An algorithmic test for diagonalizability of finite-dimensional PT-invariant systems

Stefan Weigert

Department of Mathematics, University of York, Heslington, York YO10 5DD, UK

E-mail: slow500@york.ac.uk

Abstract
A non-Hermitian operator does not necessarily have a complete set of eigenstates, contrary to a Hermitian one. An algorithm is presented which allows one to decide whether the eigenstates of a given PT-invariant operator on a finite-dimensional space are complete or not. In other words, the algorithm checks whether a given PT-symmetric matrix is diagonalizable. The procedure neither requires to calculate any single eigenvalue nor any numerical approximation.

PACS number: 03.67.−w

1. Introduction

The physical interpretation of PT-invariant operators—and hence their relevance for the description of physical systems—continues to be debated [1–3]. There is, however, no doubt about the cathartic role of PT-symmetry: it has become more evident what it means to let go hermiticity in exchange for a weaker property such as PT-invariance. The success and ease to describe quantum mechanical systems in terms of Hermitian operators is based on two of their generic properties, namely, the existence of *real* eigenvalues and their *diagonalizability*, i.e. the completeness of their orthonormal eigenstates. These properties do not necessarily persist if a quantum system was described by a PT-symmetric Hamiltonian: its eigenvalues could be complex, and its eigenfunctions would, in general, neither be pairwise orthogonal nor form a complete set. Given a PT-invariant operator, it thus appears desirable to decide whether it is diagonalizable or not.

The purpose of this contribution is to provide an algorithm answering the question of whether a given PT-invariant Hamiltonian operator in a finite-dimensional space *does* or *does not* possess a complete set of eigenstates. It is convenient to represent such an operator as a PT-symmetric matrix M, say. A procedure will be outlined which, after a finite number of steps, will announce whether the matrix M at hand is diagonalizable or not. In principle, the

algorithm can be carried out by hand for matrices of any dimension, and no approximations are necessary.

Often, the question of diagonalizability will arise in a more general setting where one considers not just a single matrix but a *family* of PT-symmetric matrices $M(\varepsilon)$, $\varepsilon \in \mathbb{R}$. The parameter ε measures the strength $\varepsilon \in \mathbb{R}$ of a 'perturbation' which destroys hermiticity while it respects PT-invariance. As the parameter varies, all of the cases described previously may occur: typically, two real eigenvalues merge into a single real one at a critical value of ε, subsequently splitting into a pair of two complex conjugate eigenvalues, or vice versa. These dramatic modifications are accompanied by changes in the nature of the eigenstates of the PT-invariant operator, possibly no longer spanning the space on which $M(\varepsilon)$ acts.

This behaviour can be understood in terms of so-called *exceptional points* [4] which are known to occur when a matrix is subjected to the perturbation depending analytically on a parameter such as ε. At an exceptional point, the corresponding matrix is not diagonalizable, and its spectrum may undergo a qualitative change. For a *Hermitian* operator subjected to a parameter-dependent *Hermitian* perturbation, exceptional points cannot occur.

If one applies the algorithm testing for diagonalizability to a parameter-dependent matrix $M(\varepsilon)$, it will output a polynomial in ε instead of a number. Its zeros correspond to those values of the perturbation parameter where the matrix family $M(\varepsilon)$ has exceptional points. The matrices corresponding to these values of the perturbation are not diagonalizable, and the spectra of matrices for nearby values of the parameter may differ qualitatively.

The following section summarizes the properties of PT-invariant systems in terms of (2×2) matrices. Then, the link between diagonalizability and the so-called *minimal polynomial* is reviewed. In section 3, the algorithmic test is presented which consists of constructing the minimal polynomial of the matrix followed by a search for degenerate roots by means of the Euclidean algorithm. Various methods are known to effectively calculate the minimal polynomial of a matrix, outlined in section 4. Simple examples are studied in section 5, leading to some general conclusions about the structure of PT-symmetric Hamiltonian operators in finite-dimensional spaces. Section 6 summarizes the results and discusses the challenge to extend them to state spaces of infinite dimension.

2. PT-invariant systems

A matrix H is PT-invariant [5],

$$[H, PT] = 0, \tag{1}$$

if it commutes with the product of parity P and the anti-unitary operation of time reversal T, represented here by complex conjugation, $T^{\dagger}HT = H^*$. Equation (1) implies that the characteristic polynomial of any PT-symmetric operator H has real coefficients only. Consequently, its roots are either real or come in complex-conjugate pairs. One way to show this is to construct a basis in which the Hamiltonian has real matrix elements only [6].

