A Computable Spectral Theorem

Martin Ziegler[1⋆] and Vasco Brattka[2]

[1] Heinz Nixdorf Institute, Fachbereich 17
University of Paderborn, 33095 Paderborn, Germany
ziegler@uni-paderborn.de

[2] Theoretische Informatik I, Informatikzentrum
FernUniversität, 58084 Hagen, Germany
vasco.brattka@fernuni-hagen.de

Abstract. Computing the spectral decomposition of a normal matrix is among the most frequent tasks to numerical mathematics. A vast range of methods are employed to do so, but all of them suffer from instabilities when applied to degenerate matrices, i.e., those having multiple eigenvalues. We investigate the spectral representation's effectivity properties on the sound formal basis of computable analysis. It turns out that in general the eigenvectors cannot be computed from a given matrix. If however the size of the matrix' spectrum (=number of different eigenvalues) is known in advance, it *can* be diagonalized effectively. Thus, in principle the spectral decomposition can be computed under remarkably weak non-degeneracy conditions.

1 Introduction

The Spectral Theorem for normal matrices is one of the most important theorems of linear algebra. It ensures the existence of an appropriately rotated coordinate system in which a normal operator becomes diagonal.

Theorem 1 (Spectral Theorem). *Let $A \in \mathbb{C}^{n\times n}$ be a normal matrix. There exists an orthogonal basis $(x_1, ..., x_n)$ of $\mathbb{C}^n$ and complex numbers $\lambda_1, ..., \lambda_n \in \mathbb{C}$ such that $Ax_i = \lambda_i x_i$ for each $i = 1, ..., n$.*

The Spectral Theorem has a large number of applications in mathematics, computer science, engineering and other disciplines of which we just mention the following:

- Mathematically it yields a nice normal form for normal linear operators.
- The Spectral Theorem induces an easy-to-use calculus for *functions* of self-adjoint matrices.
- It enables the explicit solvability of vector-valued linear differential equations.
- In quantum physics, it provides the basic tool for describing a measurement process ('collapse of the system into an eigenstate').

⋆ Work partially supported by DFG Grant Me872/7-3

- Whether some dynamical systems are stable or instable depends on the eigenvalues of their dynamics.
- In mechanical engineering it is particularly important to align a rotating body strictly along its mass centroid axis in order to avoid dynamic imbalances which might otherwise destroy its bearing.
- We also mention the numerous applications to computer science, e.g. in graph theory [4] and combinatorial optimization [12].
- Finally, from the complexity theoretic point of view it can increase efficiency to diagonalize a matrix A in order to compute A^n [7, 3].

However the Spectral Theorem only asserts the *existence* of the spectral decomposition — actually *finding* it is a different task. Numerical mathematics offers a vast range of methods and software libraries for doing so in the *non-degenerate* case, that is, provided the λ_i are pairwise different. For the general case on the other hand, no satisfying algorithm has been found yet: all known methods suffer from numerical instabilities and convergence problems if eigenvalues coincide.

The present work puts these experiences onto the formal basis of computable analysis, which is the theory of real number computation as it has been developed by Turing, Banach and Mazur, Grzegorczyk, Lacombe, Pour-El and Richards, Kreitz and Weihrauch, Ko and many others [16, 11, 5, 10, 13, 9, 8]. We will follow Weihrauch's approach to computable analysis (the so-called *Type-2 Theory of Effectivity*) since it offers a very uniform frame for computations on real numbers, functions and subsets [17]. Using this theory we prove that the spectral resolution of non-degenerate normal matrices can be computed but in the general case it *cannot*. The reason for this intractability lies in the discontinuous behaviour of eigenvectors: even infinitely small perturbations of the input matrix (e.g. due to floating point approximations) may entirely change the eigenvectors. On the other hand, our following main result says that it suffices to know the cardinality of the spectrum $\sigma(A)$ in advance, in order to compute the spectral resolution of A (later on we will give a precise reformulation as Theorem 13).

Theorem 2 (Computable Spectral Theorem). *Given as input a normal matrix $A \in \mathbb{C}^{n\times n}$ and the cardinality of its spectrum $|\sigma(A)|$, we can compute an orthogonal basis $(x_1, ..., x_n)$ of $\mathbb{C}^n$ and complex numbers $\lambda_1, ..., \lambda_n \in \mathbb{C}$ such that $Ax_i = \lambda_i x_i$ for each $i = 1, ..., n$.*

Our work differs from that of Pour-El and Richards [13] in that we consider *uniform* computability. This means that we investigate computability of the spectral representation as (multi-valued) function of the input matrix A. In contrast, Pour-El and Richards look for sufficient and necessary conditions on A such that the eigenvalues and eigenvectors be objects which are computable as points, i.e., without considering the computability of their dependence on A itself. Furthermore the counter-examples of Pour-El and Richards rely on so called *ad hoc* structures which have been defined for *infinite*-dimensional Hilbert spaces and thus do not apply to the finite-dimensional case we consider.

