On existence of a biorthonormal basis composed of eigenvectors of non-Hermitian operators

2006 J. Phys. A: Math. Gen. 39 7757

(http://iopscience.iop.org/0305-4470/39/24/012)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **39** (2006) 7757–7761 doi:10.1088/0305-4470/39/24/012

On existence of a biorthonormal basis composed of eigenvectors of non-Hermitian operators

Toshiaki Tanaka

Department of Physics, Tamkang University, Tamsui 25137, Taiwan, Republic of China

E-mail: ttanaka@mail.tku.edu.tw

Received 4 April 2006
Published 31 May 2006
Online at stacks.iop.org/JPhysA/39/7757

Abstract
We present a set of necessary conditions for the existence of a biorthonormal basis composed of eigenvectors of non-Hermitian operators. As an illustration, we examine these conditions in the case of normal operators. We also provide a generalization of the conditions which is applicable to non-diagonalizable operators by considering not only eigenvectors but also all root vectors.

PACS numbers: 02.30.Sa, 02.30.Tb, 03.65.Db

Self-adjointness (or Hermiticity[1]) of operators corresponding to physical observables is well known as one of the postulates of quantum theory. Nevertheless, non-Hermitian Hamiltonians have been often employed in various areas of physical applications. This is mostly due to the fact that the self-adjointness is too restrictive for practical needs to describe physical systems which have in particular dissipative and unstable nature. Several years ago, a new explosion of research activities on non-Hermitian Hamiltonians was triggered by [1] where the authors introduced the new notion of $\mathcal{PT}$ symmetry. Later, Mostafazadeh attempted to reformulate $\mathcal{PT}$-symmetric theory within the concept of pseudo-Hermiticity [2] and derived various properties of pseudo-Hermitian theory mostly under the assumption of the existence of a biorthonormal basis composed of eigenvectors (or, root vectors) of operators; see [3] and references cited therein. In response to these works, there have appeared an increasing number of papers in which the same assumption has been employed.

To the best of our knowledge, the validity of the assumption was first questioned in [4] where the authors recalled an important caution for metric operators in quasi- and pseudo-Hermitian theories, which originated in [5] and also applies to $\mathcal{C}$ operators in $\mathcal{PT}$-symmetric theory [6, 7], though the paper has not been duly appreciated in the literature appeared after it until now. Then, it was independently pointed out in [8] that ascertaining whether the assumption is indeed satisfied by a given operator would be far from trivial.

[1] What physicists usually call *Hermitian* operators correspond to what mathematicians usually call *symmetric* operators, which is a less restrictive concept than *self-adjoint* operators.

A naive derivation of a biorthonormal system composed of eigenvectors of non-Hermitian operators already appeared in the classic book [9], and some attempts with much mathematical care dates back to (at the latest) the late 1960s [10, 11]. In [11], for instance, the authors considered such a Hamiltonian H which satisfies that (i) H is dissipative, (ii) H can be expressed as a sum of a self-adjoint (unperturbed) part H_0 and a bounded (interaction) part H_1, such that (iii) $H_1^{1/2}(H_0 - \lambda I)^{-1}H_1^{1/2}$ is compact for every regular point $\lambda \in \rho(H_0)$ of H_0.[2] Later, however, another naive derivation of a biorthonormal system under a quite loose condition reappeared [12]. As far as the list of the references indicates, it was this work that the author of [2] relied on for the existence of a 'complete biorthonormal basis'.

The aim of this paper is to present what kinds of conditions are necessary for the existence of a complete biorthonormal basis by clarifying subtleties which lurk behind the naive derivation in [12]. Although they would be rather well-known facts among mathematicians, they would not be duly recognized by physicists as the situation described above clearly indicates. For the later discussions, we first introduce the following notation:

$$\mathfrak{S}_\lambda(A) = \bigcup_{n=0}^{\infty} \mathrm{Ker}((A - \lambda I)^n), \tag{1}$$

$$\mathfrak{E}_0(A) = \overline{\langle \mathrm{Ker}(A - \lambda I) | \lambda \in \sigma_p(A)\rangle}, \tag{2}$$

$$\mathfrak{E}(A) = \overline{\langle \mathfrak{S}_\lambda(A) | \lambda \in \sigma_p(A)\rangle}. \tag{3}$$

