Chapter 3

Computational Complexity in the Analysis of Quantum Operations

Miłosz Michalski

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/56159

1. Introduction

Mathematical theory of infinite dimensional Hilbert spaces and the theory of operator algebras acting in such spaces (or C^* algebras in a more abstract approach) provide a standard setting for the formulation of modern quantum mechanics. On the other hand, experimental and theoretical progress achieved in the field of quantum information theory in the last two decades has indicated the practical and technological importance of low-dimensional quantum systems, where only a few basic modes play a significant role. Such modes can often be effectively decoupled from the rest of the system and controlled separately, providing physical realizations of qubits, qutrits and other basic information carriers. Regardless of concrete physical realization, be it photon polarization, electron or nuclear spin, charge in Josephson junctions to name just a few, the mathematical description of such systems requires only finite-dimensional Hilbert space language and finite-dimensional matrix algebras. Such structures are in principle computationally manageable in sharp contrast to the infinite dimensional ones.

It has to be pointed out, however, that there is a lot of misconception concerning the above mentioned "manageability" notion in today's quantum information literature. For instance, one of the most fundamental errors appearing in innumerable papers is to indiscriminately resort to the spectral resolution technique for hermitian matrices. Such an operation cannot be considered computationally effective if the size of the matrix exceeds 4: then it unavoidably involves solving an algebraic equation of degree 5 or more. Such task can be achieved only by an approximate numerical process, and therefore any emerging questions can be answered only up to numerical precision. The latter can be critical, for example, in checking whether a hermitian matrix has a negative eigenvalue.

Fortunately, in many situations similar to the one just described there are other alternative ways to obtain a *precise* answer, avoiding the approximate numerical computations. This

is achieved by limiting oneself to the so-called finite *rational* computational procedures, involving only finitely many arithmetic operations on initial data so that the data as well as all intermediate and final computational results belong to the same number field. In particular, the use of transcendental functions is thus excluded.

The present chapter will be devoted to a review of a few such procedures, important for applications in quantum information theory. We will concentrate on the questions concerning not only the effectiveness of such procedures, but also on more detailed computational complexity issues. To describe better the subject of our considerations and to fix the terminology, let us consider the already invoked example of checking whether a given selfadjoint matrix has a negative eigenvalue, which in particular is a crucial ingredient in entanglement detection procedures. Note that the problem is posed so that the precise knowledge of the eigenvalue is not essential, it is its sign that matters.

Let A be a hermitian matrix in question and $\mathcal{H}$ be the respective Hilbert space. One can formulate the negative eigenvalue problem in an equivalent form by asking whether A is or is not positive semidefinite. As it is well known, positive semidefiniteness can be characterized by several equivalent criteria, each of them being an example of a different effectiveness or complexity issue. The list of relevant criteria is the following.

1. For each normalized vector $|\psi\rangle \in \mathcal{H}$ one has $\langle \psi \,|\, A\psi \rangle \geq 0$. The test based on this criterion is *ineffective* as it involves infinite number of conditions to verify, one for each $|\psi\rangle$.
2. All eigenvalues of A are nonnegative. As we have argued above, such test cannot be considered an effective one either. In general, the correctness of the answer hinges upon the numerical precision being used. We can call such tests *asymptotically effective*, meaning that increased numerical accuracy can yield the definite yes/no answer, but no a priori fixed precision is sufficient for the correctness of the whole class of such tests.
3. All principal minors of A are nonnegative. This is certainly an *effective* criterion as it involves the evaluation of finitely many subdeterminants of A. The computation of a determinant itself is a finite rational procedure. Note however, that direct application of the present criterion requires the evaluation of nearly 2^n minors, n being the size of A. Although finite, this number grows very rapidly with n, making the test *inefficient*. In practical terms, it may easily take years to complete such a test on the fastest computers, even for A of moderate size. For example, if A results from an application of some entanglement test to a mixed state of a system composed of merely 6 qubits, then $n = 64$ and hence the number of minors to compute is about $2^{64} \approx 10^{19}$. Assuming that our computing device can evaluate 10^6 minors per second on average, the time required to complete such a test would be of the order of 10^5 years. We characterize the computational complexity of such procedures by saying that they are *nonpolynomial* in n. Problems for which only nonpolynomial solution methods are available are termed *intractable*.
4. While the test of positive definiteness (Sylvester's criterion) is much simpler, for it involves only n *leading* principal minors of A, it has no counterpart for positive semidefinite matrices. However, one can easily check that the following recursive procedure based on Gaussian elimination can be used in this case. By A_{11} we denote here the submatrix of $A = [a_{ij}]$ obtained by the deletion of its 1st row and 1st column.
   (a) If $a_{11} < 0$ then A is *not* positive semidefinite.

(b) If $a_{11} = 0$ then A is *not* positive semidefinite unless its entire 1st column is null and A_{11} is positive semidefinite.

(c) If $a_{11} > 0$ then we first perform row-elimination of the entire 1st column of A. Then A is positive semidefinite iff the resulting A_{11} is such.

This is again a finite rational procedure. The largest computational effort in completing such a check is needed when there are no 0 entries in the first column of A and, likewise, no zeros are produced in A_{11} by the elimination. Then the recursive check uses the variant (c) repeatedly, so that the total number of arithmetic operations performed is of the order of n^3. The complexity of the method is thus *polynomial* and its efficiency is much higher than that of criterion 3. If as before $n = 64$, the test will complete in less than 1 second, assuming the computer speed of 10^6 rational arithmetic operations per second.

Positive semidefiniteness is certainly a very simple issue, however the above example highlights a few characteristic aspects of computational complexity. Mathematical problems often admit many different solution methods which, similarly as in our example, may range from ineffective to very efficient ones. Effective procedures however can often prove useless in practice if the computational effort involved grows too fast with the size of input data. The complexity of problems themselves can be characterized relative to the most efficient solution methods known for them. In some cases theoretical complexity bounds can be derived for classes of problems.

In the next section we provide a brief review of fundamental notions of computational complexity theory.