We close this section with a short survey on the organization of this paper. In the following section we recall some basic definitions from computable analysis and in Section 3 we will present relevant parts of our previous results from [18, 2], which have been developed for the canonical Euclidean space $\mathbb{R}^n$. Here we will discuss the transfer of these results to the complex case $\mathbb{C}^n$. In Section 4 we will present basic results on eigenvalues and eigenvectors which are mainly based on the Computable Fundamental Theorem of Algebra. In Section 5 we discuss and prove our main result, the Computable Spectral Theorem.

2 Computable Analysis

In this section we briefly present some basic notions from computable analysis (based on the approach of Type-2 Theory of Effectivity) and some direct consequences of well-known facts. For a precise and comprehensive reference we refer the reader to [17]. Roughly speaking, a partial real number function $f :\subseteq \mathbb{R}^n \to \mathbb{R}$ is computable, if there exists a Turing machine which transfers each sequence $p \in \Sigma^\omega$ that represents some input $x \in \mathbb{R}^n$ into some sequence $F_M(p)$ which represents the output $f(x)$. Since the set of real numbers has continuum cardinality, real numbers can only be represented by infinite sequences $p \in \Sigma^\omega$ (over some finite alphabet Σ) and thus, such a Turing machine M has to compute infinitely long. But in the long run it transfers each input sequence p into an appropriate output sequence $F_M(p)$. It is reasonable to allow only one-way output tapes for infinite computations since otherwise the output after finite time would be useless (because it could possibly be replaced later by the machine). It is straightforward how this notion of computability can be generalized to other sets X with a corresponding *representation*, that is a surjective partial mapping $\delta :\subseteq \Sigma^\omega \to X$.

Definition 3 (Computable functions). Let δ, δ' be representations of X, Y, respectively. A function $f :\subseteq X \to Y$ is called (δ, δ')–*computable*, if there exists some Turing machine M such that $\delta' F_M(p) = f\delta(p)$ for all $p \in \mathrm{dom}(f\delta)$.

Here, $F_M :\subseteq \Sigma^\omega \to \Sigma^\omega$ denotes the partial function, computed by the Turing machine M. It is straightforward how to generalize this definition to functions with several inputs and it can even be generalized to multi-valued operations $f :\subseteq X \rightrightarrows Y$, where $f(x)$ is a subset of Y instead of a single value. In this case we replace the condition in the definition above by $\delta' F_M(p) \in f\delta(p)$. We can also define the notion of (δ, δ')–*continuity* by replacing F_M by a continuous function $F :\subseteq \Sigma^\omega \to \Sigma^\omega$ (w.r.t. the Cantor topology on Σ^ω).

Already in case of the real numbers it appears that the defined notion of computability sensitively relies on the chosen representation of the real numbers. The theory of *admissible* representations completely answers the question how to find "reasonable" representations of topological spaces [17]. Let us just mention that for admissible representations δ, δ' each (δ, δ')–computable function is necessarily continuous (w.r.t. the final topologies of δ, δ').

An example of an admissible representation of the real numbers is the so-called *Cauchy representation* $\rho :\subseteq \Sigma^\omega \to \mathbb{R}$, where roughly speaking, $\rho(p) = x$ if p is an (appropriately encoded) sequence of rational numbers $(q_i)_{i\in\mathbb{N}}$ which converges rapidly to x, i.e. $|q_k - x| \leq 2^{-k}$ for all k. By standard coding techniques this representation can easily be generalized to a representation of the n-dimensional Euclidean space $\rho^n :\subseteq \Sigma^\omega \to \mathbb{R}^n$ and to a representation of $m \times n$ matrices $\rho^{m\times n} :\subseteq \Sigma^\omega \to \mathbb{R}^{m\times n}$. From the computational point of view we can identify the set of complex numbers $\mathbb{C}$ with $\mathbb{R}^2$ and in this way we obtain canonically a representation $\rho^n_{\mathbb{C}} := \rho^{2n}$ of $\mathbb{C}^n$. Analogously, we consider $\rho^{m\times n}_{\mathbb{C}}$ as a representation of the set of $m \times n$–matrices $\mathbb{C}^{m\times n}$. A vector $x \in \mathbb{C}^n$ or a matrix $A \in \mathbb{C}^{m\times n}$ will be called *computable*, if it has a computable $\rho^n_{\mathbb{C}}$–, $\rho^{m\times n}_{\mathbb{C}}$*–name*, i.e. if there exists a computable $p \in \Sigma^\omega$ such that $x = \rho^n_{\mathbb{C}}(p)$ or $A = \rho^{m\times n}_{\mathbb{C}}(p)$, respectively. A function $f :\subseteq \mathbb{C}^n \to \mathbb{R}$ is called just *computable*, if it is $(\rho^n_{\mathbb{C}}, \rho)$–computable. Analogous notions are used over the real numbers.