That is, $\mathfrak{S}_\lambda(A)$ is the root subspace spanned by the root vectors of A belonging to the eigenvalue λ, and $\mathfrak{E}_0(A)$ (respectively $\mathfrak{E}(A)$) is the completion of the vector space spanned by all the eigenvectors (respectively root vectors) of the operator A. The quantities $m_\lambda^{(a)}(A) = \dim \mathfrak{S}_\lambda(A)$ and $m_\lambda^{(g)}(A) = \dim \mathrm{Ker}(A - \lambda I)$ are called the algebraic and geometric multiplicities of λ, respectively. By definition, $\mathfrak{E}(A) \supset \mathfrak{E}_0(A)$ and $m_\lambda^{(a)} \geqslant m_\lambda^{(g)}$. An eigenvalue is called *semi-simple* if $m_\lambda^{(a)} = m_\lambda^{(g)}$. Two subspaces $\mathfrak{L}_1$ and $\mathfrak{L}_2$ of a Hilbert space are said to be *skewly linked* and denoted by $\mathfrak{L}_1 \# \mathfrak{L}_2$ if $\mathfrak{L}_1 \cap \mathfrak{L}_2^\perp = \mathfrak{L}_1^\perp \cap \mathfrak{L}_2 = \{0\}$. With these preliminaries, we shall carefully reexamine the biorthonormal relation derived in [12].

The argument in [12] is as follows. Let λ_i and ψ_i be an eigenvalue and the corresponding eigenvector of a non-self-adjoint operator H in a Hilbert space $\mathfrak{H}$ equipped with an inner product $(\cdot, \cdot)$. Then λ_i^* belongs to the spectrum of $H^\dagger$, where $*$ and $\dagger$ denote complex conjugate and adjoint, respectively. So let χ_i be the eigenvector of $H^\dagger$ corresponding to λ_i^*, namely,

$$H\psi_i = \lambda_i \psi_i, \qquad H^\dagger \chi_i = \lambda_i^* \chi_i. \tag{4}$$

Considering the relation $(\chi_j, H\psi_i) = (H^\dagger \chi_j, \psi_i)$, we find $(\lambda_i - \lambda_j)(\chi_j, \psi_i) = 0$ and thus can choose the eigenvectors such that

$$(\chi_j, \psi_i) = (\psi_i, \chi_j) = \delta_{ij} \qquad \forall i, j. \tag{5}$$

Let f be an arbitrary vector belonging to the space defined by the complete set of states ψ_i or χ_i, namely, $f \in \mathfrak{E}_0(A)$ or $\mathfrak{E}_0(A^\dagger)$. We may expand

$$f = \sum_i A_i \psi_i \qquad \text{or} \qquad f = \sum_i B_i \chi_i, \tag{6}$$

where A_i, B_i are constants. From the biorthonormality (5) we easily find $A_i = (\chi_i, f)$ and $B_i = (\psi_i, f)$. Hence we obtain 'resolutions of the identity'

[2] See any elementary textbook on functional analysis for the precise definition of mathematical concepts such as dissipative (bounded, compact) operators, regular point, etc and the commonly used notation such as kernel Ker A, resolvent set $\rho(A)$, point spectrum $\sigma_p(A)$ of an operator A, etc appeared in this paper without explicit explanation.

$$\sum_i \psi_i(\chi_i, \cdot) = I\, \cdot \qquad \text{or} \qquad \sum_i \chi_i(\psi_i, \cdot) = I\, \cdot, \tag{7}$$

in terms of the biorthonormal vectors ψ_i and χ_i.

Although the above derivation may satisfy not a few physicists, it cannot be justified mathematically. In fact, to justify it rigorously, we must verify that (at least) the following conditions are all satisfied:

(1) The point spectra of A and $A^\dagger$ satisfy $\sigma_p(A^\dagger) = \sigma_p(A)^*$.

(2) Let $\Sigma(A)$ be the subset of $\sigma_p(A)$ such that $\mathrm{Ker}(A - \lambda I) \neq \mathrm{Ker}(A^\dagger - \lambda^* I)$ for all $\lambda \in \Sigma(A) \subset \sigma_p(A)$. Then for all $\lambda \in \Sigma(A)$, each pair of eigenspaces satisfies $\mathrm{Ker}(A - \lambda I) \# \mathrm{Ker}(A^\dagger - \lambda^* I)$.

(3) The geometric multiplicities satisfy $m_\lambda^{(g)}(A) = m_{\lambda^*}^{(g)}(A^\dagger)$ for all $\lambda \in \Sigma(A)$ (when neither of them is finite).