Let us briefly review the properties of PT-symmetric systems by considering the most general PT-invariant matrix of dimension 2,

$$H = \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix}, \qquad a, b \in \mathbb{C}, \tag{2}$$

with parity given by the Pauli matrix σ_x in the standard representation. For real numbers a and b, the matrix H is not only PT-invariant but also Hermitian. Thus, its eigenvalues are *real*, and

its orthonormal eigenstates span $\mathbb{C}^2$. For $a^* \neq a$ and $b = 0$, H has a pair of complex conjugate eigenvalues and two orthonormal eigenstates. Matrices of the form

$$H = \begin{pmatrix} i & b \\ b & -i \end{pmatrix}, \qquad b \in [-1, 1], \tag{3}$$

are particularly interesting. For $|b| < 1$, one finds a pair of two complex conjugate eigenvalues,

$$E_\pm = \pm\sqrt{b^2 - 1} \in i\mathbb{R}, \tag{4}$$

associated with two non-orthogonal eigenstates,

$$\frac{1}{\sqrt{2b}} \begin{pmatrix} b \\ i - \sqrt{b^2 - 1} \end{pmatrix}, \qquad \frac{1}{\sqrt{2b}} \begin{pmatrix} b \\ i + \sqrt{b^2 - 1} \end{pmatrix}. \tag{5}$$

When $b = \pm 1$ in (3), H has a two-fold degenerate eigenvalue, $E_0 = 0$, and there is only *one* eigenstate, namely,

$$\frac{1}{\sqrt{2}} \begin{pmatrix} \mp 1 \\ i \end{pmatrix}. \tag{6}$$

This situation, impossible for a Hermitian matrix, is usually described by saying that the *algebraic* multiplicity of the eigenvalue E_0 is 2 while its *geometric* multiplicity equals 1: the characteristic polynomial of M has a double root associated with a single eigenvector only. In this case, the matrix H is not diagonalizable: a similarity transformation sending it to a *diagonal* matrix cannot exist since its eigenstates *would* span the space $\mathbb{C}^2$.

3. Diagonalizability and the minimal polynomial of a matrix

Each square matrix M of dimension N satisfies the identity

$$p_M(M) = 0, \tag{7}$$

where $p_M(\lambda)$ is the *characteristic polynomial* of M,

$$p_M(\lambda) = \det(\lambda E - M), \tag{8}$$

with E being the unit matrix of dimension N. In other words, the characteristic polynomial of M *annihilates* the matrix M. The polynomial $p_M(\lambda)$ has degree N and it is a *monic* polynomial, that is, the coefficient multiplying the highest power of λ is equal to 1. Obviously, many other monic polynomials of *higher* degree also annihilate M: simply take $p_M^2(\lambda)$, $p_M^3(\lambda)$, ... It is less obvious, however, whether one can find polynomials of degree *less* than N which annihilate M. This, in fact, depends on the properties of the matrix M.

Define [7] the *minimal polynomial* of the matrix M as the monic polynomial $m_M(\lambda)$ of *least* degree which annihilates M:

$$m_M(M) = 0. \tag{9}$$

The minimal polynomial $m_M(\lambda)$ is unique [7], and its degree N_0 is less than or equal to the degree of the characteristic polynomial, $N_0 \leqslant N$. The minimal polynomial divides the characteristic polynomial without a remainder,

$$p_M(\lambda) = d_M(\lambda)m_M(\lambda), \tag{10}$$

where $d_M(\lambda)$ is a non-zero polynomial of degree less than N. The characteristic and the minimal polynomial of the matrix M coincide in the case where $d_M(\lambda) \equiv 1$.

In general, the minimal polynomial has ν_0 roots M_ν,

$$m_M(\lambda) = \prod_{\nu=1}^{\nu_0} (\lambda - M_\nu)^{\mu_\nu}, \qquad \nu_0 \leqslant N, \tag{11}$$

with multiplicities μ_ν summing to $N_0 = \mu_1 + \mu_2 + \cdots + \mu_{\nu_0}$. Here is the important property of the polynomial $m_M(\lambda)$: the matrix M is diagonalizable if and only if each root M_ν in (11) has multiplicity 1, $\mu_\nu \equiv 1$, $\nu = 1, \ldots, \nu_0$, that is,

$$m_M(\lambda) = \prod_{\nu=1}^{\nu_0}(\lambda - M_\nu), \quad \text{all } M_\nu \text{ distinct.} \tag{12}$$

No polynomial of degree less than $m_M(\lambda)$ annihilates the matrix M.

Let us illustrate the properties of minimal polynomials using low-dimensional matrices. Consider the matrix A with entries $(1, 1, 2)$ on the diagonal, and zero elsewhere. Its characteristic polynomial is given by

$$p_A(\lambda) = (\lambda - 1)^2(\lambda - 2), \tag{13}$$

while its minimal polynomial reads

$$m_A(\lambda) = (\lambda - 1)(\lambda - 2), \tag{14}$$

being of the form (11), with $N_0 = \nu_0 = 2$. This is easy to verify since $m_A(A) = A^2 - 3A + 2E = 0$ holds, while none of its factors annihilates M: both $(M - E)$ and $(M - 2E)$ are different from zero. Thus, the minimal polynomial divides the characteristic one, $p_A(\lambda) = (\lambda-1)m_A(\lambda)$, leading to $d_A(\lambda) = (\lambda - 1)$. Due to (14), the matrix A *is* diagonalizable—a correct but hardly surprising result since the matrix A has been diagonal from the outset.