If δ, δ' are admissible representations of topological spaces X, Y, respectively, then there exists a canonical representation $[\delta, \delta'] :\subseteq \Sigma^\omega \to X \times Y$ of the product $X \times Y$ and a canonical representation $[\delta \to \delta'] :\subseteq \Sigma^\omega \to C(X, Y)$ of the space $C(X, Y)$ of the total continuous functions $f : X \to Y$. We just mention that these representations allow evaluation and type conversion (which correspond to an utm- and smn-Theorem). Evaluation means that the evaluation function $C(X, Y)\times X \to Y, (f, x) \mapsto f(x)$ is $([[\delta \to \delta'], \delta], \delta')$–computable and type conversion means that a function $f : Z \times X \to Y$ is $([\delta'', \delta], \delta')$–computable, if and only if the canonically associated function $f' : Z \to C(X, Y)$ with $f'(z)(x) := f(z, x)$ is $(\delta'', [\delta \to \delta'])$–computable. As a direct consequence we obtain that matrices $A \in \mathbb{C}^{m\times n}$ can effectively be identified with linear mappings $f \in \mathrm{Lin}(\mathbb{C}^n, \mathbb{C}^m)$, see Proposition 4.1 and 4.2 below. Especially, a matrix A is computable, if and only if the corresponding linear mapping is a computable function.

To express weaker computability properties, we will use two further representations $\rho_<, \rho_> :\subseteq \Sigma^\omega \to \mathbb{R}$. Roughly speaking, $\rho_<(p) = x$ if p is an (appropriately encoded) list of all rational numbers $q < x$. (Analogously, $\rho_>$ is defined with $q > x$.) It is known that a mapping $f :\subseteq X \to \mathbb{R}$ is (δ, ρ)–computable, if and only if it is $(\delta, \rho_<)$– and $(\delta, \rho_>)$–computable [17]. The $(\rho^n_{\mathbb{C}}, \rho_<)$–, $(\rho^n_{\mathbb{C}}, \rho_>)$–computable functions $f : \mathbb{C}^n \to \mathbb{R}$ are called *lower, upper semi-computable*, respectively. (Analogous notions are used for functions $f : \mathbb{R}^n \to \mathbb{R}$).

Occasionally, we will also use some standard representation $\nu_{\mathbb{N}}, \nu_{\mathbb{Q}}$ of the natural numbers $\mathbb{N} = \{0, 1, 2, ...\}$ and the rational numbers $\mathbb{Q}$, respectively.

Moreover, we will also need a representation for the space $\mathcal{L}^n_{\mathbb{C}}$ of linear subspaces $V \subseteq \mathbb{C}^n$. Since all linear subspaces are non-empty closed spaces, we can use well-known representations of the hyperspace $\mathcal{A}^n_{\mathbb{C}}$ of all closed non-empty subsets $A \subseteq \mathbb{C}^n$ (cf. [1, 17]). One way to represent such spaces is via the distance function $d_A : \mathbb{C}^n \to \mathbb{R}$, defined by $d_A(x) := \inf_{a\in A} d(x, a)$, where $d : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{R}$ denotes the canonical metric of $\mathbb{C}^n$, defined by $d(x, y) := |x-y|$. Altogether, we define three representations $\psi^n_{\mathbb{C}}, \psi^n_{\mathbb{C}<}, \psi^n_{\mathbb{C}>} :\subseteq \Sigma^\omega \to \mathcal{A}^n_{\mathbb{C}}$. We let $\psi^n_{\mathbb{C}}(p) = A$, if and only if $[\rho^n_{\mathbb{C}} \to \rho](p) = d_A$. In other words, p encodes a set A w.r.t. $\psi^n_{\mathbb{C}}$, if it encodes the distance function d_A w.r.t. $[\rho^n_{\mathbb{C}} \to \rho]$. Analogously, let

$\psi^n_{\mathbb{C}<}(p) = A$, if and only if $[\rho^n_{\mathbb{C}} \to \rho_>](p) = d_A$ and let $\psi^n_{\mathbb{C}>}(p) = A$, if and only if $[\rho^n_{\mathbb{C}} \to \rho_<](p) = d_A$. One can prove that $\psi^n_{\mathbb{C}<}$ encodes "positive" information about the set A (all open rational balls $B(q,r) := \{x \in \mathbb{C}^n : d(x,q) < r\}$ which intersect A can be enumerated), and $\psi^n_{\mathbb{C}>}$ encodes "negative" information about A (all closed rational balls $\overline{B}(q,r)$ which do not intersect A can be enumerated). The final topology induced by $\psi^n_{\mathbb{C}}$ on $\mathcal{A}^n_{\mathbb{C}}$ is the Fell topology. It is a known fact that a mapping $f :\subseteq X \to \mathcal{A}^n_{\mathbb{C}}$ is $(\delta, \psi^n_{\mathbb{C}})$–computable, if and only if it is $(\delta, \psi^n_{\mathbb{C}<})$– and $(\delta, \psi^n_{\mathbb{C}>})$–computable [17]. We mention that

1. the operation $(f, A) \mapsto f^{-1}(A) \subseteq \mathbb{C}^n$ is $\big([[\rho^n_{\mathbb{C}} \to \rho^m_{\mathbb{C}}], \psi^m_{\mathbb{C}>}], \psi^n_{\mathbb{C}>}\big)$–computable,
2. the operation $(f, B) \mapsto \overline{f(B)} \subseteq \mathbb{C}^m$ is $\big([[\rho^n_{\mathbb{C}} \to \rho^m_{\mathbb{C}}], \psi^n_{\mathbb{C}<}], \psi^m_{\mathbb{C}<}\big)$–computable.