2. Basic notions of computational complexity theory

In theoretical computer science, algorithms are classified according to their *time* or *space complexity*. Time complexity gives an estimate of how does the number of elementary steps in the algorithm scale with the size of input data defining an instance of the problem. Space complexity refers to the scaling of the amount of workspace or extra memory (in one convention the memory storing input data is not counted) needed in the course of computation. The complexity of *problems* is related to their inherent difficulty and is a theoretical estimate of the computational cost indispensable for their solution. Often only some lower or upper complexity bounds are known for classes of problems. The complexity theory uses the formalism of abstract Turing machines to ensure the universality of conclusions.

It is not our goal to review the complexity theory in its general abstract formulation here, but rather to provide necessary intuitions for an unacquainted reader. Those familiar with computational complexity may well skip the current section.

The scaling of solution time or workspace with problem size is expressed using the "big O" notation.

Definition 1. *For two functions* $f, g\colon \mathbb{N} \to \mathbb{R}$ *one writes* $f(n) = O\big(g(n)\big)$ *for* $n \to \infty$ *if and only if*

$$\exists\, M \in \mathbb{R} \;\; \textit{and} \;\; n_0 \in \mathbb{N} \quad \textit{such that} \quad |f(n)| \leq M|g(n)| \quad \textit{for } n > n_0\,.$$

For example, standard square matrix multiplication requires $O(n^3)$ arithmetic operations, n being the matrix size. Since no extra memory beyond that for data storage is needed for performing the multiplication, the space complexity here is $O(n^2)$. The Fast Fourier Transform performs $O(n \log n)$ arithmetic operations on an n element data vector. Evaluation of a determinant directly from its definition would involve the summation of $n!$ terms, however more efficient method using Gauss elimination reduces the effort to $O(n^3)$ arithmetic operations. Evaluation of a permanent on the other hand appears more complex (except for the case of computations over $\mathbb{Z}_2$, where $-1 \equiv 1(\mathrm{mod}2)$ and hence $\det A = \mathrm{per}\, A$): the best methods known so far [8, 23] have the complexity of $O(n2^n)$.

One of the objectives of the theory is to identify *complexity classes*, consisting of problems which can be solved by using only limited type of computational resources, which are abstractly characterized by restricted classes of Turing machines, most notably the classes P and NP. The class P consists of problems which can be solved by a deterministic Turing machine executing a number of steps bounded by a polynomial in the input data size. The class NP on the other hand consists of problems solvable in polynomial time by a *nondeterministic* Turing machine. As the latter can be simulated by a deterministic machine in exponential time, NP is often conventionally (yet not quite correctly) identified with the class of problems solved by exponential (nonpolynomial) time deterministic algorithms. Strictly speaking however, the essential feature of NP problems is that given a random candidate for a solution it takes no more than *polynomial* number of steps to verify its correctness or to reject it. Exponential time deterministic algorithms in NP can be thought of as performing an extensive "blind" search in the space of potential solutions (which is the actual source of nonpolynomial complexity) checking each of them at low (i.e. polynomial time) cost. In contrast, problems in P admit "clever" constructive solution methods.

In practical terms, problems of type P can be solved relatively fast regardless of their size, while for the NP type ones solution times become impractically long even for moderate size of input data, c.f. our discussion of positive semidefiniteness verification in the previous section. The distinction between efficient and inefficient methods is often used as a synonym for that between P and NP classes.

Obviously $\mathrm{P} \subset \mathrm{NP}$, but it is a famous open question (although today hardly believed to hold true) whether $\mathrm{P} = \mathrm{NP}$. The quest for an answer to the latter has led to the definition of various special complexity classes, in particular the class of NP-complete problems, NP-C. We say that a problem π can be polynomially transformed to another problem π', in written $\pi \propto \pi'$, if the solution of π for input data of size n can be obtained by means of the execution of an algorithm for π' at most a polynomial in n number of times on new data translated from the original input with at most polynomial effort. So if $\pi' \in \mathrm{P}$ (resp. in NP), then any π such that $\pi \propto \pi'$ is necessarily also in P (resp. NP).

Definition 2. *A problem π is* NP*-complete iff*

(i) $\pi \in \mathrm{NP}$;

(ii) $\forall\, \sigma \in \mathrm{NP} \quad \sigma \propto \pi$.

If π satisfies only condition (ii), it is said to be NP*-hard.*

http://dx.doi.org/10.5772/56159

It may appear that NP-C can well be empty, but it is not so as shown by Cook in [7]. The first NP-C problem identified by Cook was the satisfiability of Boolean functions: given a Boolean function F in variables $x_1, \ldots, x_n$ doest there exist a truth/false assignment to all x_i making the value of F true? Cook's proof gives a method of how to cast, at polynomial cost, an arbitrary nondeterministic Turing machine into the one computing Boolean functions.

Knowing at least one NP-C problem it becomes easier to identify other ones: if $\pi \in$ NP is such that $\sigma \propto \pi$ for some $\sigma \in$ NP-C, then $\pi \in$ NP-C. The list of known NP-complete problems exceeds now 3000 items. By definition, providing a polynomial time solution to any single NP-C problem would automatically prove that P = NP. Because of this, NP-complete problems are considered the hardest among NP ones. In other words, it is generally believed that the search for exact polynomial time solution methods of NP-C problems is a waste of time. On the other hand, there are numerous problems of practical interest for which neither a proof of NP-completeness nor an efficient polynomial time solution method are known. The most notable example is the problem of finding factors of large integers.

It is interesting that a large class of problems in matrix theory which possess an efficient solution can be reduced to an evaluation of a small number of determinants or, equivalently, can be expressed, as above, in terms of Gaussian elimination or — still more elementary reduction — by a series of matrix multiplications. This point of view motivates the interest in the design of fast matrix multiplication algorithms. Perhaps the best known schema of this kind is due to V. Strassen (1969) and its complexity is $O(n^{2.81})$, while more recent method of Coppersmith and Winograd (1987) improves the efficiency to $O(n^{2.367})$, the theoretical lower bound being $O(n^2)$.