Here is the instructive part of the example: consider the matrix

$$B = \begin{pmatrix} 1 & b & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \qquad b \in \mathbb{C}, \tag{15}$$

which is different from A as long as b is different from zero. The characteristic polynomial of B equals that of A but the matrix B must have a different minimal polynomial since $m_A(B) \neq 0$. It is not difficult to verify that no linear or quadratic polynomial annihilates B as long as $b \neq 0$. This implies that its minimal polynomial coincides with its characteristic polynomial,

$$p_B(\lambda) = m_B(\lambda), \qquad d_B(\lambda) \equiv 1. \tag{16}$$

Consequently, the minimal polynomial of B does *not* have the form specified in (12), and the matrix B is *not* similar to a diagonal matrix. Inspection shows that B indeed contains a (2×2) Jordan block for any non-zero value of b.

For PT-invariant matrices, both the polynomials $d_H(\lambda)$ and $m_H(\lambda)$ have real coefficients only, just as the characteristic polynomial. This will be shown once the function $d_M(\lambda)$ in (10) has been defined in general (cf section 4.2).

4. An algorithmic test for diagonalizability

A square matrix M of dimension N is diagonalizable if its minimal polynomial is a product of factors $(\lambda - M_\nu)$ with all numbers $M_\nu, \nu = 1, 2, \ldots, n \leqslant \nu_0$, distinct, as shown in equation (12). Consequently, to test for diagonalizability of a given matrix M, one needs to

 (i) find the minimal polynomial $m_M(\lambda)$ of the matrix M;
 (ii) determine whether the polynomial $m_M(\lambda)$ has single roots only.

To calculate numerically the roots of either the characteristic or the minimal polynomial is not a valid approach since, in general, the *exact* roots of a polynomial cannot be specified in a finite procedure. Any algorithmic implementation must generate answers to (i) and (ii) in a

finite number of steps. Note that even if the first step has been implemented, it is unlikely that the minimal polynomial will emerge in factorized form.

Interestingly, it is possible to construct the minimal polynomial of a matrix and to check for degenerate roots in a finite number of steps. In both cases one searches for common factors of polynomials, which is achieved algorithmically by the *Euclidean division algorithm*. These results seem to have been put together for the first time in [8] in order to decide algorithmically whether a given matrix is diagonalizable. As it stands, it could be applied to the non-Hermitian matrix governing the motion of two-coupled damped classical oscillators studied in [9].

In the following, a slightly simplified approach to the problem of diagonalizability is presented, adapted to matrices with PT-symmetry. Before implementing the steps (i) and (ii), the Euclidean algorithm for polynomials will be presented briefly to establish notation.

4.1. The Euclidean division algorithm for polynomials

Given two integer numbers $p_0 > p_1$, say, the Euclidean division algorithm outputs their greatest common divisor, denoted by $\gcd(p_0, p_1) \in \mathbb{N}_0$, after a finite number of steps. It works as follows: first, you need to express the larger number as q_1-fold multiple of the smaller number plus a remainder p_2,

$$p_0 = q_1 p_1 + p_2, \qquad q_1, p_2 \in \mathbb{N}_0, \qquad p_1 > p_2 \geqslant 0. \tag{17}$$

This relation implies that any common divisor of p_0 and p_1 divides p_2 as well, hence $\gcd(p_0, p_1) = \gcd(p_1, p_2)$. Thus, it is sufficient to search for the greatest common divisor of the pair (p_1, p_2). This can be achieved by increasing each index in (17) by one and feeding in the pair (p_1, p_2) instead of (p_0, p_1), etc. Since $p_0 > p_1$ and $p_1 > p_2$, the algorithm will stop after a finite number of iterations and produce a remainder equal to zero, $p_{k+1} = 0$, say. The *non-zero* remainder p_k generated in the penultimate step is the desired result, $\gcd(p_0, p_1) = p_k$. If $\gcd(p_0, p_1) = 1$, the numbers p_0 and p_1 are relatively prime, otherwise a common divisor different from one has been identified.

A polynomial in the variable λ can be written as a unique product of linear factors $(\lambda - \lambda_n)$ where the numbers $\lambda_n \in \mathbb{C}$ are its roots. This representation makes polynomials similar to integer numbers in some respects. The equivalent of the Euclidean algorithm, when applied to two polynomials, outputs their greatest common divisor, which is a polynomial itself. This result is based on the fact that any two polynomials $p_0(\lambda)$ and $p_1(\lambda)$, with $\deg p_0(\lambda) > \deg p_1(\lambda)$ are related by

$$p_0(\lambda) = q_1(\lambda) p_1(\lambda) + p_2(\lambda), \qquad \deg p_1(\lambda) > \deg p_2(\lambda) \geqslant 0, \tag{18}$$