From these properties one can deduce some computability properties of kernel and image, see Proposition 4.3 and 4.4 below.

A closed set $A \subseteq \mathbb{C}^n$ is called *recursive*, if it is empty or if there is a computable $p \in \Sigma^\omega$ such that $A = \psi^n_{\mathbb{C}}(p)$. Thus, the non-empty recursive subsets $A \subseteq \mathbb{C}^n$ are exactly those with computable distance function $d_A : \mathbb{C}^n \to \mathbb{R}$. We will apply all defined notions analogously to closed subsets of $\mathbb{R}^n$ and we will denote the corresponding representations simply without index "$\mathbb{C}$" (cf. [18, 2]).

3 Computable Linear Algebra

In our previous papers [18, 2] we have investigated some basic computability properties of linear algebra on the canonical Euclidean vector space $\mathbb{R}^n$. The purpose of this section is to transfer these results to the complex vector space $\mathbb{C}^n$. We will argue that all proofs can be transfered in a one-to-one manner without any essentially new aspects.

In the following we assume that the unitary vector space $\mathbb{C}^n$ is endowed with the canonical inner product, defined by $x \cdot y := \sum_{i=1}^n x_i \overline{y_i}$ for vectors $x = (x_1, ..., x_n), y = (y_1, ..., y_n) \in \mathbb{C}^n$. Here $\overline{z} := a - ib$ denotes the *conjugate* of the complex number $z = a + ib \in \mathbb{C}$. Correspondingly, we will use the norm $|x| := \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^n x_i \overline{x_i}}$. The following proposition reformulates results from [18] for the complex case.

Proposition 4. *Consider the following canonical mappings from linear algebra:*

1. $\mathrm{Lin}(\mathbb{C}^n, \mathbb{C}^m) \to \mathbb{C}^{m\times n}$ *is* $\big([\rho^n_{\mathbb{C}} \to \rho^m_{\mathbb{C}}], \rho^{m\times n}_{\mathbb{C}}\big)$*–computable,*
2. $\mathbb{C}^{m\times n} \to \mathrm{Lin}(\mathbb{C}^n, \mathbb{C}^m)$ *is* $\big(\rho^{m\times n}_{\mathbb{C}}, [\rho^n_{\mathbb{C}} \to \rho^m_{\mathbb{C}}]\big)$*–computable,*
3. $\ker : \mathbb{C}^{m\times n} \to \mathcal{A}^n_{\mathbb{C}}$ *is* $\big(\rho^{m\times n}_{\mathbb{C}}, \psi^n_{\mathbb{C}>}\big)$*–computable, but neither* $(\rho^{m\times n}_{\mathbb{C}}, \psi^n_{\mathbb{C}<})$*–computable, nor continuous,*
4. $\mathrm{span} : \mathbb{C}^{m\times n} \to \mathcal{A}^m_{\mathbb{C}}$ *is* $(\rho^{m\times n}_{\mathbb{C}}, \psi^m_{\mathbb{C}<})$*–computable, but neither* $(\rho^{m\times n}_{\mathbb{C}}, \psi^m_{\mathbb{C}>})$*–computable, nor continuous,*
5. $\det : \mathbb{C}^{n\times n} \to \mathbb{R}$ *is* $\big(\rho^{n\times n}_{\mathbb{C}}, \rho\big)$*–computable,*

6. $\operatorname{rank} : \mathbb{C}^{m\times n} \to \mathbb{R}$ *is* $\left(\rho_{\mathbb{C}}^{m\times n}, \rho_<\right)$*–computable, but neither* $\left(\rho_{\mathbb{C}}^{m\times n}, \rho_>\right)$*–computable, nor continuous,*
7. $\dim :\subseteq \mathcal{A}_{\mathbb{C}}^n \to \mathbb{R}$ *is* $\left(\psi_{\mathbb{C}<}^n, \rho_<\right)$*– and* $\left(\psi_{\mathbb{C}>}^n, \rho_>\right)$*–computable.*

All the proofs given in [18] can be transfered to the complex case one-to-one. This is mainly due to the fact that topologically and computationally we can identify $\mathbb{C}$ with $\mathbb{R}^2$. Wherever the rational numbers $\mathbb{Q}$ have been used, we substitute $\mathbb{Q}[i]$ for $\mathbb{Q}$ (and we replace S^{n-1} by the border of the unit ball of $\mathbb{C}^n$ in the proof of Lemma 8 in [18]).