An example of NP-complete matrix algebra problem is the following [5]: given an $n \times m$ matrix A over $\mathbb{Z}$ with $n \leq m$, decide whether there exist a vanishing $n \times n$ subdeterminant of A. The evaluation of a permanent is NP-hard, for it is most likely not in NP class. Again, the existence of a polynomial algorithm for the computation of per A would infer the equality P = NP. Many complicated counting problems in combinatorics and graph theory can be reduced to an evaluation of a permanent. Actually, permanent evaluation is #P-complete, meaning that all counting functions which can be defined in terms of NP problems can be polynomially reduced to it, [25].

Another important complexity category, from a physicist's point of view, is the so-called BPP class (bounded error probabilistic polynomial time) consisting of decision problems solvable in polynomial time by a *probabilistic* Turing machine, with the probability of producing wrong answer bounded from above by a constant $0 \leq p < 1/2$. Less formally, this class corresponds to Monte Carlo algorithms likely to yield correct answers and running in polynomial time. Such conditions guarantee that in practice one can perform a relatively short series of independent runs of the method to learn the correct answer with very high probability. By Chernoff bound, the probability that incorrect answer appears in a series of runs most of the time decays exponentially with the series length. If instead of probabilistic one uses *quantum* Turing machines, the resulting class is called BQP (bounded error quantum polynomial time). It is shown that BPP $\subset$ BQP, but little is known so far about the relation of either of the classes to NP.

Finally, PSPACE is a class of problems solvable by deterministic Turing machines using at most polynomial in the data size amount of workspace. It is proved that adding nondeterminism does not alter this class, namely PSPACE = NPSPACE. NP is thus clearly

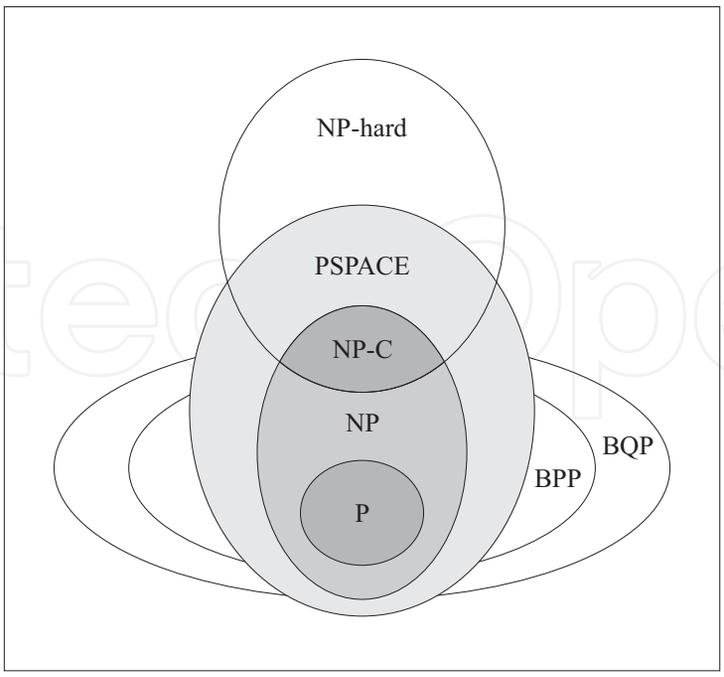

Figure 1. Hypothetical relations among complexity classes.

contained in PSPACE since using workspace of nonpolynomial size would automatically require nonpolynomial time. Fig. 1 summarizes what has been said above about the complexity classes.

Last but not least, there are problems which are provably undecidable, meaning that no finite algorithm can ever resolve them. Among such tasks there is the fascinating tiling problem [26].

Let us mention also that to date no general effective criteria are known for one of the most fundamental decision problems in quantum information, namely the determination whether a given mixed state of a bipartite system is entangled or not. All known exact methods, apart from those for low-dimensional systems, namely for $n = 4 = 2 \times 2$ and $n = 6 = 2 \times 3$, involve infinite number of computational tests (local actions of positive maps or, equivalently, evaluation of expectations of entanglement witnesses). Moreover, no effective method is in sight despite the two decades of intensive research efforts worldwide.

3. Some computational problems of quantum information theory

Quantum information (QI) theory regards quantum states as information carriers and quantum evolution of states as acts of information processing. As we have already mentioned in the Introduction, QI research focuses on low-dimensional quantum systems, qubits, qutrits and likewise, which appear to be most interesting from the point of view of potential future large-scale technological applications. Such low dimensional structures can be combined into multipartite quantum systems, realizing quantum registers and memories. Namely, given a low-dimensional Hilbert space, e.g. $\mathcal{H}_2 \simeq \mathbb{C}^2$ for a qubit, the space of the compound

http://dx.doi.org/10.5772/56159

multipartite system is then

$$\mathcal{H} \;=\; \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_2 \;=\; \mathcal{H}_2^{\otimes n} \;\simeq\; \mathbb{C}^{2^n}\,.$$

Genuinely quantum properties of such systems, most importantly the *entanglement* of their states, are proved to underlie the extraordinary efficiency of quantum information processing, surpassing that of the classical one. In what follows we shall silently assume finite-dimensionality of all quantum systems in question.

Let us recall that *pure states* of a quantum systems are represented by vectors in the respective Hilbert space, $|\psi\rangle \in \mathcal{H}$, while *observables*, i.e. measurable physical quantities, correspond to selfadjoint operators acting on $\mathcal{H}$, i.e. $A \in \mathcal{B}(\mathcal{H})$ such that $A = A^*$. In the finite-dimensional setting they can be identified with Hermitian matrices in the matrix algebra $\mathcal{M}_n(\mathbb{C})$, $n = \dim \mathcal{H}$. In passing to *mixed states* one replaces pure states with the corresponding 1-dimensional projection operators, $|\psi\rangle\langle\psi| \in \mathcal{B}(\mathcal{H})$, and one defines the mixed states as statistical sums of mutually orthogonal projections, $\varrho = \sum p_i |\psi_i\rangle\langle\psi_i|$ with real positive p_i summing up to 1. So defined, mixed states are quantum counterparts of classical discrete probability distributions. Their representatives are called density matrices. It can be easily seen that density matrices form a convex subset $\Sigma = \Sigma(\mathcal{H})$ of $\mathcal{B}(\mathcal{H})$ characterized by positive semidefiniteness and normalization of trace[1]

$$\varrho \in \mathcal{B}(\mathcal{H}) \qquad \text{such that} \qquad \varrho \geq 0 \quad \text{and} \quad \operatorname{Tr}\varrho = 1\,.$$

According to the postulates of quantum mechanics, dynamical evolution of quantum systems is described by the Schrödinger equation, which, when reformulated for mixed states, takes the form of von Neumann equation

$$\dot{\varrho} \;=\; -i[H,\varrho] \;=\; -i(H\varrho - \varrho H)\,.$$

Here H denotes the Hamiltonian of the system in question and we have assumed the convention $\hbar = 1$. This equation is solved by

$$\varrho(t) = U(t)\varrho(0)U^*(t),$$

where the unitary propagator has the form $U(t) = e^{-iHt}$.

