A remark on the mathematics of the seesaw mechanism

To cite this article: Fabien Besnard 2017 J. Phys. Commun. 1 015005

View the article online for updates and enhancements.

Related content

- New results for exponential synchronization of linearly coupled ordinary differential systems
  Ping Tong and Shi-Hua Chen

- Flexible sparse regularization
  Dirk A Lorenz and Elena Resmerita

- String-averaging incremental subgradients for constrained convex optimization with applications to reconstruction of tomographic images
  Rafael Massambone de Oliveira, Elias Salomão Helou and Eduardo Fontoura Costa
A remark on the mathematics of the seesaw mechanism

Fabien Besnard
Pôle de recherche M.L. Paris, EPF, 3 bis rue Lakanal, F-92330 Sceaux, France
E-mail: fabien.besnard@epf.fr

Keywords: seesaw mechanism, matrices of seesaw type, singular values estimates

Abstract
To demonstrate that matrices of seesaw type lead to a hierarchy in the neutrino masses, i.e. that there is a large gap in the singular spectrum of these matrices, one generally uses an approximate block-diagonalization procedure. In this note we show that no approximation is required to prove this gap property if the Courant–Fisher–Weyl theorem is used instead. This simple observation might not be original, however it does not seem to show up in the literature. We also sketch the proof of additional inequalities for the singular values of matrices of seesaw type.

1. Introduction

The terms in the Standard Model Lagrangian\(^1\) giving mass to the neutrinos can be gathered in a matrix of the general form

\[
M_\nu = \begin{pmatrix} m_L & m_D \\ m_D^\dagger & m_R \end{pmatrix}
\]

(1)

which has to be symmetric and where each entry is a complex \(3 \times 3\) matrix acting on the generation (or flavor) space. The requirement of renormalizability gives the further constraint that \(m_L = 0\). (Note that the noncommutative geometry approach to the Standard Model naturally predicts a matrix of this type with \(m_L = 0\) without any consideration of renormalizability, see [1, 2].)

The neutrino masses are the singular values of \(M_\nu\), that is to say the eigenvalues of the positive definite matrix \(M_\nu^* M_\nu\), where the star means matrix adjoint.

To explain the smallness of the observed neutrino masses it is generally argued that if \(m_D\) is small with respect to \(m_R\), then the singular values of \(M_\nu\) split into two families: one very small, and one large (of the order of \(m_R\)). This is the seesaw mechanism. It is easy to show explicitly for one generation, since in that case \(M_\nu\) is a \(2 \times 2\) matrix. Moreover, all the phases can be put into \(M_R\) so that \(m_D\) is a real number and \(m_R := |M_R|\) is the single singular value of \(M_R\). It is then found that the smallest singular value \(m_1\) of \(M_\nu\) satisfies

\[
\frac{m_1^1}{m_R} \approx \left( \frac{m_D}{m_R} \right)^2
\]

at second order in \(\frac{m_D}{m_R}\) while the largest singular value \(m_2\) is approximately equal to \(m_R\). When the matrix is real, one can explicitly compute its eigenvalues and find, using the properties of the square root function, that

\[
\frac{m_1^1}{m_R} \leq \left( \frac{m_D}{m_R} \right)^2, \quad \frac{m_1^2}{m_R} > 1.
\]

We call this ‘the gap property’. For three generations though, \(M_\nu\) cannot be diagonalized by an analytical formula and one appeals to an approximate block diagonalization (see for instance [6] or the appendix of [5]),

\(^1\) Extended with right-handed neutrinos.
that is to say that $M_r$ is brought to block-diagonal form thanks to approximately unitary matrices. One can then prove the gap property up to higher order terms.

Our purpose here is just to make the simple observation that if one uses the Courant–Fischer–Weyl theorem then no approximation is needed to prove the gap property for the singular values for an arbitrary number of generations, in the form of exact inequalities like (3).

