_3) = 2 \mp \sqrt{2} \approx 0.59,\ 3.4;$$

one can observe the interlacing inequalities (1). One can then verify (2) for all $i, j = 1, 2, 3$:

$$\frac{2}{3} = |v_{1,1}|^2 = \frac{(0-2)(0-4)}{(0-3)(0-4)},$$
$$\frac{1}{6} = |v_{1,2}|^2 = \frac{(0-2-\sqrt{2})(0-2+\sqrt{2})}{(0-3)(0-4)},$$
$$\frac{1}{6} = |v_{1,3}|^2 = \frac{(0-2-\sqrt{2})(0-2+\sqrt{2})}{(0-3)(0-4)},$$

$$\frac{1}{3} = |v_{2,1}|^2 = \frac{(3-2)(3-4)}{(3-0)(3-4)},$$
$$\frac{1}{3} = |v_{2,2}|^2 = \frac{(3-2-\sqrt{2})(3-2+\sqrt{2})}{(3-0)(3-4)},$$
$$\frac{1}{3} = |v_{2,3}|^2 = \frac{(3-2-\sqrt{2})(3-2+\sqrt{2})}{(3-0)(3-4)},$$

$$0 = |v_{3,1}|^2 = \frac{(4-2)(4-4)}{(4-0)(4-3)},$$
$$\frac{1}{2} = |v_{3,2}|^2 = \frac{(4-2-\sqrt{2})(4-2+\sqrt{2})}{(4-0)(4-3)},$$
$$\frac{1}{2} = |v_{3,3}|^2 = \frac{(4-2-\sqrt{2})(4-2+\sqrt{2})}{(4-0)(4-3)}.$$

One can also verify (4) for this example after computing

$$p'_A(\lambda) = 3\lambda^2 - 14\lambda + 12,$$
$$p_{M_1}(\lambda) = \lambda^2 - 6\lambda + 8,$$
$$p_{M_2}(\lambda) = \lambda^2 - 4\lambda + 2,$$
$$p_{M_3}(\lambda) = \lambda^2 - 4\lambda + 2.$$

Note, that $p'_A(\lambda) = p_{M_1}(\lambda) + p_{M_2}(\lambda) + p_{M_3}(\lambda)$, which is needed for the column normalization, see (x) in the consistency checks below.

Numerical code to verify the identity can be found at [Den2019].

Consistency Checks. Theorem 1 passes a number of basic consistency checks.

(i) *Dilation symmetry.* If one multiplies the matrix A by a real scalar c, then the eigenvalues of A and M_j also get multiplied by c, while the coefficients $v_{i,j}$ remain unchanged, which does not affect the truth of (2). To put it

another way, if one assigns units to the entries of A, then the eigenvalues of A, M_j acquire the same units, while $v_{i,j}$ remains dimensionless, and the identity (2) is dimensionally consistent.

(ii) *Translation symmetry.* If one adds a scalar multiple of the identity λI_n to A, then the eigenvalues of A and M_j are shifted by λ, while the coefficient $v_{i,j}$ remains unchanged. Thus both sides of (2) remain unaffected by such a transformation.

(iii) *Permutation symmetry.* Permuting the eigenvalues of A or M_j does not affect either side of (2) (provided one also permutes the index i accordingly). Permuting the ordering of the rows (and colums), as well as the index j, similarly has no effect on (2).

(iv) *First degenerate case.* If $v_{i,j}$ vanishes, then the eigenvector v_i for A also becomes an eigenvector for M_j with the same eigenvalue $\lambda_i(A)$ after deleting the jth coefficient. In this case, both sides of (2) vanish.

(v) *Second degenerate case.* If the eigenvalue $\lambda_i(A)$ of A occurs with multiplicity greater than one, then by the interlacing inequalities (1) it also occurs as an eigenvalue of M_j. Again in this case, both sides of (2) vanish.

(vi) *Compatibility with interlacing.* More generally, the identity (2) is consistent with the interlacing (1) because the component $v_{i,j}$ of the unit eigenvector v_i has magnitude at most 1.

(vii) *Phase symmetry.* One has the freedom to multiply each eigenvector v_i by an arbitrary complex phase $e^{\sqrt{-1}\theta_i}$ without affecting the matrix A or its minors M_j. But both sides of (2) remain unchanged when one does so.

(viii) *Diagonal case.* If A is a diagonal matrix with diagonal entries $\lambda_1(A), \dots, \lambda_n(A)$, then $|v_{i,j}|$ equals 1 when $i = j$ and zero otherwise, while the eigenvalues of M_j are formed from those of A by deleting one copy of $\lambda_i(A)$. In this case one can easily verify (2) by hand.

(ix) *Row normalization.* As the eigenvectors $v_1, \dots, v_n$ form the columns of an orthogonal matrix, one must have the identity $\sum_{i=1}^n |v_{i,j}|^2 = 1$ for all $j = 1, \dots, n$. Assuming for simplicity that the eigenvalues $\lambda_i(A)$ are distinct, this follows easily from the algebraic identity $\sum_{i=1}^n \frac{\lambda_i^m}{\prod_{k=1;k\neq i}^n (\lambda_i - \lambda_k)} = \delta_{m,n-1}$ for $m = 0, \dots, n-1$ and any distinct complex numbers $\lambda_1, \dots, \lambda_n$, which can be seen by integrating the rational function $\frac{z^m}{\prod_{k=1}^n (z-\lambda_k)}$ along a large circle $\{|z| = R\}$ and applying the residue theorem. See also Remark 9 below.

(x) *Column normalization.* As the eigenvectors $v_1, \dots, v_n$ are unit vectors, one must have $\sum_{j=1}^n |v_{i,j}|^2 = 1$ for all $i = 1, \dots, n$. To verify this, use the translation symmetry (ii) to normalize $\lambda_i(A) = 0$, and then observe (e.g., from (7)) that $(-1)^n \sum_{j=1}^n p_{M_j}(0) = \sum_{j=1}^n \det(M_j) = \operatorname{tr} \operatorname{adj}(A)$ is the $(n-1)$-th elementary symmetric function of the eigenvalues and thus equal (since $\lambda_i(A)$ vanishes) to $\prod_{k=1;k\neq i}^n \lambda_k(A) = (-1)^n p'_A(0)$. Comparing this with (4), we obtain $\sum_{j=1}^n |v_{i,j}|^2 = 1$ as desired. Alternatively, one can see from Jacobi's formula $\frac{d}{dt} \det(A(t)) = \operatorname{tr}(\operatorname{adj}(A(t)) \frac{dA(t)}{dt})$ that $p'_A(\lambda) = \operatorname{tr}(\operatorname{adj}(\lambda I_n - A)) = \sum_j p_{M_j}(\lambda)$ which when combined with (4) also recovers the identity $\sum_{j=1}^n |v_{i,j}|^2 = 1$. Jacobi's formula give us the need relationships between the eigenvalues of A and the eigenvalues of the M_j's. They are $(n-k)S_k(A) = \sum_{j=1}^n S_k(M_j)$ for $k = 1, \dots, n-1$, where $S_k(A)$ is the kth

elementary symmetric polynomial of the eigenvalues of A, e.g., $S_1(A) = \mathrm{tr}(A), \ldots, S_n(A) = \det(A)$.

(xi) *Relative phase information.* As mentioned in (vii) above, the phase of any individual eigenvector v_i is arbitrary, therefore the relative phase between $v_{i,k}$ and $v_{j,k}$, $i \neq j$, is arbitrary. However, the relative phases between the components of any v_i, say between $v_{i,j}$ and $v_{i,k}$ for $j \neq k$, is not arbitrary. Identity (2) can be used to extract these relative phases as follows: consider a unitary transformation on the matrix A and its eigenvectors such that $v_{i,j} \to \frac{1}{\sqrt{2}}(v_{i,j} + \omega\, v_{i,k})$ and $v_{i,k} \to \frac{1}{\sqrt{2}}(v_{i,k} - \omega^*\, v_{i,j})$ with $\omega = 1$ or $\sqrt{-1}$ where ω^* is the complex conjugate of ω. Applying (2) to the original A and to the two unitary transformed A's, gives us the information need to extract $\arg(v_{i,j} v_{i,k}^*)$. Note $p_{M_j}(\lambda)$ and $p_{M_k}(\lambda)$ are not invariant under this particular unitary transformation, but $p'_A(\lambda)$ and other $p_{M_l}(\lambda)$, $l \neq j$ or k, are invariant. Furthermore, the unitarity condition that the $\sum_{i=1}^n v_{i,j} v_{i,k}^* = 0$ for $j \neq k$, can also be derived in this fashion. For further discussion on the relative phases, see the beginning of Section 4.

We also note that, since (2) is a continuous function of the matrix A, it is possible to treat all eigenvalues as simple via the usual limiting argument.

The eigenvector-eigenvalue identity has a surprisingly complicated history in the literature, having appeared in some form or another (albeit often in a lightly disguised form) in over two dozen references, and being independently rediscovered a half-dozen times, in fields as diverse as numerical linear algebra, random matrix theory, inverse eigenvalue problems, graph theory (including chemical graph theory, graph reconstruction, and walks on graphs), and neutrino physics; see Figure 1. While the identity applies to all Hermitian matrices, and extends in fact to normal matrices and more generally to diagonalizable matrices, it has found particular application in the special case of symmetric tridiagonal matrices (such as Jacobi matrices), which are of particular importance in several fundamental algorithms in numerical linear algebra.

While the eigenvector-eigenvalue identity is moderately familiar to some mathematical communities, it is not as broadly well known as other standard identities in linear algebra, such as Cramer's rule [Cra1750] or the Cauchy determinant formula [Cau1841] (though, as we shall shortly see, it can be readily derived from either of these identities). While several of the papers in which some form of the identity was discovered went on to be cited several times by subsequent work, the citation graph is only very weakly connected; in particular, Figure 1 reveals that many of the citations coming from the earliest work on the identity did not propagate to later works, which instead were based on independent rediscoveries of the identity (or one of its variants). In many cases, the identity was not highlighted as a relation between eigenvectors and eigenvalues, but was instead introduced in passing as a tool to establish some other application; also, the form of the identity and the notation used varied widely from appearance to appearance, making it difficult to search for occurrences of the identity by standard search engines. The situation changed after a popular science article by Wolchover [Wol2019] reported on the most recent rediscovery [DPZ2020, DPTZ2019] of the identity by ourselves. In the wake of the publicity generated by that article, we received many notifications (see Section 5) of the disparate places in the literature where the eigenvector-eigenvalue

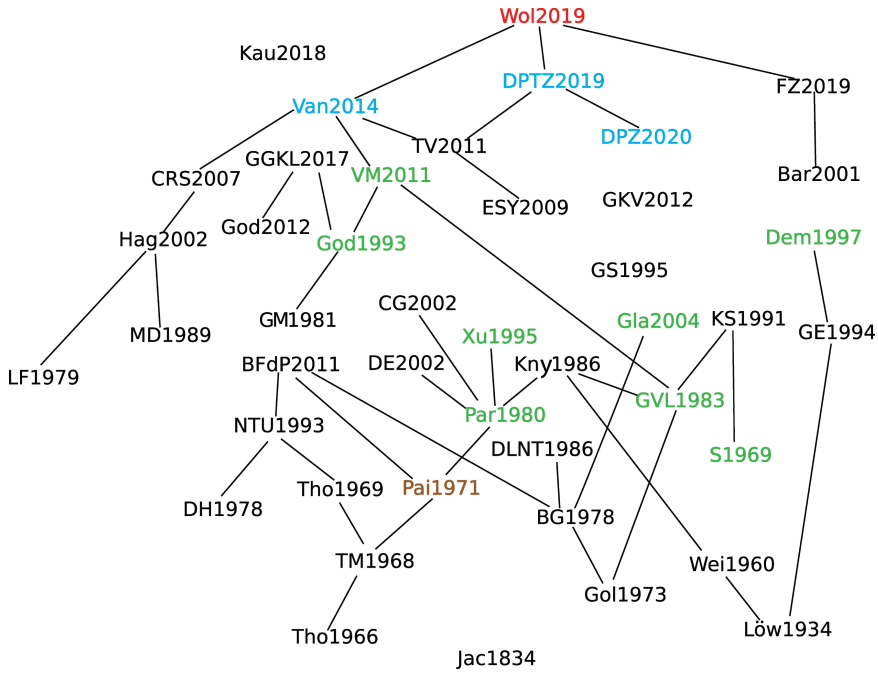

FIGURE 1. The citation graph of all the references in the literature we are aware of (predating the current survey) that mention some variant of the eigenvector-eigenvalue identity. To reduce clutter, transitive references (e.g., a citation of a paper already cited by another paper in the bibliography) are omitted. Note the very weakly connected nature of the graph, with many early initial references not being (transitively) cited by many of the more recent references. Blue references are preprints, green references are books, the brown reference is a thesis, and the red reference is a popular science article. This graph was mostly crowdsourced from feedback received by the authors after the publication of [Wol2019]. The reference [Jac1834] predates all others found by a century!

identity, or an identity closely related to it, was discovered. Effectively, this "crowdsourced" the task of collating all these references together. In this paper, we survey all the appearances of the eigenvector-eigenvalue identity that we are aware of as a consequence of these efforts, as well as provide several proofs, generalizations, and applications of the identity. Finally, we speculate on some reasons for the limited nature of the dissemination of this identity in prior literature.