Often, when the continuous time dependence of the system state is not the main issue, one resorts to discretized dynamics, using e.g. the "time one" mapping, $\varrho' = U\varrho U^*$. It turns out that general *quantum operations*, providing an adequate mathematical description of complex

[1] More consistently, mixed states should be regarded as elements of the Hilbert-Schmidt *dual* of $\mathcal{B}(\mathcal{H})$, that is linear functionals on $\mathcal{B}(\mathcal{H})$ acting on observables of the system by expectation $\varrho(A) = \operatorname{Tr}(\varrho A)$. For finite-dimensional $\mathcal{H}$ both $\mathcal{B}$ and $\mathcal{B}^*$ are in fact identical with $\mathcal{M}_n(\mathbb{C})$, the algebra of complex $n \times n$ matrices.

multi-stage quantum processes, experiments or computations acting on system states have a more general form of an operator sum

$$\Phi(\varrho) \;=\; \sum K_i \varrho K_i^* \,. \tag{1}$$

These include, for instance, quantum measurements or transmission of states through noisy quantum channels. The above so-called *Kraus representation* is the most general form of a linear *completely positive map* $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$. From the point of view of quantum theory we are interested in the restriction of Φ to the set of density matrices $\Sigma(\mathcal{H})$. Complete positivity of Φ ensures that it preserves positivity of states, while an extra assumption is needed to guarantee the preservation of trace, namely $\sum K_i^* K_i = I$, where I denotes the identity matrix. So, for such Φ we have $\Phi : \Sigma \to \Sigma$. In QI theory such maps represent general quantum communication channels and typical questions studied in this context concern e.g. the effect of Φ on the initial entanglement of the transmitted states, the impact of noise, decoherence, etc. Let us mention also that Kraus representation, though very useful, has the defect of not being unique for a given quantum map Φ.

It should be stressed that quantum operations in the above sense are as a rule nonunitary. Even in the simplest case of Φ represented by two unitary (up to scaling) Kraus terms, $\Phi(\varrho) = U\varrho U^* + V\varrho V^*$, the action of Φ is *not* unitary unless $U = V$ up to a constant factor. However, this is does not pose a contradiction with postulates of quantum mechanics. Let us sketch briefly a typical *open system* scenario leading to nonunitary dynamics.

Suppose that we realistically consider a quantum system not as isolated one, but as remaining in contact with an external bath, so that the underlying Hilbert space has the structure $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_B$, with $\mathcal{H}_S$ and $\mathcal{H}_B$ being respectively the system and the bath spaces. It is natural then to cast the overall Hamiltonian in the following form:

$$H \;=\; H_S \otimes I_B \;+\; I_S \otimes H_B \;+\; H_I \,,$$

where H_S and H_B are the Hamiltonians describing the evolution of the system and bath alone, H_I represents the interaction between them and I_S, I_B are the respective identity operators. While the overall system dynamics is unitary

$$\varrho(t) \;=\; U(t)\varrho(0)U^*(t)\,, \qquad\qquad U(t) = e^{-iHt}\,,$$

it is intractable in such an exact form due to typically huge number of degrees of freedom of the bath. It is then natural to pass to a statistical description of the system evolution by averaging the bath out, assuming in addition that initially the system and the bath are decoupled, that is

$$\varrho_S(t) \;=\; \mathrm{Tr}_B\Big(U(t)\,\varrho_S(0)\otimes\varrho_B(0)\,U^*(t)\Big) \;=\; \sum_\alpha A_\alpha(t)\varrho_S(0)A_\alpha^*(t)\,, \tag{2}$$

where the Kraus operators emerge as $A_\alpha = c_\alpha\langle\beta_i|U|\beta_j\rangle$ with α enumerating index pairs (i,j) and $|\beta_i\rangle$ being the bath basis states. This is clearly a nonunitary evolution unless all A_α are the same up to scalar factors — an unlikely event.

http://dx.doi.org/10.5772/56159

Nevertheless, there may exist a smaller subspace $\mathcal{H}_{DF}$ of $\mathcal{H}_S$ where the reduced dynamics (2) actually *is* unitary. This is equivalent to saying that there exists a basis of $\mathcal{H}_S$ in which all Kraus operators A_α have simultaneously the block form

$$A_\alpha = \left[\begin{array}{c|c} s_\alpha V & 0 \\ \hline 0 & \tilde{A}_\alpha \end{array}\right], \tag{3}$$

where V is unitary on $\mathcal{H}_{DF}$, s_α are scaling factors and $\tilde{A}_\alpha$ are arbitrary operators on $\mathcal{H}_{DF}^{\perp}$, the orthocomplement of $\mathcal{H}_{DF}$ in $\mathcal{H}_S$. Such a space is called *decoherence-free* as the coherent state evolution in this space is isolated from the destructive impact of the bath.

