Matrices of seesaw type and the Courant-Fischer-Weyl theorem

Fabien Besnard

December 17, 2015

Abstract

To demonstrate that matrices of seesaw type lead to a hieararchy 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. We do not claim originality for this observation which however 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 (see for instance [3])

$$M_\nu = \begin{pmatrix} m_L & {}^t m_D \\ m_D & M_R \end{pmatrix} \tag{1}$$

which has to be symmetric and where each entry is a complex 3×3 matrix acting on the generation (or flavor) space. The requirement of renormalizability gives the further contraint 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 ([1], [2]).

The neutrino masses are the singular values of M_ν, that is to say the eigenvalues of the positive definite matrix $\sqrt{M_\nu^* M_\nu}$, 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_ν 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_ν is a 2×2 matrix. It is then found that the smallest singular value m_ν^1 of M_ν satisfies

[1] extended with right-handed neutrinos

$$\frac{m_\nu^1}{m_R} \approx \left(\frac{m_D}{m_R}\right)^2 \tag{2}$$

at second order in $\frac{m_D}{m_R}$ while the largest singular value m_ν^2 is approximately equal to m_R. With just a little more work one finds in fact that (with m_D and m_R real)

$$\begin{aligned} \frac{m_\nu^1}{m_R} &\leq \left(\frac{m_D}{m_R}\right)^2 \\ \frac{m_\nu^2}{m_R} &> 1 \end{aligned} \tag{3}$$

We call this "the gap property". For three generations though, M_ν cannot be diagonalized by an analytical formula and one appeals to an approximate block diagonalization (see for instance [5] or the appendix of [4]), that is to say that M_ν 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_ν be the symmetric complex $2n \times 2n$ matrix

$$M_\nu = \begin{pmatrix} 0 & {}^t m_D \\ m_D & M_R \end{pmatrix} \tag{4}$$

where n is the number of generations. Let $m_\nu^1, \ldots, m_\nu^{2n}$ be the singular values of M_ν written in ascending order. (Hence if $n = 3$, $m_\nu^1 \leq m_\nu^2 \leq m_\nu^3$ are the masses of the 3 light neutrinos at tree level, and $m_\nu^4 \leq m_\nu^5 \leq m_\nu^6$ 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 further suppose that m_D is not singular (hence $m_D^1 > 0$) and that $m_D^n < m_R^1$. Then we will show that

$$\begin{aligned} m_\nu^n &\leq \frac{(m_D^n)^2}{\sqrt{(m_D^n)^2 + (m_R^1)^2}} \\ m_\nu^{n+1} &\geq \sqrt{(m_R^1)^2 + (m_D^1)^2} \end{aligned} \tag{5}$$

which immediately entails

$$m_\nu^1 \leq \ldots \leq m_\nu^n < \frac{(m_D^n)^2}{m_R^1} < m_R^1 < m_\nu^{n+1} \leq \ldots \leq m_\nu^{2n} \tag{6}$$

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. However a matrix with large entries can have very small (even vanishing) singular values. The method exposed here could 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-adoint $N \times N$ matrix with eigenvalues $m_1 \leq \ldots \leq m_N$. Then:*

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

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

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

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

This yields the following well-known corollary that we will need. For a self-adjoint matrix M let us write $\min(M)$ 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

$$m_k \leq q_k \leq m_{N-n+k} \tag{7}$$

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

$$\min(Q) \leq m_{n+1} \tag{8}$$

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:

$$\begin{aligned} \langle B^* ABX, X \rangle &= \langle ABX, BX \rangle \\ &= \langle A \frac{BX}{\|BX\|}, \frac{BX}{\|BX\|} \rangle \|BX\|^2 \\ &\geq \min(A) \|BX\|^2, \text{ by the min-max theorem} \end{aligned} \tag{9}$$

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 $\langle B^* ABX, X \rangle \geq \min(A) \min(B^* B)$, frow which the results follows using the min-max theorem again.