2. Proofs of the identity

The identity (2) can be readily established from existing standard identities in the linear algebra literature. We now give several such proofs.

2.1. **The adjugate proof.** We first give a proof using adjugate matrices, which is a purely "polynomial" proof that avoids any invertibility, division, or nondegeneracy hypotheses in the argument. In particular, as we remark below, it has an extension

to (diagonalizable) matrices that take values in arbitrary commutative rings. This argument appears for instance in [Par1980, Section 7.9].

Recall that if A is an $n \times n$ matrix, the *adjugate matrix* $\operatorname{adj}(A)$ is given by the formula

$$\operatorname{adj}(A) := \big((-1)^{i+j}\det(M_{ji})\big)_{1\leq i,j\leq n}, \tag{5}$$

where M_{ji} is the $n-1\times n-1$ matrix formed by deleting the jth row and ith column from A. From Cramer's rule we have the identity

$$\operatorname{adj}(A)A = A\operatorname{adj}(A) = \det(A)I_n.$$

If A is a diagonal matrix with (complex) entries $\lambda_1(A), \dots, \lambda_n(A)$, then $\operatorname{adj}(A)$ is also a diagonal matrix with ii entry $\prod_{k=1;k\neq i}^n \lambda_k(A)$. More generally, if A is a normal matrix with diagonalization

$$A = \sum_{i=1}^n \lambda_i(A) v_i v_i^*, \tag{6}$$

where $v_1, \dots, v_n$ are an orthonormal basis of eigenvectors of A and v_i^* is the conjugate transpose of v_i, then $\operatorname{adj}(A)$ has the same basis of eigenvectors with diagonalization

$$\operatorname{adj}(A) = \sum_{i=1}^n \Big(\prod_{k=1;k\neq i}^n \lambda_k(A)\Big) v_i v_i^*. \tag{7}$$

If one replaces A by $\lambda I_n - A$ for any complex number λ, we therefore have[1]

$$\operatorname{adj}(\lambda I_n - A) = \sum_{i=1}^n \Big(\prod_{k=1;k\neq i}^n (\lambda - \lambda_k(A))\Big) v_i v_i^*.$$

If one specializes to the case $\lambda = \lambda_i(A)$ for some $i = 1, \dots, n$, then all but one of the summands on the right-hand side vanish, and the adjugate matrix becomes a scalar multiple of the rank one projection $v_i v_i^*$,

$$\operatorname{adj}(\lambda_i(A) I_n - A) = \Big(\prod_{k=1;k\neq i}^n (\lambda_i(A) - \lambda_k(A))\Big) v_i v_i^*. \tag{8}$$

Extracting out the jj component of this identity using (5), we conclude that

$$\det(\lambda_i(A) I_{n-1} - M_j) = \Big(\prod_{k=1;k\neq i}^n (\lambda_i(A) - \lambda_k(A))\Big) |v_{i,j}|^2, \tag{9}$$

which is equivalent to (2). In fact this shows that the eigenvector-eigenvalue identity holds for normal matrices A as well as Hermitian matrices (despite the fact that the minor M_j need not be Hermitian or normal in this case). Of course in this case the eigenvalues are not necessarily real and thus cannot be arranged in increasing order, but the order of the eigenvalues plays no role in the identity (2).

Remark 3. The same argument also yields an off-diagonal variant

$$(-1)^{j+j'}\det(\lambda_i(A)(I_n)_{j'j} - M_{j'j}) = \Big(\prod_{k=1;k\neq i}^n (\lambda_i(A) - \lambda_k(A))\Big) v_{i,j}\overline{v_{i,j'}} \tag{10}$$

[1] To our knowledge, this identity first appears in [Hal1942, p. 157].

for any $1 \leq j, j' \leq n$, where $(I_n)_{j'j}$ is the $n-1 \times n-1$ minor of the identity matrix I_n. When $j = j'$, this minor $(I_n)_{j'j}$ is simply equal to I_{n-1} and the determinant can be expressed in terms of the eigenvalues of the minor M_j; however when $j \neq j'$ there is no obvious way to express the left-hand side of (10) in terms of eigenvalues of $M_{j'j}$ (though one can still of course write the determinant as the product of the eigenvalues of $\lambda_i(A)(I_n)_{j'j} - M_{j'j}$). Another way of viewing (10) is that for every $1 \leq j' \leq n$, the vector with jth entry

$$(-1)^j \det(\lambda_i(A)(I_n)_{j'j} - M_{j'j})$$

is a nonnormalized eigenvector associated to the eigenvalue $\lambda_i(A)$; this observation appears for instance in [Gan1959, pp. 85–86]. See [Van2014] for some further identities relating the components $v_{i,j}$ of the eigenvector v_i to various determinants.

Remark 4. This remark is due to Vassilis Papanicolaou.[2] The above argument also applies to nonnormal matrices A, so long as they are diagonalizable with some eigenvalues $\lambda_1(A), \dots, \lambda_n(A)$ (not necessarily real or distinct). Indeed, if we let $v_1, \dots, v_n$ be a basis of right eigenvectors of A (so that $Av_i = \lambda_i(A)v_i$ for all $i = 1, \dots, n$), and let $w_1, \dots, w_n$ be the corresponding dual basis[3] of left eigenvectors (so $w_i^T A = w_i^T \lambda_i(A)$, and $w_i^T v_j$ is equal to 1 when $i = j$ and 0 otherwise), then we have the diagonalization

$$A = \sum_{i=1}^n \lambda_i(A) v_i w_i^T,$$

and one can generalize (8) to

$$\operatorname{adj}(\lambda_i(A)I_n - A) = \Bigg(\prod_{k=1;k\neq i}^n (\lambda_i(A) - \lambda_k(A)) \Bigg) v_i w_i^T \tag{11}$$

leading to an extension

$$\det(\lambda_i(A)I_{n-1} - M_j) = \Bigg(\prod_{k=1;k\neq i}^n (\lambda_i(A) - \lambda_k(A)) \Bigg) v_{i,j} w_{i,j} \tag{12}$$

of (9) to arbitrary diagonalizable matrices. We remark that this argument shows that the identity (12) is in fact valid for any diagonalizable matrix taking values in any commutative ring (not just the complex numbers). The identity (10) may be generalized in a similar fashion; we leave the details to the interested reader.

Remark 5. As pointed out to us by Darij Grinberg,[4] the identity (4) may be generalized to the nondiagonalizable setting. Namely, one can prove that

$$v_j w_j p'_A(\lambda) = (w^T v) p_{M_j}(\lambda)$$

for an arbitrary $n \times n$ matrix A (with entries in an arbitrary commutative ring), any $j = 1, \dots, n$ (with M_j the $n-1 \times n-1$ minor formed from A by removing the jth row and column, and any right-eigenvector $Av = \lambda v$ and left-eigenvector $w^T A = w^T \lambda$ with a common eigenvalue λ. After reducing to the case $\lambda = 0$, the main step in the proof is to establish two variants of (8), namely that $w_j(\operatorname{adj} A)_{i,k} = w_k(\operatorname{adj} A)_{i,j}$

[2] `terrytao.wordpress.com/2019/08/13/eigenvectors-from-eigenvalues/#comment-519905`

[3] In the case when A is a normal matrix and the v_i are unit eigenvectors, the dual eigenvector w_i would be the complex conjugate of v_i.

[4] `terrytao.wordpress.com/2019/12/03`

and $(\mathrm{adj}A)_{k,i}v_j = (\mathrm{adj}A)_{j,i}v_k$ for all $i,j,k = 1,\dots,n$. We refer the reader to the comment of Grinberg for further details.

Remark 6. As observed in [Van2014, Appendix A], one can obtain an equivalent identity to (2) by working with the powers $A^m, m = 0,\dots,n-1$ in place of $\mathrm{adj}(A)$. Indeed, from (6) we have

$$A^m = \sum_{i=1}^n \lambda_i(A)^m v_i v_i^*,$$

and hence on extracting the jj component

$$(A^m)_{jj} = \sum_{i=1}^n \lambda_i(A)^m |v_{i,j}|^2$$

for all $j = 1,\dots,n$ and $m = 0,\dots,n-1$. Using Vandermonde determinants (and assuming for sake of argument that the eigenvalues $\lambda_i(A)$ are distinct), one can then solve for the $|v_{i,j}|^2$ in terms of the $(A^m)_{jj}$, eventually reaching an identity [Van2014, Theorem 2] equivalent to (2) (or (10)), which in the case when A is the adjacency matrix of a graph can also be expressed in terms of counts of walks of various lengths between pairs of vertices. We refer the reader to [Van2014] for further details.

2.2. **The Cramer's rule proof.** Returning now to the case of Hermitian matrices, we give a variant of the above proof of (2) that still relies primarily on Cramer's rule, but makes no explicit mention of the adjugate matrix. As discussed in the next section, variants of this argument have appeared multiple times in the literature. We first observe that to prove (2) for Hermitian matrices A, it suffices to do so under the additional hypothesis that A has simple spectrum (all eigenvalues occur with multiplicity one), or equivalently that

$$\lambda_1(A) < \lambda_2(A) < \dots < \lambda_n(A).$$

This is because any Hermitian matrix with repeated eigenvalues can be approximated to arbitrary accuracy by a Hermitian matrix with simple spectrum, and both sides of (2) vary continuously with A (at least if we avoid the case when $\lambda_i(A)$ occurs with multiplicity greater than one, which is easy to handle anyway by the second degenerate case (iv) noted in the introduction).

As before, we diagonalize A in the form (6). For any complex parameter λ not equal to one of the eigenvalues $\lambda_i(A)$, the resolvent $(\lambda I_n - A)^{-1}$ can then also be diagonalized as

$$(\lambda I_n - A)^{-1} = \sum_{i=1}^n \frac{v_i v_i^*}{\lambda - \lambda_i(A)}. \tag{13}$$

Extracting out the jj component of this matrix identity using Cramer's rule [Cra1750], we conclude that

$$\frac{\det(\lambda I_{n-1} - M_j)}{\det(\lambda I_n - A)} = \sum_{i=1}^n \frac{|v_{i,j}|^2}{\lambda - \lambda_i(A)},$$

which we can express in terms of eigenvalues as

$$\frac{\prod_{k=1}^{n-1}(\lambda - \lambda_k(M_j))}{\prod_{k=1}^{n}(\lambda - \lambda_k(A))} = \sum_{i=1}^n \frac{|v_{i,j}|^2}{\lambda - \lambda_i(A)}. \tag{14}$$

Both sides of this identity are rational functions in λ, and have a pole at $\lambda = \lambda_i(A)$ for any given $i = 1, \dots, n$. Extracting the residue at this pole, we conclude that

$$\frac{\prod_{k=1}^{n-1}(\lambda_i(A) - \lambda_k(M_j))}{\prod_{k=1:k\neq i}^{n}(\lambda_i(A) - \lambda_k(A))} = |v_{i,j}|^2,$$

which rearranges to give (2).