While Proposition 4.3 shows that the solution space of a (homogeneous) linear equation $Ax = 0$ does not depend continuously on the matrix A (w.r.t. $\psi_{\mathbb{C}}^n$), we have proved in [2] that we can compute the solution space if its dimension is known in advance. We reformulate this result for the complex case.

Theorem 5 (Computable Solvability of Linear Equations). *The function*

$$S :\subseteq \mathbb{C}^{m\times n} \times \mathbb{R} \to \mathcal{A}_{\mathbb{C}}^n, \quad (A, d) \mapsto \ker(A)$$

with $\operatorname{dom}(S) := \{(A, d) : d = \dim\ker(A)\}$ *is* $\left([\rho_{\mathbb{C}}^{m\times n}, \rho], \psi_{\mathbb{C}}^n\right)$*–computable.*

Again the proof can be transfered directly. Whenever the Euclidean norm and the canonical inner product of $\mathbb{R}^n$ has been used in [2], we replace them by the corresponding operations of $\mathbb{C}^n$ (the Cauchy-Schwarz inequality holds analogously in this case). Also the Gram-Schmidt orthogonalization process can be used in the complex case analogously. Finally, we mention a corollary, which shows that we can also effectively find a basis, in case that the dimension of the space is known.

Corollary 6. *The multi-valued mapping*

$$B :\subseteq \mathcal{A}_{\mathbb{C}}^n \times \mathbb{R} \rightrightarrows \mathcal{A}_{\mathbb{C}}^n, (V, d) \mapsto \left\{\{b_1, ..., b_d\} \subseteq \mathbb{C}^n : (b_1, ..., b_d) \text{ is a basis of } V\right\}$$

with $\operatorname{dom}(B) := \{(V, d) : d = \dim(V)\}$ *is* $([\psi_{\mathbb{C}}^n, \rho], \psi_{\mathbb{C}}^n)$*–computable.*

4 Eigenvalues, Eigenvectors and Eigenspaces

Recall that the *adjoint* of a complex $n\times n$-matrix $A = \left(a_{ij}\right) \in \mathbb{C}^{n\times n}$ is defined by $A^* = \left(\overline{a}_{ji}\right)$. A number $\lambda \in \mathbb{C}$ with $Az = \lambda z$ for some non-zero vector $z \in \mathbb{C}^n$ is called *eigenvalue* of A and z a corresponding *eigenvector*. The set of eigenvalues, called *spectrum* $\sigma(A)$, is precisely the set of zeros of the *characteristic polynomial* $\det(zI - A) \in \mathbb{C}[z]$, where $I \in \mathbb{C}^{n\times n}$ denotes the $n \times n$-unit matrix. The set of eigenvectors to $\lambda \in \sigma(A)$ — extended by $0 \in \mathbb{C}^n$ — forms a vector space, the *eigenspace* $\ker(\lambda I - A)$. Our first observation states that given a matrix A, we can compute its characteristic polynomial.

Proposition 7. *Given a complex $n \times n$-matrix $A \in \mathbb{C}^{n\times n}$, we can compute its characteristic polynomial*

$$z^n + \sum_{i=0}^{n-1} a_i z^i := \det(zI - A).$$

More precisely, the mapping $\chi : \mathbb{C}^{n\times n} \to \mathbb{C}^n, A \mapsto (a_0, ..., a_{n-1})$ is $(\rho_{\mathbb{C}}^{n\times n}, \rho_{\mathbb{C}}^{n+1})$–computable.

The proof is a straightforward calculation using Leibniz' formula

$$\det(B) = \sum_{\sigma\in\mathcal{S}_n} \mathrm{sign}(\sigma) \prod_{i=1}^{n} b_{\sigma(i)i},$$

for matrices $B = (b_{ij}) \in \mathbb{C}^{n\times n}$, where $\mathcal{S}_n$ denotes the set of all permutations $\sigma : \{1, ..., n\} \to \{1, ..., n\}$.

In the next step a computable version of the Fundamental Theorem of Algebra, due to Specker [15], can help us to compute the eigenvalues of a given matrix. It says that, on input of a polynomial $p \in \mathbb{C}[z]$ of degree $n \in \mathbb{N}$, its n complex zeroes can be computed.

Theorem 8 (Computable Fundamental Theorem of Algebra). *Consider the unique normed polynomial having exactly the zeros $\lambda_1, ..., \lambda_n \in \mathbb{C}$ (including multiplicities):*

$$z^n + \sum_{i=0}^{n-1} a_i z^i := \prod_{i=1}^{n} (z - \lambda_i)$$

The mapping $\mathbb{C}^n \ni (\lambda_1, ..., \lambda_n) \mapsto (a_0, ..., a_{n-1}) \in \mathbb{C}^n$ is surjective and has a multi-valued $(\rho_{\mathbb{C}}^n, \rho_{\mathbb{C}}^n)$–computable right inverse $Z : \mathbb{C}^n \rightrightarrows \mathbb{C}^n$.