Explanations of each condition are in order. First of all, we must note the fact that although the spectra of A and $A^\dagger$ always satisfy the identity $\sigma(A^\dagger) = \sigma(A)^*$, it does not necessarily guarantee the first condition. This is because of the possible existence of the *residual* spectrum $\sigma_r(A)$ of A, which is defined by

$$\sigma_r(A) = \{\lambda \in \mathbb{C} | \mathrm{Ker}(A - \lambda I) = \{0\}, \overline{\mathfrak{R}(A - \lambda I)} \neq \mathfrak{H}\}, \tag{8}$$

where $\mathfrak{R}(A) \subset \mathfrak{H}$ denotes the range of A. In general, the relation

$$\sigma_r(A)^* \subset \sigma_p(A^\dagger) \subset \sigma_r(A)^* \cup \sigma_p(A)^* \tag{9}$$

follows. It is an immediate consequence of the identity $\overline{\mathfrak{R}(A - \lambda I)} \oplus \mathrm{Ker}(A^\dagger - \lambda^* I) = \mathfrak{H}$, which holds whenever $A^\dagger$ exists. Hence, for instance, some eigenvalue $\lambda^* \in \sigma_p(A^\dagger)$ can be related to a point of the residual spectrum $\lambda \in \sigma_r(A)$ for which no corresponding eigenvector ψ_i exists. A similar situation can take place between $\sigma_p(A)$ and $\sigma_r(A^\dagger)$. We note that in [2] the fulfillment of this first condition is also assumed by considering only operators with a discrete spectrum when a complete biorthonormal eigenbasis is introduced.

Next, we recall the fact that in the case of the ordinary orthonormality of eigenvectors $\{\psi_i\}$, e.g., of a self-adjoint operator, where each eigenvector χ_i in equations (4)–(7) is just $\chi_i \propto \psi_i$ for all i, the relation (5) for $i = j$ is guaranteed by the positive definiteness of the inner product. In our general case $\chi_i \not\propto \psi_i$, however, the inner product (5) for $i = j$ can take any finite complex number and in particular can be zero. Therefore, we cannot choose the eigenvectors such that relation (5) holds without ascertaining that there exist no vectors in $\mathrm{Ker}(A - \lambda I)$ (respectively in $\mathrm{Ker}(A^\dagger - \lambda^* I)$) which is orthogonal to $\mathrm{Ker}(A^\dagger - \lambda^* I)$ (respectively to $\mathrm{Ker}(A - \lambda I)$). This is the reason why the second condition is necessary.

Finally, the third condition must be satisfied since relations (4) and (5) indicate the one-to-one correspondence between the sets $\{\psi_i\}$ and $\{\chi_j\}$. If at least one of $m_\lambda^{(g)}(A)$ and $m_{\lambda^*}^{(g)}(A^\dagger)$ is finite, this condition is automatically satisfied under the fulfillment of the second condition [13]. In this case, the second condition is also sufficient [13] for the existence of a biorthonormal basis $\{\psi_{\lambda,i}, \chi_{\lambda,i}\}_1^{m_\lambda}$, satisfying $(\chi_{\lambda,j}, \psi_{\lambda,i}) = \delta_{ij}$ for all $i, j = 1, \ldots, m_\lambda$, *in each sector* $\lambda \in \Sigma(A)$, where $m_\lambda \equiv m_\lambda^{(g)}(A) = m_{\lambda^*}^{(g)}(A^\dagger) < \infty$, $\langle \psi_{\lambda,1}, \ldots, \psi_{\lambda,m_\lambda} \rangle = \mathrm{Ker}(A - \lambda I)$, and $\langle \chi_{\lambda,1}, \ldots, \chi_{\lambda,m_\lambda} \rangle = \mathrm{Ker}(A^\dagger - \lambda^* I)$.

We note that the above conditions are automatically satisfied in the case of finite-dimensional spaces. In fact, the use of a biorthonormal system is justified whenever the dimension of the space is finite. We also note that, when the subset $\Sigma(A)$ is empty, then we have $\mathrm{Ker}(A - \lambda_i I) = \mathrm{Ker}(A^\dagger - \lambda_i^* I)$ for all $\lambda_i \in \sigma_p(A)$ and thus can identify χ_i with ψ_i for all i. In this case, the biorthonormal relation (5) reduces to just the ordinary orthonormal one.