Remark 7. One can view the above derivation of (2) from (14) as a special case of the partial fractions decomposition

$$\frac{P(t)}{Q(t)} = \sum_{Q(\alpha)=0} \frac{P(\alpha)}{Q'(\alpha)} \frac{1}{t-\alpha},$$

whenever Q is a polynomial with distinct roots $\alpha_1, \dots, \alpha_n$ and P is a polynomial of degree less than that of Q. Equivalently, this derivation can be viewed as a special case of the Lagrange interpolation formula (see, e.g., [AS1964, §25.2])

$$P(t) = \sum_{i=1}^{n} P(\alpha_i) \prod_{j=1;j\neq i}^{n} \frac{t-\alpha_j}{\alpha_i - \alpha_j},$$

whenever $\alpha_1, \dots, \alpha_n$ are distinct and P is a polynomial of degree less than n.

Remark 8. A slight variant of this proof was observed by Aram Harrow,[5] inspired by the inverse power method for approximately computing eigenvectors numerically. We again assume simple spectrum. Using the translation invariance noted in consistency check (ii) of the introduction, we may assume without loss of generality that $\lambda_i(A) = 0$. Applying the resolvent identity (13) with λ equal to a small nonzero quantity ε, we conclude that

$$(\varepsilon I_n - A)^{-1} = \frac{v_i v_i^*}{\varepsilon} + O(1). \tag{15}$$

On the other hand, by Cramer's rule, the jj component of the left-hand side is

$$\frac{\det(\varepsilon I_{n-1} - M_j)}{\det(\varepsilon I_n - A)} = \frac{p_{M_j}(\varepsilon)}{p_A(\varepsilon)} = \frac{p_{M_j}(0) + O(\varepsilon)}{\varepsilon p'_A(\varepsilon) + O(\varepsilon^2)}.$$

Extracting out the top order terms in ε, one obtains (4) and hence (2). A variant of this argument also gives the more general identity

$$\sum_{i:\lambda_i(A)=\lambda_*} |v_{i,j}|^2 = \lim_{\lambda\to\lambda_*} \frac{p_{M_j}(\lambda)(\lambda - \lambda_*)}{p_A(\lambda)}, \tag{16}$$

whenever λ_* is an eigenvalue of A of some multiplicity $m \geq 1$. Note when $m = 1$ we can recover (4) thanks to L'Hôpital's rule. The right-hand side of (16) can also be interpreted as the residue of the rational function p_{M_j}/p_A at λ_*.

An alternate way to arrive at (2) from (14) is as follows. Assume for the sake of this argument that the eigenvalues of M_j are all distinct from the eigenvalues of A. Then we can substitute $\lambda = \lambda_k(M_j)$ in (14) and conclude that

$$\sum_{i=1}^{n} \frac{|v_{i,j}|^2}{\lambda_k(M_j) - \lambda_i(A)} = 0 \tag{17}$$

[5] `twitter.com/quantum_aram/status/1195185551667847170`

for $k = 1, \dots, n-1$. Also, since the v_i form an orthonormal basis, we have from expressing the unit vector e_j in this basis that

$$\sum_{i=1}^{n} |v_{i,j}|^2 = 1. \tag{18}$$

This is a system of n linear equations in n unknowns $|v_{i,j}|^2$. For sake of notation let us use permutation symmetry to set $i = n$. From a further application of Cramer's rule, one can then write

$$|v_{n,j}|^2 = \frac{\det(S')}{\det(S)},$$

where S is the $n \times n$ matrix with ki entry equal to $\frac{1}{\lambda_k(M_j) - \lambda_i(A)}$ when $k = 1, \dots, n-1$, and equal to 1 when $k = n$, and S' is the minor of S formed by removing the nth row and column. Using the well-known Cauchy determinant identity [Cau1841]

$$\det\left(\frac{1}{x_i - y_j}\right)_{1 \le i,j \le n} = \frac{\prod_{1 \le j < i \le n}(x_i - x_j)(y_i - y_j)}{\prod_{i=1}^{n}\prod_{j=1}^{n}(x_i - y_j)} \tag{19}$$

and inspecting the asymptotics as $x_n \to \infty$, we soon arrive at the identities

$$\det(S) = \frac{\prod_{1 \le l < k \le n-1}(\lambda_k(M_j) - \lambda_l(M_j)) \prod_{1 \le l < k \le n}(\lambda_k(A) - \lambda_l(A))}{\prod_{l=1}^{n-1}\prod_{k=1}^{n}(\lambda_l(M_j) - \lambda_k(A))}$$

and

$$\det(S') = \frac{\prod_{1 \le l < k \le n-1}(\lambda_k(M_j) - \lambda_l(M_j))(\lambda_k(A) - \lambda_l(A))}{\prod_{l=1}^{n-1}\prod_{k=1}^{n-1}(\lambda_l(M_j) - \lambda_k(A))},$$

and the identity (2) then follows after a brief calculation.

Remark 9. The derivation of the eigenvector-eigenvalue identity (2) from (17), as well as the obvious normalization (18), is reversible. Indeed, the identity (2) implies that the rational functions on both sides of (14) have the same residues at each of their (simple) poles, and these functions decay to zero at infinity, hence they must agree by Liouville's theorem. Specializing (14) to $\lambda = \lambda_k(M_j)$ then recovers (17), while comparing the leading asymptotics of both sides of (14) as $\lambda \to \infty$ recovers (18) (note this also establishes the consistency check (ix) from the introduction). As the identity (17) involves the same quantities $v_{i,j}, \lambda_k(M_j), \lambda_i(A)$ as (2), one can thus view (17) as an equivalent formulation of the eigenvector-eigenvalue identity, at least in the case when all the eigenvalues of A are distinct. The identity (14) (viewing λ as a free parameter) can also be interpreted in this fashion.

Remark 10. The above resolvent-based arguments have a good chance of being extended to certain classes of infinite matrices (e.g., Jacobi matrices), or other Hermitian operators, particularly if they have good spectral properties (e.g., they are trace class). Certainly it is well known that spectral projections of an operator to a single eigenvalue λ can often be viewed as residues of the resolvent at that eigenvalue, in the spirit of (15), under various spectral hypotheses of the operator in question. The main difficulty is to find a suitable extension of Cramer's rule to infinite-dimensional settings, which would presumably require some sort of regularized determinant such as the Fredholm determinant. We will not explore this question further here, however, as pointed out to us by Carlos Tomei (personal communication), for reasonable Hermitian infinite matrices A such as Jacobi matrices, one can formulate an identity similar to (14) for the upper left coefficient

$\langle e_1, (\lambda - A)e_1\rangle$ of the resolvent for λ in the upper half-plane, which can for instance be used (in conjunction with the Herglotz representation theorem) to recover the spectral theorem for such matrices.

2.3. **Coordinate-free proof.** We now give a proof that largely avoids the use of coordinates or matrices, essentially due to Bo Berndtsson.[6] For this proof we assume familiarity with exterior algebra (see, e.g., [BM1941, Chapter XVI]). The key identity is the following statement.

Lemma 11 (Coordinate-free eigenvector-eigenvalue identity)**.** *Let $T : \mathbb{C}^n \to \mathbb{C}^n$ be a self-adjoint linear map that annihilates a unit vector v. For each unit vector $f \in \mathbb{C}^n$, let $\Delta_T(f)$ be the determinant of the quadratic form $w \mapsto (Tw, w)_{\mathbb{C}^n}$ restricted to the orthogonal complement $f^\perp := \{w \in \mathbb{C}^n : (f, w)_{\mathbb{C}^n} = 0\}$, where $(,)_{\mathbb{C}^n}$ denotes the Hermitian inner product on $\mathbb{C}^n$. Then one has*

$$|(v, f)_{\mathbb{C}^n}|^2 \Delta_T(v) = \Delta_T(f) \tag{20}$$

for all unit vectors $f \in \mathbb{C}^n$.

Proof. The determinant of a quadratic form $w \mapsto (Tw, w)_{\mathbb{C}^n}$ on a k-dimensional subspace V of $\mathbb{C}^n$ can be expressed as $(T\alpha, \alpha)_{\bigwedge^k \mathbb{C}^n}/(\alpha, \alpha)_{\bigwedge^k \mathbb{C}^n}$ for any nondegenerate element α of the kth exterior power $\bigwedge^k V \subset \bigwedge^k \mathbb{C}^n$ (equipped with the usual Hermitian inner product $(,)_{\bigwedge^k \mathbb{C}^n}$), where the operator T is extended to $\bigwedge^k \mathbb{C}^n$ in the usual fashion. If $f \in \mathbb{C}^n$ is a unit vector, then the Hodge dual $*f \in \bigwedge^{n-1} \mathbb{C}^n$ is a unit vector in $\bigwedge^{n-1}(f^\perp)$, so that we have the identity

$$\Delta_T(f) = (T(*f), *f)_{\bigwedge^{n-1} \mathbb{C}^n}. \tag{21}$$

To prove (20), it thus suffices to establish the more general identity

$$(f, v)_{\mathbb{C}^n} \Delta_T(v)(v, g)_{\mathbb{C}^n} = (T(*f), (*g))_{\bigwedge^{n-1} \mathbb{C}^n} \tag{22}$$

for all $f, g \in \mathbb{C}^n$. If f is orthogonal to v, then $*f$ can be expressed as a wedge product of v with an element of $\bigwedge^{n-2} \mathbb{C}^n$, and hence $T(*f)$ vanishes, so that (22) holds in this case. If g is orthogonal to v, then we again obtain (22) thanks to the self-adjoint nature of T. Finally, when $f = g = v$, the claim follows from (21). Since the identity (22) is sesquilinear in f, g, the claim follows. □

Now we can prove (2). Using translation symmetry we may normalize $\lambda_i(A) = 0$. We apply Lemma 11 to the self-adjoint map $T : w \mapsto Aw$, setting v to be the null vector $v = v_i$ and f to be the standard basis vector e_j. Working in the orthonormal eigenvector basis $v_1, \dots, v_n$, we have $\Delta(v) = \prod_{k=1;k\neq i}^n \lambda_k(A)$; working in the standard basis $e_1, \dots, e_n$, we have $\Delta_T(f) = \det(M_j) = \prod_{k=1}^{n-1} \lambda_k(M_j)$. Finally we have $(v, f)_{\mathbb{C}^n} = v_{i,j}$. The claim follows.

Remark 12. In coordinates, identity (21) may be rewritten as $\Delta_T(f) = f^* \mathrm{adj}(A) f$. Thus we see that Lemma 11 is basically (8) in disguise. It seems likely that the variant identity in Remark 5 can also be established in a similar coordinate-free fashion.