More precisely, let $M_r$ be the symmetric complex $2n \times 2n$ matrix

$$M_r = \begin{pmatrix} 0 & m_D \\ m_D & M_T \end{pmatrix}$$

(4)

where $n$ is the number of generations. Let $m_1^1, \ldots, m_1^{2n}$ be the singular values of $M_r$, written in ascending order. (Hence if $n = 3$, $m_1^n \leq m_2^n \leq m_3^n$ are the masses of the 3 light neutrinos at tree level, and $m_1^4 \leq m_2^4 \leq m_3^4$ are the masses of the 3 heavy ones.) Let $m_D^1 \leq \ldots \leq m_D^n$ be the singular values of $m_D$ (Dirac masses) and $m_R^1 \leq \ldots \leq m_R^n$ be the singular values of $M_R$ (Majorana masses). We suppose that

$$m_D^n < m_R^1.$$ (5)

This is our way of saying that $m_D$ is smaller than $M_R$, but note that it does not have to be much smaller. Then we will show that

$$m_1^n \leq \ldots \leq m_n^n \leq \frac{(m_D^n)^2}{\sqrt{(m_D^n)^2 + (m_R^1)^2}},$$

$$m_{n+1}^1 \geq \sqrt{(m_R^1)^2 + (m_D^n)^2}$$

(6)

which immediately entails

$$m_1^1 \leq \ldots \leq m_1^n \leq \frac{(m_D^n)^2}{m_R^1} \leq m_1^1 < m_1^{n+1} \leq \ldots \leq m_1^{2n}$$

(7)

which directly generalizes (3) to $n$ generations.

It is our hope that these exact formulas can be of some use to physicists. Here are some motivations for this hope:

1. In the usual approximate method one reasons on the order of magnitude of the entries of the matrices $m_D$ and $M_R$. Under the assumption that the latter are much larger, it is then found that $M_r$ can be approximately block-diagonalized to $\begin{pmatrix} M_l & 0 \\ 0 & M_T \end{pmatrix}$, where $M_l$ is supposed to contain the light masses and $M_T$ the heavy ones (see [6], paragraph IIB). One finds

$$M_l \approx -m_D M_T^{-1} m_D,$$

(8)

$$M_T \approx M_T$$

(9)

([6], equations (13) and (14)). However a matrix with large entries can have very small (even vanishing) singular values. For a specific example with two generations one can take for $m_D$, the $2 \times 2$ identity matrix and $M_T = \Lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 + \epsilon \end{pmatrix}$, where $\Lambda$ is large and $\epsilon$ small. For $\epsilon \rightarrow 0$, and for any value of $\Lambda$, one eigenvalue of $M_T^{-1}$ goes to infinity, so that it no longer makes sense to argue from (8) that the masses extracted from $M_l$ are small (the authors of [6] are well-aware of this fact and note that ‘specific structures in the mass matrix’ can spoil the reasoning). Our method could thus be used to see what are the most general relations to be expected among the singular values without getting our hands dirty by delving into the algebraic relations satisfied by the matrices.

2. This method is also fairly general (in particular it does not depend on any ansatz as the approximate block-diagonalization does). It might be useful in other contexts where the ratio of the entries of $m_D$, on those of $M_R$, though smaller than one, is not so small as to completely neglect all the multiplicative constants introduced in every step of the approximation.

The paper is organized as follows: in section 2 we recall the necessary mathematical background keeping it to the minimum required to prove the gap property in section 3. In section 4 we sketch the proof of additional inequalities for the singular values thanks to immediate generalizations of the formulas in section 2.

2. Min-max theorem and matrix inequalities

We recall here the following theorem.
Theorem 1 (Courant–Fischer–Weyl min-max theorem). Let $M$ be a self-adjoint $N \times N$ matrix with eigenvalues $m_1 \leq \ldots \leq m_N$. Then:

$$m_k = \min (\max \{|\langle MX, X \rangle| X \in W, \|X\| = 1\})$$

where $W$ runs over all vector subspaces of $\mathbb{C}^N$ of dimension $k$, and

$$m_k = \max (\min \{|\langle MX, X \rangle| X \in W, \|X\| = 1\})$$

where $W$ runs over all vector subspaces of $\mathbb{C}^N$ of dimension $N - k + 1$.