Similarly, one can derive conditions for the existence of a decoherence-free subspace in the framework of Markovian approximation of an open system dynamics, and they turn out to have a form consistent with (3) above. Let us recall that the following master equation in the Gorini-Kosakowski-Sudarshan form provides the most general description of a completely positive Markovian time evolution of a quantum system interacting with its environment [11, 20],

$$\dot{\varrho} \;=\; -i[H,\varrho] \;+\; \frac{1}{2}\sum_{ij} c_{ij}\Big([F_i,\,\varrho F_j^*] + [F_i\varrho,\,F_j^*]\Big)\,, \tag{4}$$

where the sum collects all the terms responsible for nonunitary decohering dynamics. Thus H is the system Hamiltonian, the operators F_i are the so-called error fields and they represent the coupling of the system with its environment, while the hermitian structure matrix $[c_{ij}]$ carries other physically relevant information. Now, if $\mathcal{H}_{DF}$ is to be a decoherence-free subspace, then for any ϱ supported on it the second term in (4) must vanish identically, so that the resulting dynamics is purely unitary. If one assumes certain robustness, or *generic property* in the terminology of [18], of this subspace, meaning that the vanishing of the nonhamiltonian part is not the result of some fine-tuning among structure parameters c_{ij} but rather the effect of simultaneous vanishing of all individual terms, it can be seen that $\mathcal{H}_{DF}$ must be spanned by common eigenvectors of all error fields. In particular, $[F_i, F_j] = 0$ on $\mathcal{H}_{DF}$.

Let us now go back to general quantum operations represented by completely positive trace preserving maps in the form (1). As we have seen, the basic issue in the search for decoherence free subspaces is the identification of common eigenvectors of all Kraus operators K_i and maximal common invariant subspaces spanned by them. For reasons outlined in the introduction, it is impractical to approach this problem by means of direct evaluation of eigenvectors. As a rule, such computations are prone to numerical errors and hence the precise identification of common eigenvectors cannot be achieved this way. In section 5, we will describe an alternative constructive method based on simple linear algebra, the so-called Shemesh criterion, which allows one to identify common invariant subspaces of several operators.

We shall conclude this section by mentioning three more situations where the identification of common invariant subspaces plays a significant role.

1. Characterization of irreducibility of quantum operations, [14, 15]. Irreducible quantum operations (superoperators) appear as a natural generalization of the notion of positive semidefinite irreducible linear operators, treated in particular by Perron-Frobenius theory. The latter provides a very useful and simple characterization of the spectra of irreducible operators. It turns out that if a quantum operation Φ is given in terms of Kraus representation (1), then it is irreducible if and only if the operators K_i do not share a nontrivial invariant subspace. In other words, no decoherence-free subspace exists for an irreducible Φ.
2. Identification of sufficient algebras of observables, [12, 13]. To identify an unknown quantum state ϱ, an experimenter has to perform a number of measurements on the system in question, collecting data that can be used subsequently in the estimation of ϱ. Each of these measurements returns an expectation of the measured observable A_i in the state ϱ, that is the quantity $\mathrm{Tr}(A_i\varrho)$. A natural question that emerges is how to optimize such a data collection, namely how to choose observables A_i to obtain maximum information with the least experimental effort. Sufficiency of an algebra generated by a finite collection of observables $\mathcal{A} = \mathcal{A}(A_1, \ldots, A_p)$ means that the information acquired in the measurement process $\mathrm{Tr}(A_i\varrho)$, $i = 1, \ldots, p$, characterizes the state ϱ completely. One of the rationally verifiable conditions which can be used here is based on Burnside's theorem, which allows one to check whether a given set of observables generates the full matrix algebra $\mathcal{M}_n$ or not. This question can again be related with the existence of a common invariant subspace for the generators of $\mathcal{A}$.
3. Error correcting codes, [6, 17]. This is a more general case than that of the existence of a decoherence-free subspace. Here, one is interested in establishing the existence of a subspace $\mathcal{H}_{EC}$, the subscript EC for *error correcting*, of $\mathcal{H}_S$ on which the action of the channel Φ can be effectively inverted, namely, there exists a quantum operation Θ such that for states ϱ supported on $\mathcal{H}_{EC}$ one has $\Theta(\Phi(\varrho)) = \varrho$. The motivation behind such a demand is that the basis states of $\mathcal{H}_{EC}$ can be regarded as "code words" which can unambiguously be unscrambled after transmission through the generally corrupting channel Φ, and thus they can be used to safely encode portions of information to be sent through the channel. As shown in [17], the necessary and sufficient condition for the existence of an EC subspace for an operation Φ resulting from (2) can be phrased in the following simple algebraic form involving the Kraus operators A_α: there exists a basis of $\mathcal{H}_S$ such that for all α, β

$$A_\alpha^* A_\beta = \left[\begin{array}{c|c} r_{\alpha\beta} I & 0 \\ \hline 0 & \tilde{A}_\alpha^* \tilde{A}_\beta \end{array}\right],$$

where as before $\tilde{A}_\alpha, \tilde{A}_\beta$ are arbitrary operators on $\mathcal{H}_{EC}^{\perp}$ and $R = [r_{\alpha\beta}]$ is a scalar matrix. I in the upper left block is the identity on $\mathcal{H}_{EC}$. Note that the decoherence-free subspace is a special case of an EC space, since then from (3) it follows that the matrix R has a very special form $r_{\alpha\beta} = \bar{s}_\alpha s_\beta$ and therefore has rank 1.

4. Characteristic and minimal polynomials

As we have mentioned in the introduction, the precise determination of eigenvalues of a matrix by means of a finite rational computation is in general impossible. The same is true for

http://dx.doi.org/10.5772/56159

eigenvectors. One can nevertheless rationally acquire exact knowledge about some spectral properties of a matrix, for instance by studying its characteristic and minimal polynomials. Numerous methods for obtaining the polynomials can be found in algebraic literature, and we are going to recall two of them here.