[6] `terrytao.wordpress.com/2019/08/13/eigenvectors-from-eigenvalues/#comment-519914`

2.4. **Proof using perturbative analysis.** Now we give a proof using perturbation theory, which to our knowledge first appears in [MD1989]. By the usual limiting argument we may assume that A has simple eigenvalues. Let ε be a small parameter, and consider the rank one perturbation $A + \varepsilon e_j e_j^*$ of A, where $e_1, \dots, e_n$ is the standard basis. From (3) and cofactor expansion, the characteristic polynomial $p_{A+\varepsilon e_j e_j^*}(\lambda)$ of this perturbation may be expanded as

$$p_{A+\varepsilon e_j e_j^*}(\lambda) = p_A(\lambda) - \varepsilon p_{M_j}(\lambda) + O(\varepsilon^2).$$

On the other hand, from perturbation theory the eigenvalue $\lambda_i(A + \varepsilon e_j e_j^*)$ may be expanded as

$$\lambda_i(A + \varepsilon e_j e_j^*) = \lambda_i(A) + \varepsilon |v_{i,j}|^2 + O(\varepsilon^2).$$

If we then Taylor-expand the identity

$$p_{A+\varepsilon e_j e_j^*}(\lambda_i(A + \varepsilon e_j e_j^*)) = 0$$

and extract the terms that are linear in ε, we conclude that

$$\varepsilon |v_{i,j}|^2 p_A'(\lambda_i(A)) - \varepsilon p_{M_j}(\lambda_i(A)) = 0,$$

which gives (4) and hence (2).

2.5. **Proof using a Cauchy–Binet type formula.** Now we give a proof based on a Cauchy–Binet type formula, which is also related to Lemma 11. This argument first appeared in [DPTZ2019].

Lemma 13 (Cauchy–Binet type formula)**.** *Let A be an $n \times n$ Hermitian matrix with a zero eigenvalue $\lambda_i(A) = 0$. Then for any $n \times n-1$ matrix B, one has*

$$\det(B^* A B) = (-1)^{n-1} p_A'(0) \left|\det \begin{pmatrix} B & v_i \end{pmatrix}\right|^2,$$

where $\begin{pmatrix} B & v_i \end{pmatrix}$ denotes the $n \times n$ matrix with right column v_i and all remaining columns given by B.

Proof. We use a perturbative argument related to that in Section 2.4. Since $Av_i = 0$, $v_i^* A = 0$, and $v_i^* v_i = 1$, we easily confirm the identity

$$\begin{pmatrix} B^* \\ v_i^* \end{pmatrix} (\varepsilon I_n - A) \begin{pmatrix} B & v_i \end{pmatrix} = \begin{pmatrix} -B^* A B + O(\varepsilon) & O(\varepsilon) \\ O(\varepsilon) & \varepsilon \end{pmatrix}$$

for any parameter ε, where the matrix on the right-hand side is given in block form, with the top left block being an $n-1 \times n-1$ matrix and the bottom right entry being a scalar. Taking determinants of both sides, we conclude that

$$p_A(\varepsilon) \left|\det \begin{pmatrix} B & v_i \end{pmatrix}\right|^2 = (-1)^{n-1} \det(B^* A B^*) \varepsilon + O(\varepsilon^2).$$

Extracting out the ε coefficient of both sides, we obtain the claim. □

Remark 14. In the case when v_i is the basis vector e_n, we may write A in block form as $A = \begin{pmatrix} M_n & 0_{n-1\times 1} \\ 0_{1\times n-1} & 0 \end{pmatrix}$, where $0_{i\times j}$ denotes the $i \times j$ zero matrix, and write $B = \begin{pmatrix} B' \\ x^* \end{pmatrix}$ for some $n-1 \times n-1$ matrix B' and $(n-1)$-dimensional vector x, in which case one can calculate

$$\det(B^* A B) = \det((B')^* M_n B') = \det(M_n) |\det(B')|^2$$

and

$$\det \begin{pmatrix} B & v_i \end{pmatrix} = \det(B').$$

Since $p'_A(0) = (-1)^{n-1}\det(M_n)$ in this case, this establishes (13) in the case $v_i = e_n$. The general case can then be established from this by replacing A by UAU^* and B by UB, where U is any unitary matrix that maps v_i to e_n.

We now prove (4) and hence (2). Using the permutation and translation symmetries, we may normalize $\lambda_i(A) = 0$ and $j = 1$. We then apply Lemma 13 with $B = \begin{pmatrix} 0_{1\times n-1} \\ I_{n-1} \end{pmatrix}$, in which case

$$\det(B^*AB) = \det(M_1) = (-1)^{n-1} p_{M_1}(0)$$

and

$$\det\begin{pmatrix} B & v_i \end{pmatrix} = v_{i,1}.$$

Applying Lemma 13, we obtain (4).

2.6. **Proof using an alternate expression for eigenvector component magnitudes.** There is an alternate formula for the square $|v_{i,j}|^2$ of the eigenvector component $v_{i,j}$ that was introduced in the paper [ESY2009, (5.8)] of Erdős, Schlein, and Yau in the context of random matrix theory, and then highlighted further in a paper [TV2011, Lemma 41] of the third author and Vu; it differs from the eigenvector-eigenvalue formula in that it involves the actual coefficients of A and M_1, rather than just their eigenvalues. It was also previously discovered by Gaveau and Schulman [GS1995, (2.6)]. For sake of notation we just give the formula in the $j = 1$ case.

Lemma 15 (Alternate expression for $v_{i,1}$)**.** *Let A be an $n \times n$ Hermitian matrix written in block matrix form as*

$$A = \begin{pmatrix} a_{11} & X^* \\ X & M_1 \end{pmatrix},$$

where X is an $(n-1)$-dimensional column vector and a_{11} is a scalar. Let $i = 1, \dots, n$, and suppose that $\lambda_i(A)$ is not an eigenvalue of M_1. Let $u_1, \dots, u_{n-1}$ denote an orthonormal basis of eigenvectors of M_1, associated to the eigenvalues $\lambda_1(M_1), \dots, \lambda_{n-1}(M_1)$. Then

$$|v_{i,1}|^2 = \frac{1}{1 + \sum_{j=1}^{n-1} X^*(\lambda_i(A) I_{n-1} - M_1)^{-2} X}. \tag{23}$$

This lemma is useful in random matrix theory for proving *delocalization* of eigenvectors of random matrices, which roughly speaking amounts to proving upper bounds on the quantity $\sup_{1\le j\le n} |v_{i,j}|$.

Proof. One can verify that this result enjoys the same translation symmetry as Theorem 1 (see consistency check (ii) from the introduction), so without loss of generality we may normalize $\lambda_i(A) = 0$. If we write $v_i = \begin{pmatrix} v_{i,1} \\ w_i \end{pmatrix}$ for an $(n-1)$-dimensional column vector w_i, then the eigenvector equation $Av_i = \lambda_i(A)v_i = 0$ can be written as the system

$$\begin{aligned} a_{11} v_{i,1} + X^* w_i &= 0, \\ X v_{i,1} + M_1 w_i &= 0. \end{aligned}$$

By hypothesis, 0 is not an eigenvalue of M_1, so we may invert M_1 and conclude that

$$w_i = -M_1^{-1} X v_{i,1}.$$

Since v_i is a unit vector, we have $|w_i|^2 + |v_{i,1}|^2 = 1$. Combining these two formulae and using some algebra, we obtain the claim. □

Now we can give an alternate proof of (4) and hence (2). By permutation symmetry (iii) it suffices to establish the $j = 1$ case. Using limiting arguments as before, we may assume that A has distinct eigenvalues; by further limiting arguments we may also assume that the eigenvalues of M_1 are distinct from those of A. By translation symmetry (ii) we may normalize $\lambda_i(A) = 0$. Comparing (4) with (23), our task reduces to establishing the identity

$$p'_A(0) = p_{M_1}(0)(1 + X^* M_1^{-2} X).$$

However, for any complex number λ not equal to an eigenvalue of M_1, we may apply Schur complementation [Cot1974] to the matrix

$$\lambda I_n - A = \begin{pmatrix} \lambda - a_{11} & -X^* \\ -X & \lambda I_{n-1} - M_1 \end{pmatrix}$$

to obtain the formula

$$\det(\lambda I_n - A) = \det(\lambda I_{n-1} - M_1)(\lambda - a_{11} - X^*(\lambda I_{n-1} - M_1)^{-1} X),$$

or equivalently

$$p_A(\lambda) = p_{M_1}(\lambda)(\lambda - a_{11} - X^*(\lambda I_{n-1} - M_1)^{-1} X),$$

which on Taylor expansion around $\lambda = 0$ using $p_A(0) = 0$ gives

$$p'_A(0)\lambda + O(\lambda^2) = (p_{M_1}(0) + O(\lambda))(\lambda - a_{11} + X^* M_1^{-1} X + \lambda X^* M_1^{-2} X + O(\lambda^2)).$$

Setting $\lambda = 0$ and using $p_{M_1}(0) \neq 0$, we conclude that $a_{11} + X^* M_1^{-1} X$ vanishes. If we then extract the λ coefficient, we obtain the claim.

Remark 16. The same calculations also give the well-known fact that the minor eigenvalues $\lambda_1(M_1), \dots, \lambda_{n-1}(M_1)$ are precisely the roots for the equation

$$\lambda - a_{11} - X^*(\lambda I_{n-1} - M_1)^{-1} X = 0.$$

Among other things, this can be used to establish the interlacing inequalities (1).

2.7. **A generalization.** The following generalization of the eigenvector-eigenvalue identity was recently observed[7] by Yu Qing Tang (private communication), relying primarily on the Cauchy–Binet formula and a duality relationship (24) between the various minors of a unitary matrix. If A is an $n \times n$ matrix and I, J are subsets of $\{1, \dots, n\}$ of the same cardinality m, let $M_{I,J}(A)$ denote the $n - m \times n - m$ minor formed by removing the m rows indexed by I and the m columns indexed by J.

Proposition 17 (Generalized eigenvector-eigenvalue identity)**.** *Let A be a normal $n \times n$ matrix diagonalized as $A = UDU^*$ for some unitary U and diagonal $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, let $1 \leq m < n$, and let $I, J, K \subset \{1, \dots, n\}$ have cardinality m. Then*

$$\overline{\det M_{J^c, I^c}(U)} (\det M_{K^c, I^c}(U)) \prod_{i \in I, j \in I^c} (\lambda_j - \lambda_i) = \det M_{J,K} \Big(\prod_{i \in I} (A - \lambda_i I_n) \Big),$$

[7] Variants of this identity have also been recently observed in [Che2019], [Sta2019].

where I^c denotes the complement of I in $\{1,\dots,n\}$, and similarly for J^c, K^c.

Note that if we set $m = 1$, $I = \{i\}$, and $J = K = \{j\}$, we recover (2). The identity (10) can be interpreted as the remaining $m = 1$ cases of this proposition.

Proof. We have

$$\prod_{i\in I}(A-\lambda_i I_n) = U\prod_{i\in I}(D-\lambda_i I_n)U^*,$$

and hence by the Cauchy–Binet formula

$$\begin{aligned}&\det M_{J,K}\Big(\prod_{i\in I}(A-\lambda_i I_n)\Big)\\ &\quad=\sum_{L,L'}(\det M_{J,L}(U))\Big(\det M_{L,L'}\Big(\prod_{i\in I}(D-\lambda_i I_n)\Big)\Big)(\det M_{L',K}(U^*)),\end{aligned}$$

where L, L' range over subsets of $\{1,\dots,n\}$ of cardinality m. A computation reveals that the quantity $\det M_{L,L'}(\prod_{i\in I}(D-\lambda_i I_n))$ vanishes unless $L = L' = I$, in which case the quantity equals $\prod_{i\in I, j\in I^c}(\lambda_j-\lambda_i)$. Thus it remains to show that

$$\overline{\det M_{J^c,I^c}(U)}(\det M_{K^c,I^c}(U)) = \det M_{J,I}(U)\det M_{I,K}(U^*).$$

Since $\det M_{I,K}(U^*) = \overline{\det M_{K,I}(U)}$, it will suffice to show that[8]

$$\det M_{J,I}(U) = \overline{\det M_{J^c,I^c}(U)}\det U \tag{24}$$

for any $J, I \subset \{1,\dots,n\}$ of cardinality m. By permuting rows and columns, we may assume that $J = I = \{1,\dots,m\}$. If we split the identity matrix I_n into the left m columns $I_n^1 := \begin{pmatrix} I_m \\ 0_{n-m\times m}\end{pmatrix}$ and the right $n-m$ columns $I_n^2 := \begin{pmatrix} 0_{m\times n-m} \\ I_{n-m}\end{pmatrix}$ and take determinants of both sides of the identity

$$U\begin{pmatrix} U^*I_n^1 & I_n^2\end{pmatrix} = \begin{pmatrix} I_n^1 & UI_n^2\end{pmatrix},$$

we conclude that

$$\det(U)\det M_{I^c,J^c}(U^*) = \det M_{J,I}(U),$$

giving the claim. □

3. History of the identity

In this section we present, roughly in chronological order, all the references to the eigenvector-eigenvalue identity (2) (or closely related results) that we are aware of in the literature. For the primary references, we shall present the identity in the notation of that reference in order to highlight the diversity of contexts and notational conventions in which this identity has appeared.