Another important point we should take care of, in addition to the above conditions, is that for a non-self-adjoint operator A the set of eigenvectors of either A or $A^\dagger$ does not generally span a dense subset of the whole Hilbert space $\mathfrak{H}$, namely, $\mathfrak{E}_0(A), \mathfrak{E}_0(A^\dagger) \subsetneq \mathfrak{H}$. Thus, even when the existence of biorthonormal system spanned by all eigenvectors of the operator under consideration is rigorously proved, it does not necessarily mean that the vector f which admits the expansion (6) can be an arbitrary vector of $\mathfrak{H}$ and that the 'resolutions of the identity' (7) are valid in a dense subset of $\mathfrak{H}$. Therefore, we cannot apply the relation like (7) in the whole Hilbert space unless the following additional condition is rigorously fulfilled:

(4) Each set of eigenvectors of A and $A^\dagger$ is complete in $\mathfrak{H}$, namely, $\mathfrak{E}_0(A) = \mathfrak{E}_0(A^\dagger) = \mathfrak{H}$.

In ordinary quantum theory, it is crucial that any state vectors in the Hilbert space L^2 can be expressed as a linear combination of a set of the eigenstates of the Hamiltonian or physical observables under consideration. However, this property, called *completeness*, is so frequently employed in vast areas of applications without any doubt that one may forget the fact that the completeness, as well as the absence of the residual spectrum, is guaranteed by the very property of self-adjointness of the operators.

As a simple illustration, we examine whether the above conditions are fulfilled by a normal operator A, namely, an operator which is closed, densely defined in $\mathfrak{H}$, and satisfies $A^\dagger A = AA^\dagger$. In finite-dimensional spaces the normality of operators (matrices) is the necessary and sufficient condition for the diagonalizability by a unitary transformation. In the general infinite-dimensional case, crucial properties of an arbitrary normal operator A in our context are as the follows (for details see, e.g., [13, 14]):

(a) $\mathrm{Ker}(A - \lambda I) \perp \mathrm{Ker}(A - \mu I)$ for all $\lambda \neq \mu$.
(b) Every eigenvalue is semi-simple.
(c) The residual spectrum is empty, $\sigma_r(A) = \emptyset$, and thus $\sigma_p(A^\dagger) = \sigma_p(A)^*$.
(d) The set of eigenvectors is complete, $\mathfrak{E}_0(A) = \mathfrak{E}_0(A^\dagger) = \mathfrak{H}$, so long as the spectrum $\sigma(A) (= \sigma(A^\dagger)^*)$ contains no more than a countable set of points of condensation.
(e) $\mathrm{Ker}(A - \lambda I) = \mathrm{Ker}(A^\dagger - \lambda^* I)$ for all $\lambda \in \sigma_p(A)$.

The first and second consequences (a) and (b) ensure the diagonalizability of every normal operator. By virtue of the properties (c) and (d), conditions 1 and 4 are (almost) always satisfied. The property (e) simply implies that the subset $\Sigma(A)$ defined in condition 2 is empty, $\Sigma(A) = \emptyset$. Thus it automatically guarantees conditions 2 and 3, and every normal operator A admits just an ordinary orthonormal basis composed of simultaneous eigenvectors of A and $A^\dagger$.

Later relation (7) was generalized to the case when non-Hermitian Hamiltonians have block diagonal structure [15]. It is a reasonable attempt since non-Hermitian (more generally and adequately, non-normal) operators in general admit non-semi-simple eigenvalues and thus anomalous Jordan cells. In this case, it is apparent by considering the explanations in the previous case that conditions 2–4 should be generalized to the following statements:

(2′) Let $\Sigma(A)$ be the subset of $\sigma_p(A)$ such that $\mathfrak{S}_\lambda(A) \neq \mathfrak{S}_{\lambda^*}(A^\dagger)$ for all $\lambda \in \Sigma(A) \subset \sigma_p(A)$. Then for all $\lambda \in \Sigma(A)$, each pair of root spaces satisfies $\mathfrak{S}_\lambda(A) \# \mathfrak{S}_{\lambda^*}(A^\dagger)$.
(3′) The algebraic multiplicities satisfy $m_\lambda^{(a)}(A) = m_{\lambda^*}^{(a)}(A^\dagger)$ for all $\lambda \in \Sigma(A)$ (when neither of them is finite).
(4′) Each set of root vectors of A and $A^\dagger$ is complete in $\mathfrak{H}$, namely, $\mathfrak{E}(A) = \mathfrak{E}(A^\dagger) = \mathfrak{H}$.