See [17] for a sketch of the proof. A combination of this theorem with the previous proposition and the fact that $\mathbb{C}^n \to \mathcal{A}_{\mathbb{C}}^n, (a_0, ..., a_{n-1}) \mapsto \{a_0, ..., a_{n-1}\}$ is $(\rho_{\mathbb{C}}^n, \psi_{\mathbb{C}}^n)$–computable, directly yields the computability of the spectrum mapping.

Corollary 9. *The map $\sigma : \mathbb{C}^{n\times n} \to \mathcal{A}_{\mathbb{C}}^n, \quad A \mapsto \sigma(A)$ is $(\rho_{\mathbb{C}}^{n\times n}, \psi_{\mathbb{C}}^n)$–computable.*

Computing eigenspaces is slightly more difficult. As we have seen in Proposition 4.3 and Theorem 5 we can compute the kernel of a matrix, provided that we know its dimension in advance. Thus, we directly obtain the following corollary on computing eigenspaces.

Corollary 10. *Given a matrix $A \in \mathbb{C}^{n\times n}$ together with some eigenvalue $\lambda \in \mathbb{C}$ and $d = \dim\ker(\lambda I - A)$, we can compute the eigenspace $\ker(\lambda I - A)$. More precisely,*

$$E :\subseteq \mathbb{C}^{n\times n} \times \mathbb{C} \times \mathbb{R} \to \mathcal{A}_{\mathbb{C}}^n, \quad (A, \lambda, d) \mapsto \ker(\lambda I - A),$$

with $\mathrm{dom}(E) := \{(A, \lambda, d) : \lambda$ is eigenvalue of A and $d = \dim\ker(\lambda I - A)\}$, is $([\rho_{\mathbb{C}}^{n\times n}, \rho_{\mathbb{C}}, \rho], \psi_{\mathbb{C}}^n)$–computable

Considering a single computable matrix $A \in \mathbb{C}^{n\times n}$ together with some eigenvalue $\lambda \in \mathbb{C}$ (which necessarily is computable too by Corollary 9), the discrete dimension is always available as a further (computable) input; thus, we obtain the following corollary.

Corollary 11. *Each computable real or complex $n \times n$-matrix has a recursive spectrum $\sigma(A)$, especially, all eigenvalues are computable. Moreover, the corresponding eigenspaces are all recursive and each eigenvalue admits a computable eigenvector.*

5 The Spectral Theorem

Recall that a complex $n \times n$-matrix $A \in \mathbb{C}^{n\times n}$ is called *self-adjoint*, if $A = A^*$ and *normal*, if $AA^* = A^*A$. For normal matrices A, the dimension $d \in \mathbb{N}$ of the eigenspace $\ker(\lambda I - A)$ of some eigenvalue $\lambda \in \mathbb{C}$ equals the algebraic *multiplicity* of the zero λ of the characteristic polynomial $\det(zI - A)$, i.e., the i-th derivative $d^i/dz^i \det(zI - A) \in \mathbb{C}[z]$ vanishes at $z = \lambda$ for $i = 0, ..., d-1$ but not for $i = d$.

Now our goal is to prove a computational version of the classical Spectral Theorem 1. Unfortunately, it turns out that the spectral resolution cannot be computed directly from a given self-adjoint matrix $A \in \mathbb{C}^{n\times n}$. The chief snag is that, although the spectrum $\sigma(A)$ can be computed from A, its cardinality $|\sigma(A)|$ does not continuously depend on A. Even worser, the following proposition shows that the eigenvectors of real symmetric 2×2-matrices A do not depend continuously on A. For the proof, we borrow an example of Rellich [14] which can also be found in [6].

Proposition 12. *There exists no $(\rho^{2\times 2}, \rho^{2\times 2})$–continuous multi-valued function*

$$F :\subseteq \mathbb{R}^{2\times 2} \rightrightarrows \mathbb{R}^{2\times 2}$$

such that each $(x, y) \in F(A)$ is an orthogonal basis of $\mathbb{R}^2$ consisting of eigenvectors of A, and $A \in \mathrm{dom}(F)$ whenever $A \in \mathbb{R}^{2\times 2}$ is symmetric.

Proof. First of all, it suffices to prove the statement for ortho*normal* bases instead of ortho*gonal* bases, since each orthogonal basis (x, y) can continuously be normalized to $(x/|x|, y/|y|)$. Let us assume, that there exists a $(\rho^{2\times 2}, \rho^{2\times 2})$–continuous multi-valued function $F :\subseteq \mathbb{R}^{2\times 2} \rightrightarrows \mathbb{R}^{2\times 2}$ which solves the problem for orthonormal bases. Now consider the continuous function $A : \mathbb{R} \to \mathbb{R}^{2\times 2}$, defined by

$$A(\varepsilon) := \exp(-1/\varepsilon^2)\begin{pmatrix} \cos(2/\varepsilon) & \sin(2/\varepsilon) \\ \sin(2/\varepsilon) & -\cos(2/\varepsilon) \end{pmatrix}, \quad A(0) := \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \ .$$