For an $n \times n$ complex matrix A let

$$\chi_A(\lambda) \;=\; \det(\lambda I - A) \;=\; \lambda^n + p_1\lambda^{n-1} + \cdots + p_{n-1}\lambda + p_n$$

be its characteristic polynomial. We will describe the method of undetermined coefficients — an efficient algorithm yielding the numbers p_i. The procedure begins with the evaluation of auxiliary constants

$$D_k \;:=\; \chi_A(k) \;=\; \det(kI - A)\,, \qquad k = 0, 1, \ldots, n-1\,.$$

Next the following system of linear equations in the unknowns $p_1, \ldots, p_n$ is formed

$$\begin{cases} & & & & p_n = D_0 \\ 1^n & + & p_1 1^{n-1} & + \cdots + & p_n = D_1 \\ 2^n & + & p_1 2^{n-1} & + \cdots + & p_n = D_2 \\ \cdots & & & \cdots & \cdots \\ (n-1)^n & + & p_1(n-1)^{n-1} & + \cdots + & p_n = D_{n-1} \end{cases}$$

or equivalently

$$\begin{bmatrix} 1^{n-1} & 1^{n-2} & \cdots & 1 \\ 2^{n-1} & 2^{n-2} & \cdots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ (n-1)^{n-1} & (n-1)^{n-2} & \cdots & n-1 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{n-1} \end{bmatrix} = \begin{bmatrix} D_1 - D_0 - 1^n \\ D_2 - D_0 - 2^n \\ \vdots \\ D_{n-1} - D_0 - (n-1)^n \end{bmatrix}.$$

Writing S_{n-1} for the matrix on the left hand side, the solution can be expressed in compact vector notation as $p = S_{n-1}^{-1} D$. Note that S_{n-1} is a constant matrix whose inverse can be computed and stored beforehand and used repeatedly for various input matrices A. The computational cost is thus limited to the determination of the vector D, and hence is bounded by $O(n^4)$. For comparison, direct expansion expressing the coefficients p_i by the sums of i-th order principal minors of A results in the computation scheme of complexity $O(2^n)$.

The minimal polynomial of a A is defined to be the least degree monic polynomial μ (i.e. with the leading coefficient 1) which annihilates A, $\mu(A) = 0$. Alternatively, it can be given in the form

$$\mu_A(\lambda) \;=\; (\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_k)^{r_k}\,,$$

where λ_i are distinct eigenvalues of A and r_i denotes the order of the largest Jordan block for λ_i in the canonical representation of A. Clearly μ_A divides χ_A.

One obvious direct method consists in checking the sequence of matrices

$$I,\ A,\ A^2,\ \ldots,\ A^r$$

for linear independence, systematically for $r = 1, 2, \ldots$. The least r for which the sequence turns out to be linearly dependent is the degree of the minimal polynomial μ_A, and the respective vanishing linear combination

$$c_r I + c_{r-1} A + \cdots + c_1 A^{r-1} + c_0 A^r \ = \ 0$$

yields, after dividing by c_0, the coefficients of μ_A. This task can be realized by applying Gauss elimination to the $r \times n^2$ matrix whose rows are the reshaped matrices $I, A, A^2, \ldots$, i.e. row vectors obtained by arranging the elements of A^i lexicographicaly row after row. The complexity of such process is $O(n^4)$.

An equivalent method often used in practice is a variant of Krylov subspace algorithm, based on the following classical theorem.

Theorem 1. *For a linear map $A : V \to V$ let $W_1, \ldots, W_k$ be subspaces of V such that*

i) $W_1 + \cdots + W_k = V$, the sum not necessarily being direct,

ii) each W_i is invariant for A,

iii) the restriction $A|_{W_i}$ has minimal polynomial m_i.

Then the minimal polynomial μ_A of A on V is the least common multiple of $m_1, \ldots, m_k$.

The algorithm has the following steps.

1. Pick nonzero $v \in V$ and iteratively compute its Krylov subspace relative to A,

   $$W = \mathrm{Span}\{v, Av, \ldots, A^{d-1}v\}\,.$$

   That is, d is the smallest number such that the vectors $v, Av, \ldots, A^d v$ are linearly dependent, namely

   $$A^d v = c_1 A^{d-1} v + \cdots c_{d-1} Av + c_d v\,.$$

   By construction, the subspace W is invariant for A. It is not difficult to justify that

   $$m(\lambda) \ = \ \lambda^d - c_1 \lambda^{d-1} - \cdots - c_{d-1}\lambda - c_d$$

   is the minimal polynomial of the restriction $A|_W$.

2. Set $W_1 = W$ and $m_1 = m$. If $W_1 = V$ we are done, otherwise pick $v' \notin W_1$ and repeat step 1 to obtain W_2 and m_2 and so on. The construction terminates when $W_1 + W_2 + \cdots + W_k = V$.

3. Find μ_A as the least common multiple of $m_1, \ldots, m_k$. This can be done rationally by using Euclid's algorithm repeatedly to find first GCD of pairs of polynomials m_i.

Most of the computational effort resides here in the construction of Krylov subspaces. For each new vector $A^i v$ added to W linear dependence is checked by Gaussian elimination. Altogether no more than n such checks are performed so the complexity bound is $O(n^4)$.

Let us conclude this section by mentioning some exemplary problems in quantum physics, where knowledge of spectral and minimal polynomials plays a role. Firstly, it is the design of optimal setups for stroboscopic tomography of states [12, 13]. Namely, one has to find a minimal set of observables and design a stroboscopic measurement, i.e. one performed at preselected time instants when the measured observables are subdued to time evolution, the objective being to collect information sufficient for the complete reconstruction of a quantum state with least experimental effort. To this end, Krylov subspaces of the observables relative to the generator of the dynamics have to be constructed. The degree of the minimal polynomial of the dynamics generator is one of the essential parameters appearing in the design process.

Second set of examples is related to the construction of common invariant subspaces for families of operators, which finds application e.g. in the identification of decoherence-free subspaces in open quantum systems. This problem will be discussed in detail in the next section. It turns out that the construction of such common invariant subspaces can be simplified considerably if one of the operators has nondegenerate spectrum. The former property can be tested for an operator A by analyzing the GCD of its characteristic polynomial and its derivative: the eigenvalues are simple iff χ_A and χ'_A are relatively prime. To detect diagonalizability, one has to perform a similar test on the minimal polynomial of A. An alternative for the Euclidean GCD algorithm is the singularity test of the so-called associated Sylvester matrix [27].

5. Common invariant subspaces

The problem we are going to discuss now in its simplest version can be formulated as follows: given two square matrices $A, B \in \mathcal{M}_n$ decide whether they have an eigenvector in common. We are interested, of course, in finite *rational* procedures solving this problem. As it was indicated in the introduction, naive direct approach by literally finding the eigenspaces of A and B and comparing them is useless because of finite accuracy of numerics. We will be concerned with a more general formulation of the problem, namely we will ask whether two matrices share an invariant subspace of dimension k and how to find such subspace.