This yields the following well-known corollary (see for instance [3], example 7.5.2) that we will need. For a self-adjoint matrix $M$ let us write $m^{\nu}$ for the smallest eigenvalue of $M$.

Corollary 1. Let $A$, $B$ be two self-adjoint $N \times N$ matrices. Then

$$\min(A + B) \geq \min(A) + \min(B).$$

Thanks to the min-max theorem one can also easily show the following interlacing property (called the Cauchy interlacing theorem): let $Q$ be a submatrix of $M$ obtained by orthogonal projection on a vector subspace generated by $n$ basis vectors. Let $q_1 \leq \ldots \leq q_n$ be the eigenvalues of $Q$. Then

$$m_k \leq q_k \leq m_{N-n+k}$$

for every $k \leq n$. In the main part of this paper we will only need the special case where $N = 2n$ and $k = 1$, yielding

$$\min(Q) \leq m_{n+1}.$$  \hspace{1cm} (11)

We will also need the following lemma:

Lemma 1. For any $A \geq 0$ and any $B \in M_n(\mathbb{C})$ one has

$$\min(B^*AB) \geq \min(A) \min(B^*B).$$

We prove the lemma. It is obvious when $B$ is singular. We then suppose that it is not. Let $X$ be a unit vector. We have:

$$\langle B^*ABX, X \rangle = \langle ABX, BX \rangle = \left( A \frac{BX}{\|BX\|}, \frac{BX}{\|BX\|} \right) \|BX\|^2 \geq \min(A) \|BX\|^2,$$

by the min-max theorem.  \hspace{1cm} (12)

Now $\|BX\|^2 = \langle B^*BX, X \rangle \geq \min(B^*B)$ also by the min-max theorem. Since $\min(A) \geq 0$ one gets $\langle B^*ABX, X \rangle \geq \min(A) \min(B^*B)$, from which the results follows using the min-max theorem again.

3. Singular value estimates for matrices of seesaw type

In the proof we suppose that $m_D$ is not singular (hence $m_D^1 > 0$). Once the inequalities have been proved in this case, they can be generalized to singular $m_D$ using the continuity of the map assigning to a matrix its singular value multiplet.

From (4) we compute

$$M_\nu M_\nu^* = \begin{pmatrix} ? & \sqrt{m_D^1} M_R^\nu \end{pmatrix} \begin{pmatrix} ?^* & \sqrt{m_D^1} M_R^\nu^* \end{pmatrix}$$

where the question marks stand for matrices we do not care about. We call $Q = m_D M_D^* + M_R M_R^*$. From (11) we get:

$$\min(Q) \leq (m^{n+1}_\nu)^2.$$  \hspace{1cm} (13)

But $\min(Q) \geq \min(M_R M_R^*) + \min(m_D M_D^*) = (m_R^1)^2 + (m_D^1)^2$ which yields the second part of (6).

To prove the first part we first need to write down the inverse of $M_\nu$. There exists a general formula for inverting $2 \times 2$ block matrices. Here we can check by direct computation that
$$M^{-1} = \begin{pmatrix} m^{-1} & m^{-1} & m^{-1} \\ m^{-1} & 0 & \end{pmatrix} \quad (15)$$

We then see that
$$\begin{pmatrix} X \end{pmatrix}^n = \begin{pmatrix} m \end{pmatrix} m^{-1} m^{n-1} (m^{-1})^* M_R^* (m^{-1})^* + (m^{-1})^* m^{-1}.$$  