The earliest appearance of identities equivalent to (2) that we know of is due to Jacobi [Jac1834, §8, (33)]. In modern notation, Jacobi diagonalizes a symmetric quadratic form $\sum_{\chi=1}^n\sum_{\lambda=1}^n a_{\chi,\lambda}x_\chi x_\lambda$ as $\sum_{m=1}^n G_m(\sum_{i=1}^n \alpha_i^{(m)}x_i)^2$ for an orthogonal matrix $(\alpha_i^{(m)})_{1\le i,m\le n}$, and then the for each m the cofactors $B_{\chi\lambda}^{(m)}$ of the form $\sum_{\chi=1}^n\sum_{\lambda=1}^n a_{\chi,\lambda}x_\chi x_\lambda - G_m\sum_{\chi=1}^n x_\chi^2$ are extracted. Noting that the columns of this

[8]This identity is also a special case of the more general identity $\det M_{I,J}(\mathrm{adj}(A)) = (\det A)^{m-1}\det M_{J^c,I^c}(A)$ of Jacobi [Jac1834], which is valid for arbitrary $n\times n$ matrices A, as can be seen after noting that $\mathrm{adj}(U) = \det(U)^{-1}U^*$.

cofactor matrix are proportional to the eigenvector $(\alpha_i^{(m)})_{1\le i\le n}$, Jacobi concluded that

$$(G_1 - G_m)\cdots(G_n - G_m)\alpha_\chi^{(m)}\alpha_\lambda^{(m)} = B_{\chi\lambda}^{(m)} \tag{25}$$

with the factor $G_m - G_m$ omitted from the left-hand side; this is essentially (8) for real symmetric matrices. In [Jac1834, §8, (36)] an identity essentially equivalent to (4) for real symmetric matrices is also given.

An identity that implies (2) as a limiting case appears a century after Jacobi in a paper of Löwner [Löw1934, (7)]. In this paper, a diagonal quadratic form

$$A_n(x,x) = \sum_{i=1}^n \lambda_i x_i^2$$

is considered, as well as a rank one perturbation

$$B_n(x,x) = A_n(x,x) + \left(\sum_{i=1}^n \alpha_i x_i\right)^2$$

for some real numbers $\lambda_1,\dots,\lambda_n,\alpha_1,\dots,\alpha_n$. The eigenvalues of the quadratic form B_n are denoted $\mu_1,\dots,\mu_n$. If the eigenvalues are arranged in nondecreasing order, one has the interlacing inequalities

$$\lambda_1 \le \mu_1 \le \lambda_2 \le \cdots \le \lambda_n \le \mu_n$$

(compare with (1)). Under the nondegeneracy hypothesis

$$\lambda_1 < \mu_1 < \lambda_2 < \cdots < \lambda_n < \mu_n,$$

the identity

$$\alpha_k^2 = \frac{(\mu_1-\lambda_k)(\mu_2-\lambda_k)\cdots(\mu_n-\lambda_k)}{(\lambda_1-\lambda_k)(\lambda_2-\lambda_k)\cdots(\lambda_{k-1}-\lambda_k)(\lambda_{k+1}-\lambda_k)\cdots(\lambda_n-\lambda_k)} \tag{26}$$

is established for $k = 1,\dots,n$, which closely resembles (2); a similar formula also appears in [Jac1834, §12]. The identity (26) is obtained via "Eine einfache Rechnung" (an easy calculation) from the standard relations

$$\sum_{k=1}^n \frac{\alpha_k^2}{\mu_i - \lambda_k} = 1$$

for $i = 1,\dots,n$ (compare with (17)), after applying Cramer's rule and the Cauchy determinant identity (19). As such, it is very similar to the proof of (2) in Section 2.2 that is also based on (19). The identity (26) was used in [Löw1934] to help classify monotone functions of matrices and has also been applied to inverse eigenvector problems and stable computation of eigenvectors [GE1994], [Dem1997, pp. 224–226]. It can be related to (2) as follows. For sake of notation let us just consider the $j = n$ case of (2). Let ε be a small parameter, and consider the perturbation $e_n e_n^* + \varepsilon A$ of the rank one matrix $e_n e_n^*$. Standard perturbative analysis reveals that the eigenvalues of this perturbation consist of $n-1$ eigenvalues of the form $\varepsilon\lambda_i(M_n)+O(\varepsilon^2)$ for $i = 1,\dots,n-1$, plus an outlier eigenvalue at $1+O(\varepsilon)$. Rescaling, we see that the rank one perturbation $A + \frac{1}{\varepsilon}e_n e_n^*$ of A has eigenvalues of the form $\lambda_i(M_n) + O(\varepsilon)$ for $i = 1,\dots,n-1$, plus an outlier eigenvalue at $\frac{1}{\varepsilon} + O(1)$. If we

let A_n, B_n be the quadratic forms associated to $A, A + \frac{1}{\varepsilon} e_n e_n^*$ expressed using the eigenvector basis $v_1, \dots, v_n$, the identity (26) becomes

$$\frac{1}{\varepsilon}|v_{k,n}|^2 = \frac{\prod_{i=1}^n (\lambda_i(A + \frac{1}{\varepsilon} e_n e_n^*) - \lambda_k(A))}{\prod_{i=1; i \neq k}^n (\lambda_i(A) - \lambda_k(A))}.$$

Extracting the $1/\varepsilon$ component of both sides of this identity using the aforementioned perturbative analysis, we recover (2) after a brief calculation.

A more complicated variant of (26) involving various quantities related to the Rayleigh–Ritz method of bounding the eigenvalues of a symmetric linear operator was stated by Weinberger [Wei1960, (2.29)], where it was noted that it can be proven using much the same method as in [Löw1934].

The first appearance of the eigenvector-eigenvalue identity in essentially the form presented here that we are aware of was by Thompson [Tho1966, (15)], which does not reference the prior work of Löwner or Weinberger. In the notation of Thompson's paper, A is a normal $n \times n$ matrix, and $\mu_1, \dots, \mu_s$ are the *distinct* eigenvalues of A, with each μ_i occurring with multiplicity e_i. To avoid nondegeneracy it is assumed that $s \geq 2$. One then diagonalizes $A = UDU^{-1}$ for a unitary U and diagonal $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, and then sets

$$\theta_{i\beta} = \sum_{j : \lambda_j = \mu_\beta} |U_{ij}|^2$$

for $i = 1, \dots, n$ and $\beta = 1, \dots, n-1$, where U_{ij} are the coefficients of U. The minor formed by removing the ith row and column from A is denoted $A(i|i)$; it has "trivial" eigenvalues in which each μ_i with $e_i > 1$ occurs with multiplicity $e_i - 1$, as well as some "nontrivial" eigenvalues $\xi_{i1}, \dots, \xi_{i,s-1}$. The equation [Tho1966, (15)] then reads

$$\theta_{i\alpha} = \prod_{j=1}^{s-1} (\mu_\alpha - \xi_{ij}) \prod_{j=1, j \neq \alpha}^{s} (\mu_\alpha - \mu_j)^{-1} \tag{27}$$

for $1 \leq \alpha \leq s$ and $1 \leq i \leq n$. If one specializes to the case when A is Hermitian with simple spectrum, so that all the multiplicities e_i are equal to 1, and set $s = n$ and $\mu_i = \lambda_i$, it is then not difficult to verify that this identity is equivalent to the eigenvector-eigenvalue identity (2) in this simple spectrum case. In the case of repeated eigenvalues, the eigenvector-eigenvalue identity (2) may degenerate (in that the left and right-hand sides both vanish), but the identity (27) remains nontrivial in this case. The proof of (27) given in [Tho1966] is written using a rather complicated notation (in part because much of the paper was concerned with more general $k \times k$ minors rather than the $n-1 \times n-1$ minors $A(i|i)$), but it is essentially the adjugate proof from Section 2.1 (where the adjugate matrix is replaced by the closely related $(n-1)$-th compound matrix). In [Tho1966], the identity (27) was not highlighted as a result of primary interest in its own right, but was instead employed to establish a large number of inequalities between the eigenvalues $\mu_1, \dots, \mu_s$ and the minor eigenvalues $\xi_{i1}, \dots, \xi_{i,s-1}$ in the Hermitian case; see [Tho1966, Section 5].

In a followup paper [TM1968] by Thompson and McEnteggert, the analysis from [Tho1966] was revisited, restricting attention specifically to the case of an $n \times n$ Hermitian matrix H with simple eigenvalues $\lambda_1 < \cdots < \lambda_n$, and with the minor $H(i|i)$ formed by deleting the ith row and column having eigenvalues

$\xi_{i1} \leq \cdots \leq \xi_{i,n-1}$. In this paper the inequalities

$$\sum_{i=1}^{n} \frac{\lambda_j - \xi_{i,j-1}}{\lambda_j - \lambda_{j-1}} \frac{\xi_{ij} - \lambda_j}{\lambda_{j+1} - \lambda_j} \geq 1 \tag{28}$$

and

$$\sum_{i=1}^{n} \frac{\lambda_j - \xi_{i,j-1}}{\lambda_j - \lambda_1} \frac{\xi_{ij} - \lambda_j}{\lambda_n - \lambda_j} \leq 1 \tag{29}$$

for $1 \leq j \leq n$ were proved (with most cases of these inequalities already established in [Tho1966]), with a key input being the identity

$$|u_{ij}|^2 = \left\{ \frac{\lambda_j - \xi_{i1}}{\lambda_j - \lambda_1} \right\} \cdots \left\{ \frac{\lambda_j - \xi_{i,j-1}}{\lambda_j - \lambda_{j-1}} \right\} \left\{ \frac{\xi_{ij} - \lambda_j}{\lambda_{j+1} - \lambda_j} \right\} \cdots \left\{ \frac{\xi_{i,n-1} - \lambda_j}{\lambda_n - \lambda_j} \right\}, \tag{30}$$

where u_{ij} are the components of the unitary matrix U used in the diagonalization $H = UDU^{-1}$ of H. Note from the Cauchy interlacing inequalities that each of the expressions in braces takes values between 0 and 1. It is not difficult to see that this identity is equivalent to (2) (or (27)) in the case of Hermitian matrices with simple eigenvalues, and the hypothesis of simple eigenvalues can then be removed by the usual limiting argument. As in [Tho1966], the identity is established using adjugate matrices, essentially by the argument given in the previous section. However, the identity (30) is only derived as an intermediate step towards establishing the inequalities (28), (29), and is not highlighted as of interest in its own right. The identity (27) was then reproduced in a further followup paper [Tho1969], in which the identity (14) was also noted; this latter identity was also independently observed in [DH1978].