which is the equivalent of (17). The polynomials $q_1(\lambda)$, with $\deg q_1(\lambda) = (\deg p_0(\lambda) - \deg p_1(\lambda))$, and hence $p_2(\lambda)$, are found from long division. If $p_0(\lambda)$ and $p_1(\lambda)$ have a common factor, then $p_2(\lambda)$ must have this factor as well. Thus, it is sufficient to search for $\gcd(p_1(\lambda), p_2(\lambda))$ instead of $\gcd(p_0(\lambda), p_1(\lambda))$ but the degrees of the polynomials involved have effectively been reduced. Consequently, this procedure can be repeated all over again and it halts once a *vanishing* remainder has been obtained, $p_{k+1}(\lambda) = 0$, say. Then, the greatest common factor of the polynomials $p_0(\lambda)$ and $p_1(\lambda)$ is given by the last non-zero remainder polynomial, $p_k(\lambda)$, calculated in the next-to-last application of the algorithm. If $\deg p_k(\lambda) = 0$, the initial polynomials are 'relatively prime,' otherwise their greatest common divisor is a polynomial of degree at least 1.

4.2. Step (i): Finding the minimal polynomial of a matrix

The function $d_M(\lambda)$ relates the minimal polynomial $m_M(\lambda)$ of the matrix M to its characteristic polynomial $p_M(\lambda)$ according to equation (10). Hence, the minimal polynomial associated with M is known once the characteristic polynomial and the function $d_M(\lambda)$ have been determined.

Two steps are required to construct the function $d_M(\lambda)$ [7]. First, you need to calculate the matrix $D_M = \mathrm{adj}\,(\lambda E - M)$, given by the transposed cofactors—or signed minors—of the matrix $(\lambda E - M)$. The adjoint of a matrix, C say, always exists, and it satisfies the relation

$$C \,\mathrm{adj}\, C = (\det C)E. \tag{19}$$

For $\det C \neq 0$, equation (19) leads to the familiar expression of the inverse matrix of C.

According to [7], the polynomial $d_M(\lambda)$ is given by the *greatest* (monic) *common divisor* of the N^2 elements of $\mathrm{adj}\,(\lambda E - M)$,

$$d_M(\lambda) = \gcd\{(D_M)_{nm}|n, m = 1, \ldots, N\}, \tag{20}$$

Thus, in a second step, you need to apply the Euclidean algorithm to all pairs of entries of the matrix $\mathrm{adj}\,(\lambda E - M)$. Having thus identified the function $d_M(\lambda)$, the minimum polynomial of M follows from (10),

$$m_M(\lambda) = \frac{p_M(\lambda)}{d_M(\lambda)}. \tag{21}$$

Now it is possible to show that, for a PT-invariant matrix, the polynomials $d_M(\lambda)$ and $m_M(\lambda)$ have real coefficients only, just as does the characteristic polynomial. Using a basis in which all elements of H are real, leads to

$$(D_H(\lambda))^* = (\mathrm{adj}(\lambda E - H))^* = \mathrm{adj}(\lambda^* E - H) = D_H(\lambda^*). \tag{22}$$

which states that in this basis the adjoint of H has only real matrix elements (except for the unknown λ). Taking (20) into account this leads to

$$(d_H(\lambda))^* = \gcd\{(D_H)_{nm}(\lambda^*)|n, m = 1, \ldots, N\} = d_H(\lambda^*), \tag{23}$$

which, in conjunction with $(p_H(\lambda))^* = p_H(\lambda^*)$ and equation (21) implies indeed $(m_H(\lambda))^* = m_H(\lambda^*)$.

Let us verify that this procedure outputs the correct minimal polynomials for the matrices A and B introduced in section 3. The adjoint of the matrix $(\lambda E - B)$ reads

$$\mathrm{adj}(\lambda E - B) = \mathrm{adj}\begin{pmatrix} \lambda - 1 & -b & 0 \\ 0 & \lambda - 1 & 0 \\ 0 & 0 & \lambda - 2 \end{pmatrix}$$
$$= \begin{pmatrix} (\lambda - 1)(\lambda - 2) & b(\lambda - 2) & 0 \\ 0 & (\lambda - 1)(\lambda - 2) & 0 \\ 0 & 0 & (\lambda - 1)^2 \end{pmatrix}. \tag{24}$$

Due to the simplicity of the matrices involved, the Euclidean algorithm can be run 'by inspection:' for $b \neq 0$, the only common factor among the entries in (24) is given by $d_B(\lambda) = 1$. Consequently, the minimal and the characteristic polynomial of B coincide as stated in equation (16). If the parameter b takes the value zero, B turns into A, and a non-constant greatest common divisor emerges, $d_A(\lambda) = (\lambda - 1)$. Using equation (10), one obtains the minimal polynomial $m_A(\lambda) = (\lambda - 1)(\lambda - 2)$, agreeing with equation (14).