3 Singular value estimates for matrices of see-saw type

From (4) we compute

$$M_\nu M_\nu^* = \begin{pmatrix} ? & ? \\ ? & m_D {m_D}^* + M_R M_R^* \end{pmatrix} \tag{10}$$

where the question marks stand for matrices we do not care about. We call $Q = m_D {m_D}^* + M_R M_R^*$. From (8) we get:

$$\min(Q) \leq m_{n+1} \tag{11}$$

where $m_{n+1} = (m_\nu^{n+1})^2$. 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 (5).

To prove the first part we first need to write down the inverse of M_ν. There exists a general formula for inverting 2×2 block matrices. Here we can check by direct computation that

$$M_\nu^{-1} = \begin{pmatrix} -m_D^{-1} M_R \, {}^t m_D^{-1} & m_D^{-1} \\ {}^t m_D^{-1} & 0 \end{pmatrix} \tag{12}$$

We then see that $(M_\nu^* M_\nu)^{-1} = \begin{pmatrix} X & ? \\ ? & ? \end{pmatrix}$, where

$$X = m_D^{-1} M_R\, {}^t m_D^{-1} ({}^t m_D^{-1})^* M_R^* (m_D^{-1})^* + (m_D^* m_D)^{-1} \tag{13}$$

Using (8) again we obtain

$$\min(X) \leq m_{n+1}$$

where this time m_{n+1} is $n+1$-th largest eigenvalue of $(M_\nu^* M_\nu)^{-1}$, that is to say $m_{n+1} = (m_\nu^n)^{-2}$. Hence

$$\min(m_D^{-1} M_R\, {}^t m_D^{-1} ({}^t m_D^{-1})^* M_R^* (m_D^{-1})^*) + \min((m_D^* m_D)^{-1}) \leq \frac{1}{(m_\nu^n)^2}$$

Now using the lemma twice we obtain

$$\begin{aligned}
\min(m_D^{-1} M_R\, {}^t m_D^{-1} ({}^t m_D^{-1})^* M_R^* (m_D^{-1})^*) &\geq \min(M_R\, {}^t m_D^{-1} ({}^t m_D^{-1})^* M_R^*) \min(m_D^{-1} (m_D^{-1})^*) \\
&\geq \min({}^t(m_D m_D^*)^{-1}) \min(M_R M_R^*) \min((m_D^* m_D)^{-1}) \\
&\geq \min((m_D m_D^*)^{-1}) \min(M_R M_R^*) \min((m_D^* m_D)^{-1}) \\
&\geq \frac{(m_R^1)^2}{(m_D^n)^4}
\end{aligned} \tag{14}$$

We thus have
$$\frac{(m_R^1)^2}{(m_D^n)^4} + \frac{1}{(m_D^n)^2} \leq \frac{1}{(m_\nu^n)^2}$$

which easily yields the first part of (5).

4 Additional inequalities

We now sketch the proof of the following inequalities:

$$m_\nu^{n+k} \geq \sqrt{(m_D^1)^2 + (m_R^k)^2} \tag{15}$$

for $k = 1, \ldots, n$ and

$$m_\nu^j \leq \frac{m_D^n m_D^j}{\sqrt{(m_D^j)^2 + (m_R^1)^2}} \tag{16}$$

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

$$a_k + b_1 \leq c_k \leq a_k + b_n \tag{17}$$

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 $C = B^*AB$. 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 $\langle ABX, BX \rangle \geq a_k \|BX\|^2$ on W. The result follows from the minmax theorem.

The inequalities (15) and (16) can then be proven by the same techniques as in the previous section.

References

[1] L. Boyle and S. Farnsworth, *Non-commutative geometry, non-associative geometry and the standard model of particle physics*, New Journal of Physics, vol. 16 (2014)

[2] C. Brouder, N. Bizi, and F. Besnard, *The Standard Model as an extension of the noncommutative algebra of forms*, arXiv:1504.03890

[3] R.N. Mohapatra and A.Y. Smirnov, *Neutrino mass and new physics*, Annu. Rev. Nucl. Part. Sci. 56:569-628 (2006)

[4] I. S. Sogami and T. Shinohara, *Universal Seesaw Mechanism for Quarks and Leptons*, Prog. Theor. Phys., Vol. 86, No.5, (1991)

[5] M. Lindner, T. Ohlsson, and G. Seidl, *See-saw Mechanisms for Dirac and Majorana Neutrino Masses*, Physical Review D, 2002