In case $\varepsilon > 0$ the eigenvalues of A are $\exp(-1/\varepsilon^2)$ and $-\exp(-1/\varepsilon^2)$; the corresponding orthonormal basis of eigenvectors of A is uniquely determined up to order and orientation and consists of the vectors

$$x(\varepsilon) := \begin{pmatrix} \cos(1/\varepsilon) \\ \sin(1/\varepsilon) \end{pmatrix}, \quad y(\varepsilon) := \begin{pmatrix} \sin(1/\varepsilon) \\ -\cos(1/\varepsilon) \end{pmatrix}.$$

Thus, the eigenspaces and eigenvectors of $A(\varepsilon)$ rotate faster and faster if ε tends to 0. By assumption $F \circ A : \mathbb{R} \rightrightarrows \mathbb{R}^{2\times 2}$ is $(\rho, \rho^{2\times 2})$–continuous. Let us fix some $z := (x, y) \in FA(0)$, which could be any orthonormal basis of $\mathbb{R}^2$. Then there are arbitrarily small $\varepsilon > 0$ such that $z(\varepsilon) := (x(\varepsilon), y(\varepsilon))$ is far away from z, no matter how $x(\varepsilon)$ and $y(\varepsilon)$ are oriented or ordered, e.g. $|z - z(\varepsilon)| > \frac{1}{2}$. Thus, $F \circ A$ cannot be $(\rho, \rho^{2\times 2})$–continuous. Contradiction! □

Any substantial information about the degree of degeneracy however does enable computability of the spectral representation. Especially the additional knowledge of the cardinality of the spectrum suffices, as our following result shows.

Theorem 13 (Computable Spectral Theorem). *There exists a multi-valued* $([\rho_{\mathbb{C}}^{n\times n}, \rho], [\rho_{\mathbb{C}}^{n\times n}, \rho_{\mathbb{C}}^{n}])$*–computable function*

$$R :\subseteq \mathbb{C}^{n\times n} \times \mathbb{R} \rightrightarrows \mathbb{C}^{n\times n} \times \mathbb{C}^n$$

with $\mathrm{dom}(R) := \{(A, s) : A$ *normal and* $s = |\sigma(A)|\}$ *such that for any normal* $A \in \mathbb{C}^{n\times n}$ *with* $s = |\sigma(A)|$ *and* $((x_1, ..., x_n), (\lambda_1, ..., \lambda_n)) \in R(A, s)$ *we obtain*

$$Ax_i = \lambda_i x_i$$

and $(x_1, ..., x_n)$ *is an orthogonal basis of* $\mathbb{C}^n$.

That means that our algorithm has to know in advance the number of pairwise different eigenvalues. Indeed from this, the multiplicities of all eigenvalues can be deduced. This follows from the following technical lemma (where, for technical simplicity, multiplicities occur as complex numbers).

Lemma 14. *The function* $T :\subseteq \mathbb{C}^{n+1} \to \mathcal{A}_{\mathbb{C}}^2$*, with*

$$T(\lambda_1, ..., \lambda_n, s) := \{(\lambda, d) : \lambda = \lambda_i \text{ for precisely } d \text{ values of } i = 1, ..., n\}$$

with $\mathrm{dom}(T) := \big\{(\lambda_1, ..., \lambda_n, s) : s = |\{\lambda_1, ..., \lambda_n\}|\big\}$ *is* $(\rho_{\mathbb{C}}^{n+1}, \psi_{\mathbb{C}}^2)$*–computable.*

Proof. We sketch the proof which is a straightforward exercise. Let us assume that $(\lambda_1, ..., \lambda_n) \in \mathbb{C}^n$ and $s = |\{\lambda_1, ..., \lambda_n\}|$ are given w.r.t. $\rho_{\mathbb{C}}$. Since the set $\{(x, y) : x \neq y\}$ is an r.e. open subsets of $\mathbb{C}^2$, we can compare all pairs (λ_i, λ_j) with $i \neq j$ until we can distinguish exactly s pairwise different "clusters" of values λ_j. At this stage we know that all d values λ in a cluster have to coincide and we can produce a name of the set of all pairs (λ, d) w.r.t. $\psi_{\mathbb{C}}^2$. □

This lemma actually enables us to complete the proof of the computable Spectral Theorem.

Proof (of Theorem 13). Let us assume that a normal matrix $A \in \mathbb{C}^{n\times n}$ is given w.r.t. $\rho_{\mathbb{C}}^{n\times n}$ and $s = |\sigma(A)|$ is given w.r.t. $\rho_{\mathbb{C}}$. By applying Proposition 7 and the computable version of the Fundamental Theorem of Algebra 8, we can compute a tuple $(\lambda_1, ..., \lambda_n) \in \mathbb{C}^n$ of eigenvalues. By the previous lemma we can compute the set $T(\lambda_1, ..., \lambda_n, s)$ of all pairs (λ, d) of eigenvalues λ of A together with their multiplicities d, i.e. $d = \dim\ker(\lambda I - A)$. For each of these k pairs (λ, d) we apply Corollary 10 in order to compute the corresponding eigenspace $E(A, \lambda, d) = \ker(\lambda I - A)$ w.r.t. $\psi^n_{\mathbb{C}}$. By applying Corollary 6 we can compute a basis $(y_1, ..., y_d) \in \mathbb{C}^n$ of the eigenspace $E(A, \lambda, d)$. Now we employ the Gram-Schmidt orthogonalization process to compute an orthogonal basis $(z_1, ..., z_d) \in \mathbb{C}^n$ by