In [8], a different approach to determine the minimal polynomial of a matrix M has been presented which, ultimately, is also based on finding the greatest common divisor of specific polynomials. According to [10], any method to determine whether the matrices

$M^0 \equiv E, M, M^2, \ldots, M^{N-1}$, are linearly dependent, can be used to construct the minimal polynomial of M; two such methods are described in this reference, and a third one can be found in [11]. The latter approaches have in common that they are *not* based on the Euclidean algorithm. For actual calculations, it is convenient to resort to a `Mathematica` program [12] to find the minimal polynomial of a matrix M.

4.3. Step (ii): Identifying degenerate roots of a polynomial

Once the minimal polynomial $m_M(\lambda)$ has been found, one needs an algorithm to decide whether it has single roots only [8]. Imagine a polynomial $m(\lambda)$ to have an s-fold root λ_0, $2 \leqslant s \leqslant N$. Its factorization reads

$$m(\lambda) = (\lambda - \lambda_0)^s \ldots, \tag{25}$$

where the dots indicate a polynomial of degree $(N - s)$. Its derivative takes the form

$$\frac{dm}{d\lambda} = (\lambda - \lambda_0)^{s-1} \ldots, \tag{26}$$

the dots standing again for some polynomial of degree $(N - s)$. Obviously, the polynomial and its derivative are not relatively prime: $m(\lambda)$ and $m'(\lambda)$ have a factor $(\lambda - \lambda_0)^{s-1}$ of order at least one in common. Thus, applying the division algorithm to the pair $(m_M(\lambda), m'_M(\lambda))$ checks whether the polynomial $m_M(\lambda)$ has the form (12). If the procedure outputs $\gcd(m_M(\lambda), m'_M(\lambda)) \propto 1$, all roots of m_M are distinct and the associated matrix M is diagonalizable, otherwise it is not.

This concludes the description of an algorithm to test for diagonalizability of a given PT-symmetric matrix M. No fundamental changes are necessary if one studies a parameter-dependent family of matrices $M(\varepsilon)$. However, the algorithm will output conditions polynomial in the parameter ε, indicating specific parameter values where diagonalizability breaks down. It is convenient to study the resulting modifications by working out some simple examples, illustrating at the same time the proposed algorithm.

5. Examples

5.1. Matrices of dimension (2×2)

Let us apply the algorithm described above to the matrix H in (2) assuming the numbers a and b to be different from zero. Its characteristic polynomial reads

$$p_H(\lambda) = \det(\lambda E - H) = \lambda^2 - 2(\text{Re}\,a)\lambda + |a|^2 - |b|^2, \tag{27}$$

while its minimal polynomial is found via the function $d_H(\lambda)$, given as the highest common factor of the entries of the matrix

$$D_H(\lambda) = \text{adj}(\lambda E - H) = \begin{pmatrix} \lambda - a^* & -b \\ -b^* & \lambda - a \end{pmatrix}. \tag{28}$$

By inspection, a non-constant factor only exists among the four entries of $D_H(\lambda)$ if $b = \text{Im}\,a = 0$. In this case, H turns into a real multiple of the identity, hence it is diagonalizable. This observation illustrates a fine point of the construction of the minimal polynomial: even upon identifying a non-constant function $d_H(\lambda)$, the minimal polynomial $m_H(\lambda)$ may still be of the form (12). For $b = \text{Im}\,a = 0$, the characteristic polynomial turns into $(\lambda - \text{Re}\,a)^2$, implying indeed $m_H(\lambda) = (\lambda - \text{Re}\,a)$. Here, the function $d_H(\lambda) = (\lambda - \text{Re}\,a)$ removes factors of the characteristic polynomial which stem from the *degeneracy* of an eigenvalue of H.

From now on, either b or $\operatorname{Im} a$ are assumed to be different from zero, hence $d_{\mathsf{H}}(\lambda) = 1$, and the minimal polynomial $m_{\mathsf{H}}(\lambda)$ is given by equation (27),

$$m_{\mathsf{H}}(\lambda) = p_{\mathsf{H}}(\lambda), \tag{29}$$

which concludes the first step of the algorithm.

In the second step of the algorithm, the search for multiple roots of $m_{\mathsf{H}}(\lambda) \equiv p_0(\lambda)$, one needs to determine the highest common factor of the minimal polynomial and its derivative,

$$m_{\mathsf{H}}'(\lambda) = 2(\lambda - \operatorname{Re} a) \equiv p_1(\lambda). \tag{30}$$

Applying the Euclidean algorithm to the pair $(m_{\mathsf{H}}(\lambda), m_{\mathsf{H}}'(\lambda))$ means to solve for a polynomial $q_1(\lambda) = A\lambda + B$ and for $p_2(\lambda)$, such that

$$m_{\mathsf{H}}(\lambda) = (A\lambda + B)m_{\mathsf{H}}'(\lambda) + p_2(\lambda), \qquad A, B \in \mathbb{R}. \tag{31}$$