In the text of Šilov [Š1969, Section 10.27], the identity (2) is established, essentially by the Cramer rule method. Namely, if $A(x,x)$ is a diagonal real quadratic form on $\mathbb{R}^n$ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$, and R_{n-1} is a hyperplane in $\mathbb{R}^n$ with unit normal vector $(\alpha_1, \ldots, \alpha_n)$, and $\mu_1 \geq \cdots \geq \mu_{n-1}$ are the eigenvalues of A on R_{n-1}, then it is observed that

$$\alpha_k^2 = \frac{\prod_{k=1}^{n-1}(\mu_k - \lambda)}{(\lambda_k - \lambda_1)\cdots(\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k+1})\cdots(\lambda_k - \lambda_n)} \tag{31}$$

for $k = 1, \ldots, n$, which is (2) after changing to the eigenvector basis; identities equivalent to (17) and (4) are also established. The text [Š1969] gives no references, but given the similarity of notation with [Löw1934] (compare (31) with (26)), one could speculate that Šilov was influenced by Löwner's work.

In a section [Pai1971, Section 8.2] of the PhD thesis of Paige entitled "A Useful Theorem on Cofactors", the identity (30) is cited as "a fascinating theorem ... that relates the elements of the eigenvectors of a symmetric to its eigenvalues and the eigenvalues of its principal submatrices", with a version of the adjugate proof given. In the notation of that thesis, one considers a $k \times k$ real symmetric tridiagonal matrix C with distinct eigenvalues $\mu_1 > \cdots > \mu_k$ with an orthonormal of eigenvectors $y_1, \ldots, y_k$. For any $0 \leq r < j \leq k$, let $C_{r,j}$ denote the $j - r \times j - r$ minor of C defined by taking the rows and columns indexed between $r+1$ and j, and let $p_{i,j}(\mu) := \det(\mu I_{j-r} - C_{r,j})$ denote the associated trigonometric polynomial. The identity

$$y_{ri}^2 = p_{0,r-1}(\mu_i) p_{r,k}(\mu_i) / f(i) \tag{32}$$

is then established for $i = 1, \dots, n$, where y_{ri} is the ith component of y_r and $f(i) := \prod_{r=1;r\neq i}^{k}(\mu_i - \mu_r)$. This is easily seen to be equivalent to (2) in the case of real symmetric tridiagonal matrices with distinct eigenvalues. One can then use this to derive (2) for more general real symmetric matrices by a version of the Lanczos algorithm for tridiagonalizing an arbitrary real symmetric matrix, followed by the usual limiting argument to remove the hypothesis of distinct eigenvalues; we leave the details to the interested reader. Returning to the case of tridiagonal matrices, Paige also notes that the method also gives the companion identity

$$f(i)y_{ri}y_{si} = \delta_{r+1}\cdots\delta_s p_{0,r-1}(\mu_i)p_{s,k}(\mu_i) \tag{33}$$

for $1 \leq r < s \leq k$, where $\delta_2, \dots, \delta_k$ are the upper diagonal entries of the tridiagonal matrix C; this can be viewed as a special case of (10). These identities were then used in [Pai1971, Section 8] as a tool to bound the behavior of errors in the symmetric Lanczos process.

Paige's identities (32), (33) for tridiagonal matrices are reproduced in the textbook of Parlett [Par1980, Theorem 7.9.2], with slightly different notation. Namely, one starts with an $n\times n$ real symmetric tridiagonal matrix T, decomposed spectrally as $S\Theta S^*$ where $S = (s_1, \dots, s_n)$ is orthogonal and $\Theta = \mathrm{diag}(\theta_1, \dots, \theta_n)$. Then for $1 \leq \mu \leq \nu \leq n$ and $1 \leq j \leq n$, the jth component $s_{\mu j}$ of s_μ is observed to obey the formula

$$s_{\mu j}^2 = \chi_{1:\mu-1}(\theta_j)\chi_{\mu+1:n}(\theta_j)/\chi'_{1:n}(\theta_j)$$

when θ_j is a simple eigenvalue, where $\chi_{i:j}$ is the characteristic polynomial of the $j-i+1 \times j-i+1$ minor of T formed by taking the rows and columns between i and j. This identity is essentially equivalent to (32). The identity (32) is similarly reproduced in this notation; much as in [Pai1971], these identities are then used to analyze various iterative methods for computing eigenvectors. The proof of the theorem is left as an exercise in [Par1980], with the adjugate method given as a hint. Essentially the same result is also stated in the text of Golub and van Loan [GVL1983, pp. 432–433] (equation (8.4.12) on page 474 in the 2013 edition), proven using a version of the Cramer rule arguments in Section 2.2. They cite as reference the earlier paper [Gol1973, (3.6)], which also uses essentially the same proof (see also [Gla2004, (4.3.17)], who cites [BG1978], who in turn cite [Gol1973]). A similar result was stated without proof by Galais, Kneller, and Volpe [GKV2012, equations (6), (7)]. They provided expressions for both $|v_{i,j}|^2$ and the off-diagonal eigenvectors as a function of cofactors in place of adjugate matrices. Their work was in the context of neutrino oscillations.

The identities of Parlett and of Golub and Van Loan are cited in the thesis of Knyazev [Kny1986, (2.2.27)], again to analyze methods for computing eigenvalues and eigenvectors; the identities of Golub and Van Loan and of Šilov are similarly cited in the paper of Knyazev and Skorokhodov [KS1991] for similar purposes. Parlett's result is also reproduced in the text of Xu [Xu1995, (3.19)]. In the survey [CG2002, (4.9)] of Chu and Golub on structured inverse eigenvalue problems, the eigenvector-eigenvalue identity is derived via the adjugate method from the results of [TM1968], and it is used to solve the inverse eigenvalue problem for Jacobi matrices; the text [Par1980] is also cited.

In the paper [DLNT1986, p. 210] of Deift, Li, Nanda, and Tomei, the eigenvector-eigenvalue identity (2) is derived by the Cramer rule method, and is used to

construct action-angle variables for the Toda flow. The paper cites [BG1978], which also reproduces (2) as equation (1.5) of that paper, and in turn cites [Gol1973].

In the paper of Mukherjee and Datta [MD1989] the eigenvector-eigenvalue identity was rediscovered, in the context of computing eigenvectors of graphs that arise in chemistry. If G is a graph on n vertices $v_1, \dots, v_n$, and $G - v_r$ is the graph on $n-1$ vertices formed by deleting a vertex $v_r, r = 1, \dots, n$, then in [MD1989, (4)] the identity

$$P(G - v_r; x_j) = P'(G; x_j) C_{rj}^2 \tag{34}$$

is established for $j, r = 1, \dots, n$, where $P(G; x)$ denotes the characteristic polynomial of the adjacency matrix of G evaluated at x, and C_{rj} is the coefficient at the rth vertex of the eigenvector corresponding to the jth eigenvalue, and one assumes that all the eigenvalues of G are distinct. This is equivalent to (4) in the case that A is an adjacency matrix of a graph. The identity is proven using the perturbative method in Section 2.4, and it appears to have been discovered independently. A similar identity was also noted in the earlier work of Li and Feng [LF1979], at least in the case $j = 1$ of the largest eigenvalue. In a later paper of Hagos [Hag2002], it is noted that the identity (34) "is probably not as well known as it should be", and it also carefully generalizes (34) to an identity (essentially the same as (16)) that holds when some of the eigenvalues are repeated. An alternate proof of (34) was given in the paper of Cvetkovic, Rowlinson, and Simic [CRS2007, Theorem 3.1], essentially using the Cramer rule type methods in Section 2.2. The identity (14) is also essentially noted at several other locations in the graph theory literature, such as [God1993, Chapter 4], [GM1981, Lemma 2.1], [God2012, Lemma 7.1, Corollary 7.2], [GGKL2017, (2)] in relation to the generating functions for walks on a graph, though in those references no direct link to the eigenvector-eigenvalue identity in the form (2) is asserted.

In [NTU1993, Section 2] the identity (2) is derived for normal matrices by the Cramer rule method, citing [Tho1969], [DH1978] as the source for the key identity (14); the papers [Tho1966], [TM1968] also appear in the bibliography but were not directly cited in this section. An extension to the case of eigenvalue multiplicity, essentially corresponding to (16), is also given. This identity is then used to give a complete description of the relations between the eigenvalues of A and of a given minor M_j when A is assumed to be normal. In [BFdP2011] a generalization of these results was given to the case of J-normal matrices for some diagonal sign matrix J; this corresponds to a special case of (12) in the case where each left eigenvector w_i is the complex conjugate of Jv_i.

The paper of Baryshnikov [Bar2001] marks the first appearance of this identity in random matrix theory. Let H be a Hermitian form on $\mathbb{C}^M$ with eigenvalues $\lambda_1 \geq \dots \geq \lambda_M$, and let L be a hyperplane of $\mathbb{C}^M$ orthogonal to some unit vector l. Let l_i be the component of l with respect to an eigenvector v_i associated to λ_i, set $w_i := |l_i|^2$, and let $\mu_1 \geq \dots \geq \mu_{M-1}$ be the eigenvalues of the Hermitian form arising from restricting H to L. Then after [Bar2001, (4.5.2)] (and correcting some typos) the identity

$$w_i = \frac{\prod_{1 \leq j \leq M-1} (\lambda_i - \mu_j)}{\prod_{1 \leq j \leq M; j \neq i} (\lambda_i - \lambda_j)}$$

is established, by an argument based on Cramer's rule and the Cauchy determinant formula (19), similar to the arguments at the end of Section 2.2, and it appears

to have been discovered independently. If one specializes to the case when l is a standard basis vector e_j, then l_i is also the e_j component of v_i, and we recover (2) after a brief calculation. This identity was employed in [Bar2001] to study the situation in which the hyperplane normal l was chosen uniformly at random on the unit sphere. This formula was rederived (using a version of the Cramer rule method in Section 2.2) in the May 2019 paper of Forrester and Zhang [FZ2019, (2.7)], who recover some of the other results in [Bar2001] as well, and study the spectrum of the sum of a Hermitian matrix and a random rank one matrix.

In the paper [DE2002, Lemma 2.7] of Dumitriu and Edelman, the identity (32) of Paige (as reproduced in [Par1980, Theorem 7.9.2]) is used to give a clean expression for the Vandermonde determinant of the eigenvalues of a tridiagonal matrix, which is used in turn to construct tridiagonal models for the widely studied *β-ensembles* in random matrix theory.

In the unpublished preprint [Van2014] of Van Mieghem, the identity (4) is prominently displayed as the main result, though in the notation of that preprint it is expressed instead as

$$(x_k)_j^2 = -\frac{1}{c_A'(\lambda_k)}\det(A_{\backslash\{j\}} - \lambda_k I_n)$$

for any $j, k = 1, \dots, n$, where A is a real symmetric matrix with distinct eigenvalues $\lambda_1, \dots, \lambda_n$ and unit eigenvectors $x_1, \dots, x_n$, $A_{\backslash\{j\}}$ is the minor formed by removing the jth row and column from A, and c_A' is the derivative of the (sign-reversed) characteristic polynomial $c_A(\lambda) = \det(A - \lambda I_n) = (-1)^n p_A(\lambda)$. Two proofs of this identity are given, one being essentially the Cramer's rule proof from Section 2.2 and attributed to the previous reference [CRS2007]; the other proof is based on Cramer's rule and the Desnanot–Jacobi identity (Dodgson condensation); this identity is used to quantify the effect of removing a node from a graph on the spectral properties of that graph. The related identity (23) from [TV2011] is also noted in this preprint. Some alternate formulae from [VM2011] for quantities such as $(x_k)_j^2$ in terms of walks of graphs are also noted, with the earlier texts [God1993], [GVL1983] also cited.