$$z_1 := y_1, \quad z_{j+1} := y_{j+1} - \sum_{i=1}^{j} \frac{y_{j+1} \cdot z_i}{|z_i|^2} z_i$$

for all $j = 1, ..., d-1$. Since eigenspaces for different eigenvalues are orthogonal to each other we can simply put the s bases of all s eigenspaces $E(A, \lambda, d)$ together and we obtain an orthogonal basis $(x_1, ..., x_n)$ of $\mathbb{C}^n$. Finally, we mention that we can arrange the tuple of eigenvalues $(\lambda'_1, ..., \lambda'_n)$ in a corresponding order such that $A\lambda'_i = \lambda'_i x_i$ for all $i = 1, ..., n$. □

6 Conclusion

In this paper we have continued our project to investigate computability properties in linear algebra with rigorous methods from computable analysis. Using previous results from [18, 2] we have proved an effective version of the finite-dimensional Spectral Theorem. One could continue this project in several different directions: on the one hand, there are other normal forms in linear algebra which have not yet been studied from the point of view of computability. On the other hand, one could ask for more general cases, such as unitary vector spaces endowed with an inner product different from the canonical one, or infinite-dimensional Hilbert spaces. Moreover, it would be a fascinating project to establish connections with applications of the Spectral Theorem.

References

1. Vasco Brattka and Klaus Weihrauch. Computability on subsets of Euclidean space I: Closed and compact subsets. *Theoretical Computer Science*, 219:65–93, 1999.
2. Vasco Brattka and Martin Ziegler. Computability of linear equations. unpublished, 2000.
3. Peter Bürgisser, Michael Clausen, and M. Amin Shokrollahi. *Algebraic complexity theory*, vol. 315 of *Grundl. der math. Wissenschaften*. Springer, Berlin, 1997.
4. Fan R.K. Chung. *Spectral Graph Theory*. CBMS Regional Conference Series, vol. 92, American Mathematical Society, Providence 1997.
5. Andrzej Grzegorczyk. On the definitions of computable real continuous functions. *Fundamenta Mathematicae*, 44:61–71, 1957.
6. Tosio Kato. *Perturbation Theory for Linear Operators*, vol. 132 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 2. edition, 1976.
7. Donald E. Knuth. *The art of computer programming*. Computer Science and Information Processing. Addison-Wesley, Reading, 1981. Volume 2, Seminumerical algorithms.

8. Ker-I Ko. *Complexity Theory of Real Functions.* Progress in Theoretical Computer Science. Birkhäuser, Boston, 1991.
9. Christoph Kreitz and Klaus Weihrauch. A unified approach to constructive and recursive analysis. In M.M. Richter, E. Börger, W. Oberschelp et al. eds., *Computation and Proof Theory*, vol. 1104 of *LNM*, 259–278, Springer, Berlin, 1984.
10. Daniel Lacombe. Les ensembles récursivement ouverts ou fermés, et leurs applications à l'Analyse récursive. *Compt. Rend. Acad. des Sci. Paris*, 246:28–31, 1958.
11. S. Mazur. *Computable Analysis*, vol. 33. Razprawy Matematyczne, Warsaw, 1963.
12. Bojan Mohar and Svatopluk Poljak. Eigenvalues in Combinatorial Optimization. In R. A. Brualdi, S. Friedland and V. Klee, (eds.) *Combinatorial and Graph-Theoretic Problems in Linear Algebra*, pages 107–151, The IMA Volumes in Mathematics and Its Applications, vol. 50, Springer, Berlin, 1993.
13. Marian B. Pour-El and J. Ian Richards. *Computability in Analysis and Physics.* Perspectives in Mathematical Logic. Springer, Berlin, 1989.
14. Franz Rellich. Störungstheorie der Spektralzerlegung I., Analytische Störung der isolierten Punkteigenwerte eines beschränkten Operators, in *Math. Ann.* 113:600-619, 1937.
15. Ernst Specker. The fundamental theorem of algebra in recursive analysis. In B. Dejon and P. Henrici, editors, *Constructive Aspects of the Fundamental Theorem of Algebra*, pages 321–329, London, 1969. Wiley-Interscience.
16. Alan M. Turing. On computable numbers, with an application to the "Entscheidungsproblem". *Proc. of the London Math. Soc.*, 42(2):230–265, 1936.
17. Klaus Weihrauch. *Computable Analysis.* Springer, Berlin, 2000.
18. Martin Ziegler and Vasco Brattka. Computing the dimension of linear subspaces. In *SOFSEM'2000: Theory and Practice of Informatics*, Lecture Notes in Computer Science, Springer, Berlin, 2000 (to appear).