The unknowns are easily obtained as

$$A = \tfrac{1}{2}, \qquad B = -\tfrac{1}{2}\operatorname{Re} a, \qquad p_2(\lambda) = (\operatorname{Im} a)^2 - |b|^2 \simeq \lambda^0. \tag{32}$$

Two possibilities now arise: either $p_2(\lambda)$ equals zero or it does not. The first case occurs if

$$(\operatorname{Im} a)^2 = |b|^2, \tag{33}$$

and the algorithm comes to a halt. As mentioned above, the greatest common divisor of the initial polynomials is then given by the penultimate (monic) remainder polynomial, i.e. $p_1(\lambda) = (\lambda - \operatorname{Re} a)$. It follows that $m_{\mathsf{H}}(\lambda)$ and its derivative do have a common non-constant divisor, so that H cannot be brought to diagonal form. It is easy to verify that $m_{\mathsf{H}}(\lambda) = (\lambda - \operatorname{Re} a)^2$ when (33) holds, confirming that the minimal polynomial of H has a *double* root $\operatorname{Re} a$. Furthermore, a simple calculation shows that the matrix H has indeed only one eigenstate if the relation $\operatorname{Im} a = \pm|b|$ holds. Note that this result covers the example of a non-diagonalizable H of section 2, where $a = \mathrm{i}$ and $b = \pm 1$ had been considered.

Finally, if a and b do not satisfy (33), the remainder polynomial $p_2(\lambda)$ does not vanish. Being a constant, the algorithm is bound to stop after the next iteration. Determine $q_2(\lambda) = (C\lambda + D)$ and $p_3(\lambda)$, such that

$$(\lambda - \operatorname{Re} a) = (C\lambda + D)((\operatorname{Im} a)^2 - |b|^2) + p_3(\lambda), \qquad C, D \in \mathbb{R}, \tag{34}$$

holds, i.e.,

$$C = \frac{1}{(\operatorname{Im} a)^2 - |b|^2}, \qquad D = \frac{-\operatorname{Re} a}{(\operatorname{Im} a)^2 - |b|^2}, \qquad p_3(\lambda) = 0. \tag{35}$$

The algorithm halts indeed due to $p_3(\lambda) = 0$, and the penultimate remainder is $p_2(\lambda) \propto 1$, indicating that the minimal polynomial does not have any degenerate roots, and H is diagonalizable.

In summary, the matrix H is diagonalizable for all parameter values except when $\operatorname{Im} a = \pm|b|$. In this case, the algebraic multiplicity of its eigenvalue is 2, while its geometric multiplicity is 1; otherwise the multiplicities both equal 2. It is important to note that these conclusions have been reached without determining the eigenvalues.

5.2. Matrices of dimension (4×4)

It is instructive to apply the algorithm to a PT-invariant matrix of dimension 4,

$$\mathsf{H} = \begin{pmatrix} \mathrm{i}\varepsilon & s & 0 & 0 \\ s & -\mathrm{i}\varepsilon & \delta & 0 \\ 0 & \delta & \mathrm{i}\varepsilon & s \\ 0 & 0 & s & -\mathrm{i}\varepsilon \end{pmatrix}, \qquad s, \delta > 0, \tag{36}$$

which depends on a perturbation parameter ε. As before, the action of T on a matrix effects complex conjugation of its entries, while P is now given by a (4×4) matrix with entries equal to 1 along its minor diagonal and zero elsewhere. Equation (1) is then readily verified. The characteristic polynomial of H is given by

$$p_{\mathsf{H}}(\lambda) = \lambda^4 + \alpha_\varepsilon \lambda^2 + \beta_\varepsilon, \tag{37}$$

with $\alpha_\varepsilon = 2\varepsilon^2 - 2s^2 - \delta^2$ and $\beta_\varepsilon = \varepsilon^4 - (2s^2 + \delta^2)\varepsilon^2 + s^4$. The minimal polynomial of H coincides with the characteristic one, $m_{\mathsf{H}}(\lambda) \equiv p_{\mathsf{H}}(\lambda)$, since the only common factor of the matrix elements of D_{H} is equal to 1, $d_{\mathsf{H}}(\lambda) = 1$. To see this, it is sufficient to calculate the two matrix elements $[\mathsf{D}_{\mathsf{H}}]_{14} = -s^2\delta$, and $[\mathsf{D}_{\mathsf{H}}]_{23} = -(\lambda^2 + \varepsilon^2)\delta$, for example. Whatever the value of ε, for non-zero s and δ the only common divisor is 1, so that $d_{\mathsf{H}}(\lambda) = 1$.