For non-normal operators acting in an infinite-dimensional Hilbert space, however, ascertaining these conditions 2′–4′ in addition to 1 is much more non-trivial and difficult. We should again note the fact that the general situation $\sigma_p(A^\dagger) \neq \sigma_p(A)^*$ for non-normal operators A persists.

To conclude, the assumption of the existence of a complete biorthonormal basis composed of eigenvectors or root vectors of an operator puts stronger conditions on the operator.

Therefore, the results obtained under this assumption, e.g., those of [2, 3, 12, 15], apply only to those operators that satisfy these conditions. We hope that this work would provide one step from naive discussions towards more fruitful and careful investigations in and around the research field. Mathematical characterization of a class of non-normal operators which certainly admit a complete biorthonormal basis composed of root vectors would be a challenging problem.

Acknowledgments

We would like to thank A Mostafazadeh for the valuable comment on the first version of this paper [16] and communication. We would also like to thank the anonymous referee for the constructive comments and suggestions. This work was partially supported by the National Science Council of the Republic of China under the grant no NSC-93-2112-M-032-009.

References

[1] Bender C M and Boettcher S 1998 Real spectra in non-Hermitian Hamiltonians having $\mathcal{PT}$ symmetry *Phys. Rev. Lett.* **80** 5243–6 (*Preprint* physics/9712001)

[2] Mostafazadeh A 2002 Pseudo-Hermiticity versus PT symmetry: the necessary condition for the reality of the spectrum of a non-Hermitian Hamiltonian *J. Math. Phys.* **43** 205–14 (*Preprint* math-ph/0107001)

[3] Mostafazadeh A 2005 $\mathcal{PT}$-symmetric cubic anharmonic oscillator as a physical model *J. Phys. A: Math. Gen.* **38** 6557–70

Mostafazadeh A 2005 *J. Phys. A: Math. Gen.* **38** 8185 (erratum) (*Preprint* quant-ph/0411137)

[4] Kretschmer R and Szymanowski L 2004 Quasi-Hermiticity in infinite-dimensional Hilbert spaces *Phys. Lett.* A **325** 112–7 (*Preprint* quant-ph/0305123)

[5] Scholtz F G, Geyer H B and Hahne F J W 1992 Quasi-Hermitian operators in quantum mechanics and the variational principle *Ann. Phys.* **213** 74–101

[6] Bender C M, Brody D C and Jones H F 2002 Complex extension of quantum mechanics *Phys. Rev. Lett.* **89** 270401

Bender C M, Brody D C and Jones H F 2004 *Phys. Rev. Lett.* **92** 119902 (erratum) (*Preprint* quant-ph/0208076)

[7] Bender C M, Brody D C and Jones H F 2004 Scalar quantum field theory with cubic interaction *Phys. Rev. Lett.* **93** 251601 (*Preprint* hep-th/0402011)

[8] González-López A and Tanaka T 2006 Nonlinear pseudo supersymmetry in the framework of $\mathcal{N}$-fold supersymmetry *J. Phys. A: Math. Gen.* **39** 3715–23 (*Preprint* quant-ph/0602177)

[9] Morse P M and Feshbach H 1953 *Methods of Theoretical Physics* (New York: McGraw-Hill)

[10] Fonda L, Ghirardi G C, Weber T and Rimini A 1966 Analytic properties of resolvents and completeness *J. Math. Phys.* **7** 1643–7

[11] Wong J 1967 Results on certain non-Hermitian Hamiltonians *J. Math. Phys.* **8** 2039–42

[12] Faisal F H M and Moloney J V 1981 Time-dependent theory of non-Hermitian Schrödinger equation: application to multiphoton-induced ionisation decay of atoms *J. Phys. B: At. Mol. Phys.* **14** 3603–20

[13] Azizov T Y and Iokhvidov I S 1989 *Linear Operators in Spaces with Indefinite Metric* (Chichester: Wiley)

[14] Birman M S and Solomjak M Z 1987 *Spectral Theory of Self-Adjoint Operators in Hilbert Space* (Dordrecht: Reidel)

[15] Mostafazadeh A 2002 Pseudo-Hermiticity for a class of nondiagonalizable Hamiltonians *J. Math. Phys.* **43** 6343–52

Mostafazadeh A 2003 *J. Math. Phys.* **44** 943 (erratum) (*Preprint* math-ph/0207009)

[16] Mostafazadeh A 2006 Comment on 'On existence of a biorthogonal basis composed of eigenvectors of non-Hermitian operators' *Preprint* quant-ph/0603096