In what follows, we will discuss certain finite rational computational procedures detecting the existence of common invariant subspaces for pairs of operators. There are no known direct generalizations of such procedures to work for more than two operators at a time. However, if one can constructively obtain common invariant subspaces for all pairs of operators in the set $A_1, \ldots, A_p$, then taking their intersection one obtains a candidate for the global solution. It has to be verified though, because the resulting space need not be invariant for some (or any!) of the operators A_i. The computational complexity of such a construction will add a factor p^2 to that of the process performed for a single pair of operators. The intersection of p^2 subspaces of dimensions bounded by n can be constructed in time bounded by $p^2 n^3$.

5.1. Shemesh criterion and related methods

The basic tool in the detection of common invariant subspaces is the so-called Shemesh criterion [24]. We use here the standard notation $[A, B]$ for the commutator of matrices A and B.

THEOREM 2 (Shemesh 1984). *Matrices $A, B \in \mathcal{M}_n$ possess a common eigenvector if and only if the subspace*

$$\mathcal{N} = \bigcap_{k,l=1}^{n-1} \ker\left[A^k, B^l\right] \tag{5}$$

is of positive dimension. Moreover, $\mathcal{N}$ is invariant with respect to both A and B and restrictions of A and B to $\mathcal{N}$ commute. Every common invariant subspace of A and B (on which they commute) is contained in $\mathcal{N}$.

Let us remark that n above can be replaced by r and s — the degrees of minimal polynomials of A and B, respectively.

We shall analyze now the complexity of a direct method of checking Shemesh criterion and that of constructing $\mathcal{N}$ — the maximal common invariant subspace of A and B. Let us stress here that while the existence of a 1-dimensional common invariant subspace (corresponding to the common eigenvector of A and B) in $\mathcal{N}$ is guaranteed by the criterion, it *does not* answer any questions concerning k-dimensional common invariant subspaces, $2 \leq k < \dim \mathcal{N}$, not to mention the problem of constructing them by finite rational procedures. Such procedure can be nevertheless easily obtained for the space $\mathcal{N}$. Let us also indicate that no finite rational method should be expected to yield the common eigenvector in $\mathcal{N}$. If there were one, we would have a finite method to compute exactly the corresponding eigenvalues of A and B which is, in general, unfeasible.

To estimate the complexity of Shemesh's criterion, let us first note that computing the commutator $[A, B]$ has the same complexity as matrix multiplication[2], namely $O(n^3)$. The number of commutators to evaluate in (5) is at most $(n-1)^2$, so that the total amount of algebra is bounded here by $O(n^5)$. Finally, finding the intersection of kernels can be done just by means of solving the system of homogeneous linear equations in n variables given by the $n(n-1)^2 \times n$ matrix

$$\begin{bmatrix} [A, B] \\ [A, B^2] \\ \vdots \\ [A^{n-1}, B^{n-1}] \end{bmatrix}. \tag{6}$$

This is achieved by the Gaussian elimination algorithm again in $O(n^5)$ steps, hence the overall complexity of finding $\mathcal{N}$ is $O(n^5)$.

[2] Of course, one can always lower the exponent 3 to some extent by resorting to fast matrix multiplication schemes. This may be of practical importance when working with large matrices, here however we are mainly interested in establishing just *polynomial* complexity of our procedures.

http://dx.doi.org/10.5772/56159

An equivalent formulation of Shemesh condition ($\dim \mathcal{N} > 0$) is that

$$\det \sum_{k,l=1}^{n-1} [A^k, B^l]^* \cdot [A^k, B^l] \;=\; 0$$

but it does not simplify the computation as the sum above involves $(n-1)^2$ terms, each one computable with the arithmetic cost of $O(n^3)$ operations.

Let us turn to a more complicated problem of verifying the existence of a common invariant subspace of prescribed dimension $2 \leq k < n$. This is partly solved by applying the Shemesh criterion to exterior powers (wedge powers) of A and B. Recall that $A^{\wedge k}$ is the restriction of $A^{\otimes k}$ to the antisymmetric subspace of $\left(\mathbb{C}^n\right)^{\otimes k}$. More explicitly, $A^{\wedge k}$ is an $m \times m$ matrix with $m = \binom{n}{k}$, the elements of which are

$$\left(A^{\wedge k}\right)_{\alpha,\beta} \;=\; \det A[\alpha|\beta]\,,$$

where α and β stand for multi-indices $\alpha = (i_1, i_2, \ldots, i_k)$, with $1 \leq i_1 < i_2 < \cdots < i_k \leq n$. $A[\alpha|\beta]$ is a $k \times k$ submatrix of A with rows and columns specified by α and respectively β. The space $\mathcal{N}_k$ corresponding to $\mathcal{N}\,(= \mathcal{N}_1)$ in (5) is now defined by analogy as

$$\mathcal{N}_k \;=\; \bigcap_{i,j=1}^{m-1} \ker\left[\left(A^{\wedge k}\right)^i,\, \left(B^{\wedge k}\right)^j\right]. \tag{7}$$

The trick of using exterior algebra takes advantage of a simple fact that if $\lambda_1, \ldots, \lambda_k$ are eigenvalues of A with (linearly independent) eigenvectors $v_1, \ldots, v_k$ then $\lambda_1\lambda_2\cdots\lambda_k$ is an eigenvalue of $A^{\wedge k}$ with eigenvector $v_1 \wedge \cdots \wedge v_k$. So if $v_1, \ldots, v_k$ span an invariant k-dimensional subspace of A and B then obviously $v_1 \wedge \cdots \wedge v_k$ is a common eigenvector of $A^{\wedge k}$ and $B^{\wedge k}$. The corresponding sufficient condition, however, turns our to be more complicated. Nontriviality of $\mathcal{N}_k$ guarantees the existence of an eigenvector shared by $A^{\wedge k}$ and $B^{\wedge k}$ but it is now an object in the exterior algebra of $\mathbb{C}^n$ and, in general, it need not be decomposable, i.e. of pure product form $v = v_1 \wedge \cdots \wedge v_k$. Consequently the reconstruction of a k-dimensional common invariant subspace of A and B from v may no longer be easy if at all possible. The source of this difficulty resides in the fact that the spectrum of $A^{\wedge k}$ or $B^{\wedge k}$ may be degenerate. This possibility has to be, therefore, excluded by an additional assumption. As we will see shortly, such an assumption can be further relaxed to another one postulating the nondegeneracy of eigenvalues of either A or B alone.