The identity (2) was independently rediscovered and then generalized by Kausel [Kau2018] as a technique to extract information about components of a generalized eigenmode without having to compute the entire eigenmode. Here the generalized eigenvalue problem

$$\mathbf{K}\psi_j = \lambda_j \mathbf{M}\psi_j$$

for $j = 1, \dots, N$ is considered, where $\mathbf{K}$ is a positive semidefinite $N \times N$ real symmetric matrix, $\mathbf{M}$ is a positive definite $N \times N$ real symmetric matrix, and the matrix $\mathbf{\Psi} = (\psi_1 \cdots \psi_N)$ of eigenfunctions is normalized so that $\mathbf{\Psi}^T\mathbf{M}\mathbf{\Psi} = \mathbf{I}$. For any $1 \le \alpha \le N$, one also solves the constrained system

$$\mathbf{K}_\alpha \psi_j^{(\alpha)} = \lambda_{\alpha j}\mathbf{M}_\alpha \psi_j^{(\alpha)},$$

where $\mathbf{K}_\alpha, \mathbf{M}_\alpha$ are the $N-1 \times N-1$ minors of $\mathbf{K}, \mathbf{M}$, respectively, formed by removing the αth row and column. Then in [Kau2018, (18)] the Cramer rule method is used to establish the identity

$$\psi_{\alpha j} = \pm\sqrt{\frac{|\mathbf{M}_\alpha|}{|\mathbf{M}|}}\sqrt{\frac{\prod_{k=1}^{N-1}(\lambda_j - \lambda_{\alpha k})}{\prod_{k=1;k\neq j}^{N-1}(\lambda_j - \lambda_k)}}$$

for the α component $\psi_{\alpha j}$ of ψ_j, where $|\mathbf{M}|$ is the notation in [Kau2018] for the determinant of $\mathbf{M}$. Specializing to the case when $\mathbf{M}$ is the identity matrix, we recover (2).

The eigenvector-eigenvalue identity was discovered by three of us [DPZ2020] in July 2019, initially in the case of 3×3 matrices, in the context of trying to find a simple and numerically stable formula for the eigenvectors of the neutrino oscillation Hamiltonian, which form a separate matrix known as the PMNS lepton mixing matrix. This identity was established in the 3×3 case by direct calculation. Despite being aware of the related identity (23), the four of us were unable to locate this identity in past literature and wrote a preprint [DPTZ2019] in August 2019 highlighting this identity and providing two proofs (the adjugate proof from Section 2.1, and the Cauchy–Binet proof from Section 2.5). The release of this preprint generated some online discussion,[9] and we were notified by Jiyuan Zhang (private communication) of the prior appearance of the identity earlier in the year in [FZ2019]. However, the numerous other places in the literature in which some form of this identity appeared did not become revealed until a popular science article [Wol2019] by Wolchover was written in November 2019. This article spread awareness of the eigenvector-eigenvalue identity to a vastly larger audience, and generated a large number of reports of previous occurrences of the identity, as well as other interesting related observations, which we have attempted to incorporate into this survey.

4. Further discussion

The eigenvector-eigenvalue identity (2) only yields information about the magnitude $|v_{i,j}|$ of the components of a given eigenvector v_i, but it does not directly reveal the phase of these components. On one hand, this is to be expected, since (as already noted in the consistency check (vii) in the introduction) one has the freedom to multiply v_i by a phase; for instance, even if one restricts attention to real symmetric matrices A and requires the eigenvectors to be real v_i, one has the freedom to replace v_i by its negation $-v_i$, so the sign of each component $v_{i,j}$ is ambiguous. However, *relative* phases, such as the phase of $v_{i,j}\overline{v_{i,j'}}$ are not subject to this ambiguity. There are several ways to try to recover these relative phases. One way is to employ the off-diagonal analogue (10) of (2), although the determinants in that formula may be difficult to compute in general. For small matrices, it was suggested in [MD1989] that the signs of the eigenvectors could often be recovered by direct inspection of the components of the eigenvector equation $Av_i = \lambda_i(A)v_i$. In the application in [DPZ2020], the additional phase could be recovered by a further neutrino specific identity [Tos1991]. For more general matrices, one way to retrieve such phase information is to apply (2) in multiple bases. For instance, suppose A was real symmetric and the $v_{i,j}$ were all real. If one were to apply the eigenvector-eigenvalue identity after changing to a basis that involved the unit vector $\frac{1}{\sqrt{2}}(e_j + e_{j'})$, then one could use the identity to evaluate the magnitude of $\frac{1}{\sqrt{2}}(v_{i,j} + v_{i,j'})$. Two further applications of the identity in the original basis would give the magnitude of $v_{i,j}, v_{i,j'}$, and this is sufficient information to determine the relative sign of $v_{i,j}$ and $v_{i,j'}$. We also remark that for real symmetric matrices

[9] `terrytao.wordpress.com/2019/08/13`, `www.reddit.com/r/math/comments/cq3en0`

that are acyclic (such as weighted adjacency matrices of graphs that do not contain loops), one can write down explicit formulae for the coefficients of eigenvectors (and not just their magnitudes) in terms of characteristic polynomials of minors; see [BK2016]. We do not know of any direct connection between such formulae and the eigenvector-eigenvalue identity (2).

For large unstructured matrices, it does not seem at present that the identity (2) provides a competitive algorithm to compute eigenvectors. Indeed, to use this identity to compute all the eigenvector component magnitudes $|v_{i,j}|$, one would need to compute all $n-1$ eigenvalues of each of the n minors $M_1, \dots, M_n$, which would be a computationally intensive task in general; furthermore, an additional method would then be needed to also calculate the signs or phases of these components. However, if the matrix is of a special form (such as a tridiagonal form), then the identity could be of more practical use, as witnessed by the uses of this identity (together with variants such as (33)) in the literature to control the rate of convergence for various algorithms to compute eigenvalues and eigenvectors of tridiagonal matrices. Also, as noted recently in [Kau2018], if one has an application that requires only the component magnitudes $|v_{1,j}|, \dots, |v_{n,j}|$ at a single location j, then one only needs to compute the characteristic polynomial of a single minor M_j of A at a single value $\lambda_i(A)$, and this may be more computationally feasible.

5. Sociology of science issues

As one sees from Section 3 and Figure 1, there was some partial dissemination of the eigenvector-eigenvalue identity amongst some mathematical communities, to the point where it was regarded as "folklore" by several of these communities. However, this process was unable to raise broader awareness of this identity, resulting in the remarkable phenomenon of multiple trees of references sprouting from independent roots, and only loosely interacting with each other. For instance, as discussed in the previous section, for two months after the release of our own preprint [DPTZ2019], we only received a single report of another reference [FZ2019] containing a form of the identity, despite some substantial online discussion and the dozens of extant papers on the identity. It was only in response to the popular science article [Wol2019] that awareness of the identity finally "went viral", leading to what was effectively an *ad hoc* crowdsourced effort to gather all the prior references to the identity in the literature. While we do not know for certain why this particular identity was not sufficiently well known prior to these recent events, we can propose the following possible explanations:

(1) *The identity was mostly used as an auxiliary tool for other purposes.* In almost all of the references discussed here, the eigenvector-eigenvalue identity was established only in order to calculate or bound some other quantity; it was rarely formalized as a theorem or even as a lemma. In particular, with a few notable exceptions, such as the preprint [Van2014], this identity would not be mentioned in the title, abstract, or even the introduction. In a few cases, the identity was reproven by authors who did not seem to be fully aware that it was already established in one of the references in their own bibliography.

(2) *The identity does not have a standard name, form, or notation, and does not involve uncommon keywords.* As one can see from Section 3, the identity comes in many variants and can be rearranged in a large number of ways;

furthermore, the notation used for the various mathematical objects appearing in the identity vary greatly depending on the intended application or on the authors involved. Also, none of the previous references attempted to give the identity a formal name, and the keywords used to describe the identity (such as "eigenvector" or "eigenvalue") are in extremely common use in mathematics. As such, there are no obvious ways to use modern search engines to locate other instances of this identity, other than by manually exploring the citation graph around known references to that identity. Perhaps a "fingerprint database" for identities [BT2013] would be needed before such automated searches could become possible.

(3) *The field of linear algebra is too mature, and its domain of applicability is too broad.* The vast majority of consumers of linear algebra are not domain experts in linear algebra itself, but instead use it as a tool for a very diverse array of other applications. As such, the diffusion of linear algebra knowledge is not guided primarily by a central core of living experts in the field, but instead relies on more mature sources of authority such as textbooks and lectures. Unfortunately, only a small handful of linear algebra textbooks mention the eigenvector-eigenvalue identity, thus preventing wider dissemination of this identity.

Online discussion forums for mathematics were only partially successful in disseminating this identity. For instance, the 2012 MathOverflow question[10] "Cramer's rule for eigenvectors", which inquired as to the existence of an eigenvector identity such as (2), received nearly ten-thousand views, but only revealed the related identity in Lemma 15. Nevertheless, this post was instrumental in bringing these four authors together to produce the preprint [DPTZ2019], via a comment[11] on a Reddit post by one of us.

It is not fully clear to us how best to attribute authorship for the eigenvector-eigenvalue identity (2). A variant of the identity was observed by Jacobi [Jac1834], but not widely propagated. An identity that implies (2) was later given by Löwner [Löw1934], but the implication is not immediate, and this reference had only a modest impact on the subsequent literature. The paper of Thompson [Tho1966] is the first place we know of in which the identity explicitly appears, and it was propagated through citations into several further papers in the literature. But this did not prevent the identity from then being independently rediscovered several further times, such as in the text [GVL1983] (with the latter restricting attention to the case of tridiagonal matrices). Furthermore, we are not able to guarantee that there is not an even earlier place in the literature where some form of this identity has appeared. We propose the name "eigenvector-eigenvalue identity" for (2) on the grounds that it is descriptive, and hopefully it is a term that can be detected through search engines by researchers looking for identities of this form.

Although, in this survey we have included approximately 50 references that mention some variant of the eigenvector-eigenvalue identity, in most cases the identity does not explicitly appear in a form such as (2) that specifically links eigenvector component magnitudes of an arbitrary Hermitian matrix to eigenvalues of the matrix and its minors. Exceptions include the papers [Tho1966], [TM1968],

[10] `mathoverflow.net/questions/96190`

[11] `www.reddit.com/r/math/comments/ci665j/linear_algebra_question_from_a_physicist/ev22xgp/`

[NTU1993], and (in the special case of tridiagonal matrices) [GVL1983], [Xu1995]. To convert the other forms of the identity appearing in the literature to a form similar to (2) requires a small but nonzero amount of additional work (such as a change of basis, passing to a limit, or expressing a characteristic polynomial or determinant in terms of eigenvalues). This may well be an additional factor that has prevented this identity from being more widely known until recently.

Acknowledgments

We thank Asghar Bahmani, Carlo Beenakker, Adam Denchfield, Percy Deift, Laurent Demanet, Alan Edelman, Chris Godsil, Aram Harrow, James Kneller, Andrew Knysaev, Manjari Narayan, Michael Nielsen, Małgorzata Stawiska, Karl Svozil, Gang Tian, Carlos Tomei, Piet Van Mieghem, Brad Willms, Fu Zhang, Jiyuan Zhang, and Zhenzhong Zhang for pointing out a number of references where some variant of the eigenvector-eigenvalue identity has appeared in the literature, or suggesting various extensions and alternate proofs of the identity. We also thank Darij Grinberg, Andrew Krause, Kristoffer Varholm, Jochen Voss and Jim Van Zandt for some corrections to previous versions of this manuscript.