Let us now determine $\gcd(m_{\mathsf{H}}(\lambda), m'_{\mathsf{H}}(\lambda))$ by the Euclidean algorithm, where $m_{\mathsf{H}}(\lambda) \equiv p_0(\lambda)$ is given in equation (37) and $m'_{\mathsf{H}}(\lambda) \equiv p_1(\lambda) = 4\lambda^3 + 2\alpha_\varepsilon\lambda$. Comparing powers of λ in equation (31) with the polynomials just defined, one obtains

$$A = \frac{1}{4}, \qquad B = 0, \qquad p_2(\lambda) = \frac{\alpha_\varepsilon}{2}\lambda^2 + \beta_\varepsilon. \tag{38}$$

The algorithm only stops here if $\alpha_\varepsilon = \beta_\varepsilon = 0$ which would require $s = \delta = 0$, contrary to both s and δ being different from zero. If $\alpha_\varepsilon = 0$ is assumed, H is diagonalizable for all ε since β_ε cannot take the value zero, and the algorithm stops after the next step, outputting $\beta_\varepsilon \propto 1$ as greatest common factor. Assume now $\alpha_\varepsilon \neq 0$ and apply the division algorithm to the pair $(p_1(\lambda), p_2(\lambda))$. The unknown constants in $q_2(\lambda) = C\lambda + D$, and the remainder polynomial $p_3(\lambda)$ are found to be

$$C = \frac{8}{\alpha_\varepsilon}, \qquad D = 0, \qquad p_3(\lambda) = \frac{2}{\alpha_\varepsilon}\left(\alpha_\varepsilon^2 - 4\beta_\varepsilon\right)\lambda. \tag{39}$$

For the algorithm to stop, one must have $p_3(\lambda) = 0$. This, however, does not happen whatever the value of ε since $\alpha_\varepsilon^2 - 4\beta_\varepsilon = \delta^2(4s^2 + \delta^2) > 0$. The next iteration of the algorithm leads to

$$E = \frac{\alpha_\varepsilon^2}{4\left(\alpha_\varepsilon^2 - 4\beta_\varepsilon\right)}, \qquad F = 0, \qquad p_4(\lambda) = \beta_\varepsilon, \tag{40}$$

where $q_3(\lambda) = E\lambda + F$. Producing a remainder polynomial of degree zero in λ, the condition for the minimal polynomial to have multiple roots is finally given by

$$\beta_\varepsilon = \varepsilon^4 - (2s^2 + \delta^2)\varepsilon^2 + s^4 = 0. \tag{41}$$

This fourth-order polynomial in ε has roots

$$\pm\varepsilon_\pm = \pm\sqrt{\sigma \pm \Delta\sigma}, \qquad \sigma = s^2 + \frac{\delta^2}{2} > 0, \qquad \Delta\sigma = \sqrt{\sigma^2 - s^4} \in (0, \sigma). \tag{42}$$

For each of these four real values of the parameter ε, the matrix H is *not* diagonalizable. In other words, the algebraic and geometric multiplicity of the eigenvalues of H do not coincide, and its eigenstates do not span the space $\mathbb{C}^4$. It is important to note that the eigenvalues of the matrix H are not known at this stage.

5.3. The global structure of H

Let us now determine the global properties of H. This is easily done upon combining (42) with the characteristic polynomial (37) in its factorized form,

$$p_{\mathsf{H}}(\lambda) = (\lambda - \lambda_+)(\lambda + \lambda_+)(\lambda - \lambda_-)(\lambda + \lambda_-) \tag{43}$$
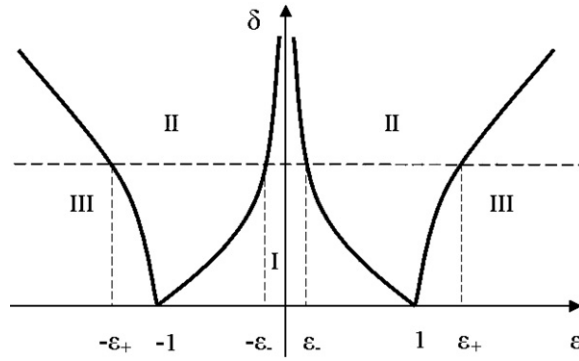
Figure 1. Parameter space of the matrix H defined in (36). The matrix H is not diagonalizable on the full lines which divide the parameter space into five regions: region I: four real eigenvalues; regions II: two real and one pair of complex-conjugate eigenvalues; regions III: two pairs of complex-conjugate eigenvalues.

with roots

$$\pm\lambda_{\pm} = \pm\sqrt{\sigma \pm \Delta\sigma - \varepsilon^2}, \tag{44}$$

expressed directly in terms of σ and $\Delta\sigma$.

To graphically represent the parameter space of H and its properties, it is convenient to eliminate the parameter s by the scaling $\varepsilon \to s\varepsilon$ and $\delta \to s\delta$. This effectively amounts to sending $s \to 1$, and equation (41) simplifies to $\varepsilon^4 - 2(1 + \delta^2/2)\varepsilon^2 + 1 = 0$, plotted in figure 1.