Using (11) again we obtain
$$\min (X) \leq m_{n+1},$$
where this time $m_{n+1}$ is $n+1$ largest eigenvalue of $(M_{*}^n M_R^{-1})^{-1}$, that is to say $m_{n+1} = (m_{*}^n)^{-1}$. Hence
$$\min (m^{-1} M_R M^{-1} (m^{-1})^* M_R^* (m^{-1})^*) + \min ((m^{-1})^* m^{-1}) \leq \frac{1}{(m_{*}^n)^2}.$$  

Now using the lemma twice we obtain
$$\min (m^{-1} M_R M^{-1} (m^{-1})^* M_R^* (m^{-1})^*) \geq \min (M_R M^{-1} (m^{-1})^* M_R^* (m^{-1})^*) \min (m^{-1} (m^{-1})^*) \geq \min ((m^{-1})^* m^{-1}) \min (m_{*}^n) \min ((m^{-1})^*) \geq \frac{(m_{*}^n)^2}{(m_{*}^n)^2}.$$  

We thus have
$$\frac{(m_{*}^n)^2}{(m_{*}^n)^2} + \frac{1}{(m_{*}^n)^2} \leq \frac{1}{(m_{*}^n)^2},$$
which easily yields the first part of (6).

4. Additional inequalities

We now sketch the proof of the following inequalities:
$$m_{n+k}^n \geq \sqrt{(m_{*}^n)^2 + (m_{*}^k)^2} \quad (18)$$
for $k = 1, \ldots, n$ and
$$m_{j}^n \leq \frac{m_{*}^n m_{*}^j}{\sqrt{(m_{*}^j)^2 + (m_{*}^k)^2}} \quad (19)$$
for $j = 1, \ldots, n$.

For this we will need to strengthen corollary 1 and lemma 1. The first strengthening is given by Weyl’s inequalities: if $A$ and $B$ are hermitian $n \times n$ matrices, and $C = A + B$, then for $1 \leq k \leq n$ one has
$$a_k + b_1 \leq c_k \leq a_k + b_n,$$
where $c_1 \leq \ldots \leq c_k, a_1 \leq \ldots \leq a_n$ and $b_1 \leq \ldots \leq b_n$ are the eigenvalues of $A, B$ and $C$.

As for lemma 1, we can extend it in the following way.

**Lemma 2.** For any $A \geq 0, B \in M_n(\mathbb{C})$ let $C = B^* A B$. Then one has (with the same notations as above)
$$c_k \geq a_k \min (B^* B).$$

The proof of this lemma follows the same line as the one of lemma 1. Suppose $B$ is non singular and let $W$ be a subspace of $\mathbb{C}^n$ of dimension $k$. Then $BW$ has dimension $k$ and must intersect the orthogonal of the subspace generated by the $k$ $1$ eigenvector of $A$ corresponding to $a_1, \ldots, a_{k-1}$. Hence $\langle \text{ABX, BX} \rangle \geq a_k \|BX\|^2$ on $W$. The result follows from the minmax theorem.

The inequalities (18) and (19) can then be proven by the same techniques as in the previous section.

**Acknowledgments**

The author wishes to thank the anonymous referees for helping him improve the original version of the text.
ORCID iDs

Fabien Besnard  https://orcid.org/0000-0003-2643-6980

References

[1] Boyle L and Farnsworth S 2014 Non-commutative geometry, non-associative geometry and the standard model of particle physics New J. Phys. 16 123027
[2] Brouder C, Bizi N and Besnard F 2015 The Standard Model as an extension of the noncommutative algebra of forms arXiv:1504.03890
[3] Meyer C 2000 Matrix Analysis and Applied Linear Algebra (Philadelphia: SIAM)
[4] Mohapatra R N and Smirnov A Y 2006 Neutrino mass and new physics Annu. Rev. Nucl. Part. Sci. 56 569–628
[5] Sogami I S and Shinohara T 1991 Universal seesaw mechanism for quarks and leptons Prog. Theor. Phys. 86 1031–51
[6] Lindner M, Ohlsson T and Seidl G 2002 See-saw mechanisms for Dirac and Majorana neutrino masses Phys. Rev. D 65 053014