References

[AS1964] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, National Bureau of Standards Applied Mathematics Series, vol. 55, For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964. MR0167642

[Bar2001] Yu. Baryshnikov, *GUEs and queues*, Probab. Theory Related Fields **119** (2001), no. 2, 256–274, DOI 10.1007/PL00008760. MR1818248

[BFdP2011] N. Bebiano, S. Furtado, and J. da Providência, *On the eigenvalues of principal submatrices of J-normal matrices*, Linear Algebra Appl. **435** (2011), no. 12, 3101–3114, DOI 10.1016/j.laa.2011.05.033. MR2831599

[BG1978] D. Boley and G. H. Golub, *Inverse eigenvalue problems for band matrices*, Numerical analysis (Proc. 7th Biennial Conf., Univ. Dundee, Dundee, 1977), Springer, Berlin, 1978, pp. 23–31. Lecture Notes in Math., Vol. 630. MR0474741

[BK2016] A. Bahmani and D. Kiani, *An explicit formula for the eigenvectors of acyclic matrices and weighted trees*, `https://arxiv.org/abs/1609.02399`, 2016.

[BM1941] G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, Macmillan Company, New York, 1941. MR0005093

[BT2013] S. C. Billey and B. E. Tenner, *Fingerprint databases for theorems*, Notices Amer. Math. Soc. **60** (2013), no. 8, 1034–1039, DOI 10.1090/noti1029. MR3113227

[Cau1841] A. L. Cauchy, *Exercices d'analyse et de physique mathématique*, vol. 2. Bachelier, page 154, 1841.

[CG2002] M. T. Chu and G. H. Golub, *Structured inverse eigenvalue problems*, Acta Numer. **11** (2002), 1–71, DOI 10.1017/S0962492902000016. MR2008966

[Che2019] X. Chen, *Note on eigenvectors from eigenvalues*, `https://arxiv.org/abs/1911.09081`, 2019.

[Cot1974] R. W. Cottle, *Manifestations of the Schur complement*, Linear Algebra Appl. **8** (1974), 189–211, DOI 10.1016/0024-3795(74)90066-4. MR354727

[Cra1750] G. Cramer, *Introduction à l'analyse des lignes courbes algébriques*, Chez les frères Cramer et C. Philibert, Geneva, pages 657–659, 1750.

[CRS2007] D. Cvetković, P. Rowlinson, and S. K. Simić, *Star complements and exceptional graphs*, Linear Algebra Appl. **423** (2007), no. 1, 146–154, DOI 10.1016/j.laa.2007.01.008. MR2312331

[DE2002] I. Dumitriu and A. Edelman, *Matrix models for beta ensembles*, J. Math. Phys. **43** (2002), no. 11, 5830–5847, DOI 10.1063/1.1507823. MR1936554

[Dem1997] J. W. Demmel, *Applied numerical linear algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997. MR1463942

[Den2019] P. B. Denton, *Eigenvalue-eigenvector identity code*, `https://github.com/PeterDenton/Eigenvector-Eigenvalue-Identity`, 2019.

[DH1978] E. Deutsch and H. Hochstadt, *On Cauchy's inequalities for Hermitian matrices*, Amer. Math. Monthly **85** (1978), no. 6, 486–487, DOI 10.2307/2320075. MR491766

[DLNT1986] P. Deift, L. C. Li, T. Nanda, and C. Tomei, *The Toda flow on a generic orbit is integrable*, Comm. Pure Appl. Math. **39** (1986), no. 2, 183–232, DOI 10.1002/cpa.3160390203. MR820068

[DPTZ2019] Peter B. Denton, Stephen J. Parke, Terence Tao, and Xining Zhang. *Eigenvectors from Eigenvalues*, arXiv e-prints, page `arXiv:1908.03795v1`, Aug 2019.

[DPZ2020] P. B. Denton, S. J. Parke, and X. Zhang, *Neutrino oscillations in matter via eigenvalues*, Phys. Rev. D **101** (2020), no. 9, 093001, 12, DOI 10.1103/physrevd.101.093001. MR4111110

[ESY2009] L. Erdős, B. Schlein, and H.-T. Yau, *Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices*, Ann. Probab. **37** (2009), no. 3, 815–852, DOI 10.1214/08-AOP421. MR2537522

[FZ2019] P. J. Forrester and J. Zhang, *Co-rank 1 projections and the randomised Horn problem*, arXiv e-prints, page `arXiv:1905.05314`, May 2019.

[Gan1959] F. R. Gantmacher, *The theory of matrices. Vols. 1, 2*, Translated by K. A. Hirsch, Chelsea Publishing Co., New York, 1959. MR0107649

[GE1994] M. Gu and S. C. Eisenstat, *A stable and efficient algorithm for the rank-one modification of the symmetric eigenproblem*, SIAM J. Matrix Anal. Appl. **15** (1994), no. 4, 1266–1276, DOI 10.1137/S089547989223924X. MR1293916

[GGKL2017] C. Godsil, K. Guo, M. Kempton, and G. Lippner, *State transfer in strongly regular graphs with an edge perturbation*, Journal of Combinatorial Theory, Series A, `https://www.sciencedirect.com/science/article/abs/pii/S0097316519301621`, 2017.

[GKV2012] Sébastien Galais, James Kneller, and Cristina Volpe, *The neutrino-neutrino interaction effects in supernovae: the point of view from the matter basis*, J. Phys., G39:035201, 2012.

[Gla2004] G. M. L. Gladwell, *Inverse problems in vibration*, 2nd ed., Solid Mechanics and its Applications, vol. 119, Kluwer Academic Publishers, Dordrecht, 2004. MR2102477

[GM1981] C. D. Godsil and B. D. McKay, *Spectral conditions for the reconstructibility of a graph*, J. Combin. Theory Ser. B **30** (1981), no. 3, 285–289, DOI 10.1016/0095-8956(81)90046-0. MR624545

[God1993] C. D. Godsil, *Algebraic combinatorics*, Chapman and Hall Mathematics Series, Chapman & Hall, New York, 1993. MR1220704

[God2012] C. Godsil, *When can perfect state transfer occur?*, Electron. J. Linear Algebra **23** (2012), 877–890, DOI 10.13001/1081-3810.1563. MR2992400

[Gol1973] G. H. Golub, *Some modified matrix eigenvalue problems*, SIAM Rev. **15** (1973), 318–334, DOI 10.1137/1015032. MR329227

[GS1995] B. Gaveau and L. S. Schulman, *Limited quantum decay*, J. Phys. A **28** (1995), no. 24, 7359–7374. MR1381425

[GVL1983] G. H. Golub and C. F. Van Loan, *Matrix computations*, Johns Hopkins Series in the Mathematical Sciences, vol. 3, Johns Hopkins University Press, Baltimore, MD, 1983. MR733103

[Hag2002] E. M. Hagos, *Some results on graph spectra*, Linear Algebra Appl. **356** (2002), 103–111, DOI 10.1016/S0024-3795(02)00324-5. Special issue on algebraic graph theory (Edinburgh, 2001). MR1944680

[Hal1942] P. R. Halmos, *Finite Dimensional Vector Spaces*, Annals of Mathematics Studies, no. 7, Princeton University Press, Princeton, N.J., 1942. MR0006591

[Jac1834] C. G. J. Jacobi, *De binis quibuslibet functionibus homogeneis secundi ordinis per substitutiones lineares in alias binas tranformandis, quae solis quadratis variabilium constant; una cum variis theorematis de tranformatione etdeterminatione integralium multiplicium* (Latin), J. Reine Angew. Math. **12** (1834), 1–69, DOI 10.1515/crll.1834.12.1. MR1577999

[Kau2018] E. Kausel, *Normalized modes at selected points without normalization* Journal of Sound and Vibration, 420:261–268, April 2018.

[Kny1986] A. Knyazev, *Computation of eigenvalues and eigenvectors for mesh problems: algorithms and error estimates* (in Russian), PhD thesis, Department of Numerical Mathematics, USSR Academy of Sciences, Moscow, 1986.

[KS1991] A. V. Knyazev and A. L. Skorokhodov, *On exact estimates of the convergence rate of the steepest ascent method in the symmetric eigenvalue problem*, Linear Algebra Appl. **154/156** (1991), 245–257, DOI 10.1016/0024-3795(91)90379-B. MR1113145

[LF1979] Q. Li and K. Q. Feng, *On the largest eigenvalue of a graph* (Chinese), Acta Math. Appl. Sinica **2** (1979), no. 2, 167–175. MR549045

[Löw1934] K. Löwner, *Über monotone Matrixfunktionen* (German), Math. Z. **38** (1934), no. 1, 177–216, DOI 10.1007/BF01170633. MR1545446

[MD1989] A. K. Mukherjee and K. K. Datta, *Two new graph-theoretical methods for generation of eigenvectors of chemical graphs*, Proceedings of the Indian Academy of Sciences–Chemical Sciences, 101(6):499–517, Dec 1989.

[NTU1993] P. Nylen, T. Y. Tam, and F. Uhlig, *On the eigenvalues of principal submatrices of normal, Hermitian and symmetric matrices*, Linear and Multilinear Algebra **36** (1993), no. 1, 69–78, DOI 10.1080/03081089308818276. MR1308910

[Pai1971] C. C. Paige, *The computation of eigenvalues and eigenvectors of very large sparse matrices*, PhD thesis, London University Institute of Computer Science, 1971.

[Par1980] B. N. Parlett, *The symmetric eigenvalue problem*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1980. Prentice-Hall Series in Computational Mathematics. MR570116

[Š1969] G. E. Šilov, *Matematicheskiĭ analiz: Vtoroĭ spetsial'nyĭ kurs* (Russian), Izdat. "Nauka", Moscow, 1965. MR0219869

[Sta2019] M. Stawiska, *More on the eigenvectors-from-eigenvalues identity*, `https://arxiv.org/abs/1912.06967`, 2019.

[Tho1966] R. C. Thompson, *Principal submatrices of normal and Hermitian matrices*, Illinois J. Math. **10** (1966), 296–308. MR190151

[Tho1969] R. C. Thompson, *Principal submatrices. IV. On the independence of the eigenvalues of different principal submatrices*, Linear Algebra Appl. **2** (1969), 355–374, DOI 10.1016/0024-3795(69)90037-8. MR245601

[TM1968] R. C. Thompson and P. McEnteggert, *Principal submatrices. II. The upper and lower quadratic inequalities*, Linear Algebra Appl. **1** (1968), 211–243, DOI 10.1016/0024-3795(68)90005-0. MR237532

[Tos1991] S. Toshev, *On T violation in matter neutrino oscillations*, Mod. Phys. Lett., A6:455–460, 1991.

[TV2011] T. Tao and V. Vu, *Random matrices: universality of local eigenvalue statistics*, Acta Math. **206** (2011), no. 1, 127–204, DOI 10.1007/s11511-011-0061-3. MR2784665

[Van2014] P. Van Mieghem, *Graph eigenvectors, fundamental weights and centrality metrics for nodes in networks*, arXiv e-prints, page `arXiv:1401.4580`, Jan 2014.

[VM2011] P. Van Mieghem, *Graph spectra for complex networks*, Cambridge University Press, Cambridge, 2011. MR2767173

[Wei1960] H. F. Weinberger, *Error bounds in the Rayleigh–Ritz approximation of eigenvectors*, J. Res. Nat. Bur. Standards Sect. B **64B** (1960), 217–225. MR129121

[Wil1963] J. H. Wilkinson, *Rounding errors in algebraic processes*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963. MR0161456

[Wol2019] N. Wolchover, *Neutrinos lead to unexpected discovery in basic math*, Quanta Magazine, Nov 2019.

[Xu1995] S. Xu, *Theories and Methods of Matrix Calculations* (in Chinese), Peking University Press, Beijing, 1995.

Department of Physics, Brookhaven National Laboratory, Upton, New York 11973
Email address: `pdenton@bnl.gov`

Theoretical Physics Department, Fermi National Accelerator Laboratory, Batavia, Illinois 60510
Email address: `parke@fnal.gov`

Department of Mathematics, University of California, Los Angeles, Los Angeles California 90095-1555
Email address: `tao@math.ucla.edu`

Enrico Fermi Institute & Department of Physics, University of Chicago, Chicago, Illinois 60637
Email address: `xining@uchicago.edu`