Imagine to move along the dashed horizontal line, determined by a fixed positive value of $\delta > 0$ and variable ε. For $\varepsilon = 0$, the matrix H is Hermitian, hence it has four distinct real eigenvalues and four orthonormal eigenstates. In region I, where $0 < |\varepsilon| < \sqrt{\sigma - \Delta\sigma}$, equation (44) says that the eigenvalues remain real and distinct; a complete, not necessarily orthonormal set of four eigenstates continues to exist since (41) does not hold. When $\varepsilon = \pm\sqrt{(\sigma - \Delta\sigma)}$, two eigenvalues coincide numerically, and the corresponding two eigenstates merge into a single one, leaving H with an incomplete basis. Then, for $\sqrt{\sigma - \Delta\sigma} < |\varepsilon| < \sqrt{\sigma + \Delta\sigma}$, in region II, two real eigenvalues and a pair of complex-conjugate ones exist, with H being diagonalizable throughout since (41) does not hold. At $\varepsilon = \pm\sqrt{(\sigma + \Delta\sigma)}$, the remaining two real eigenvalues degenerate to a single one, leaving H non-diagonalizable again with only three eigenstates. Finally, in region III, defined by $\sqrt{\sigma + \Delta\sigma} < |\varepsilon|$, the matrix H is diagonalizable and it comes with two pairs of complex-conjugate eigenvalues.

Finally, for $\delta = 0$, the matrix H in (36) decouples into a pair of identical two-dimensional matrices. The left boundary of region I sees the real eigenvalues of H degenerate pairwise which is consistent with the observations made earlier. At $\varepsilon = \pm 1$, only two eigenstates exist while all four eigenvalues coincide numerically. Beyond this value of ε, there are two pairs of identical complex-conjugate eigenvalues, and the associated basis is complete.

For PT-symmetric systems described by matrices of higher dimensions it is, in general, not possible to find the roots of the characteristic polynomial. Nevertheless, a discussion of the parameter space can still be given: to this end one needs to detect the number of real and complex eigenvalues for each set of parameter values; an algorithm capable of doing this will be presented in [13].

6. Discussion and outlook

An algorithm has been presented which allows one to determine whether a given PT-invariant matrix M is diagonalizable. To do so, it is not necessary to determine the roots of its characteristic polynomial. In terms of linear algebra, the algorithm decides whether the given matrix is *similar* [7] to a diagonal matrix or to a matrix containing at least one Jordan block of dimension two or more. Somewhat surprisingly, this question seems to have been addressed only recently from an algorithmic point of view.

It seems worthwhile to point out that the test for multiple roots of a polynomials can, in fact, be used without any change to determine whether the eigenvalues of a given Hermitian matrix are degenerate or not. The present author is not aware that this observation has been made before.

When applied to a family of PT-symmetric matrices, the algorithm outputs polynomial conditions polynomial in the perturbation parameter. These conditions are satisfied for sets of matrices all of which are not diagonalizable, and they divide the full parameter space into regions of diagonalizable matrices with qualitatively different spectra. When combined with an algorithm to identify the number of real and complex eigenvalues of M, a complete picture of the system's properties in the entire parameter space can be established.

Many PT-symmetric systems—including the first one studied from this perspective [5]—have been defined on Hilbert spaces with countably infinite dimension. Various concepts such as eigenvalues and eigenstates or the difference between algebraic and geometric multiplicities of degenerate eigenvalues continue to exist in the more general case [4]. Although Hermitian operators in finite- and infinite-dimensional spaces, have many properties in common, various other concepts of the matrix case are not easily carried over to the more general situation. For any *algorithm*, finiteness is a crucial feature: the number of steps required to identify a potential common factor of two polynomials is always finite, no matter what their degree. It will be interesting to see whether algorithmic tests for diagonalizability of operators acting on spaces with countably infinite dimension can be found.

References

 [1] Bender C M 2004 *Czech. J. Phys.* **54** 1027
 [2] Mostafazadeh A 2004 *Czech. J. Phys.* **54** 1125
 [3] Weigert St 2004 *Czech. J. Phys.* **54** 1139
 [4] Kato T 1984 *Perturbation Theory for Linear Operators* (New York: Springer)
 [5] Bender C M and Boettcher S 1998 *Phys. Rev. Lett.* **24** 5243
 [6] Berry M V, Bender C M and Mandilara A 2002 *J. Phys. A: Math. Gen.* **35** L467
 [7] Lancaster P and Tismenetsky M 1985 *The Theory of Matrices* 2nd edn (Orlando, FL: Academic)
 [8] Abate M 1997 *Am. Math. Mon.* **104** 824
 [9] Heiss W D 2004 *J. Phys. A: Math. Gen.* **37** 2455
[10] Horn R A and Lopatkin Alexey K 1999 *Linear Algebr. Appl.* **299** 153
[11] Horn R A and Johnson C R 1985 *Matrix Analysis* (New York: Cambridge University Press)
[12] Weisstein E W *et al Matrix minimal polynomial* http://mathworld.wolfram.com/MatrixMinimalPolynomial.html
[13] Weigert S *An Algorithmic Test for the 'Breaking' of PT-Invariance in Finite-Dimensional Systems* (in preparation)