The generalized Shemesh criterion [9] takes the following form.

Theorem 3 (Generalized Shemesh Criterion).

Necessity: *If A and B have a common invariant subspace of dimension $2 \leq k < n$, then $\mathcal{N}_k$ as defined in (7) has positive dimension (i.e. $A^{\wedge k}$ and $B^{\wedge k}$ share an eigenvector).*

Sufficiency: *Suppose that $A^{\wedge k}$ has nondegenerate eigenvalues and $\det B \neq 0$. Then if $\mathcal{N}_k \neq \{0\}$, there exists a common k-dimensional invariant subspace of A and B.*

In order to show how one can simplify the extra conditions in the sufficiency part of the above theorem, let us note that for an arbitrary matrix C the spectral shift transformation $C \mapsto C_t = C - tI$ does not alter its invariant subspaces. The following two facts proved in [9] allow one to preprocess, if necessary, the initial matrices A and B so that the extra requirements are fulfilled, at the same time leaving their invariant subspaces intact.

Fact 1. For any singular complex matrix B, a shift $t \in \mathbb{N}$ can be computed by a finite rational procedure so that $\det(B - tI) \neq 0$.

The procedure is very simple: it computes $\det(B - tI)$ for $t = 1, 2, \ldots$ until a nonzero value is found. Since the characteristic polynomial of B has no more than n distinct roots, the computation must terminate in no more than $n - 1$ steps.

Fact 2. If all eigenvalues of A are nondegenerate and $2 \leq k < n$, then a shift $t \in \mathbb{N}$ can be computed by a finite rational procedure so that the matrix $(A - tI)^{\wedge k}$ also has only simple eigenvalues.

See [9] for the proof of Fact 2. Its essence is that one can probe subsequent values of the shift parameter $t = 0, 1, \ldots$ until nondegeneracy of eigenvalues occurs, which is shown to happen after no more than $\frac{1}{2}kn^{2k}$ of such tests.

We are equipped now to describe the complete algorithm determining the existence of k-dimensional invariant subspace common to A and B. Let ϕ_A denote the characteristic polynomial of A.

1. Check whether A has distinct eigenvalues by computing the resultant of ϕ_A and ϕ'_A (as we have mentioned in Section 3, this can be done conveniently by expressing it as the determinant of the Sylvester matrix [27] of ϕ_A and ϕ'_A) and checking whether it is nonzero. If the test fails for A, try the same for B and switch A and B if B has simple eigenvalues. If both tests fail, the generalized Shemesh criterion cannot be used.
2. If B is singular, apply the spectral shift t as in Fact 1. Replace B with $B - tI$.
3. Compute the matrix $A^{\wedge k}$ and check whether it has nondegenerate eigenvalues (see step 1). If so, go to step 4, otherwise apply the spectral shift to A as described in Fact 2 and repeat step 3.
4. Compute $B^{\wedge k}$ and $\mathcal{N}_k$ as in (7). If $\mathcal{N}_k$ has positive dimension, then A and B have common k-dimensional invariant subspace.

It should be stressed again that Shemesh criterion yields a "yes/no" answer about the *extistence* of a common eigenvector (or, respectively, of k-dimensional common invariant subspace), but does not help in *constructing* them.

The complexity of the above algorithm is determined by n and k. The most time-consuming operations are those performed on the exterior powers of A and B because of their size $m = \binom{n}{k}$, which grows roughly like n^k for $k = 2, \ldots, \left[\frac{n}{2}\right]$. To obtain $A^{\wedge k}$, one has to evaluate m^2 minors of A of size $k \times k$, hence the computational cost is bounded by $O(k^3\, n^{2k})$. Checking for nondegeneracy of eigenvalues of $A \in \mathcal{M}_n$ costs as much as the evaluation of ϕ_A, which can be done in $O(n^3)$ steps, plus the cost of computing the $(2n-1) \times (2n-1)$ determinant of the respective Sylvester matrix, so its overall complexity is $O(n^3)$. Step 3 of the algorithm

http://dx.doi.org/10.5772/56159

probing possible shift parameters performs no more than $O(km^2)$ of nondegeneracy tests, each at the expense of $O(m^3)$ arithmetic operations. Therefore the complexity of step 3 evaluates to $O(kn^{5k})$. Finally, the complexity of constructing $\mathcal{N}_k$ by (7) is, as shown before, $O(m^5)$ or in terms of n and k $O(n^{5k})$.

The estimation above shows that even for small values of k, although of polynomial time complexity, the method is not very practical. Already for $k = 2$, the computational effort is of the order $O(n^{10})$ in the worst case.

Let us mention one more recent result [16] which shows that the nondegeneracy condition can in fact simplify the original Shemesh criterion, slightly reducing its computational complexity.

Theorem 4 (Jamiołkowski, 2012). *Let A have only simple eigenvalues. Then the formula (5) for the space $\mathcal{N}$ in the original Shemesh criterion can be simplified to*

$$\mathcal{N} \;=\; \bigcap_{k=1}^{n-1} \ker\left[A^k, B\right] \tag{8}$$

which reduces the complexity of its construction to $O(n^4)$.

Indeed, the number of commutators to evaluate in (8) is now at most $n-1$, $O(n^3)$ arithmetic operations each, and the system of homogeneous equations defining $\mathcal{N}$ is of the size $n(n-1) \times n$, so the complexity of solving it is also $O(n^4)$.

As the sufficiency part of the generalized Shemesh condition requires the nondegeneracy of the spectrum of $A^{\wedge k}$, so the formula (7) automatically simplifies analogously to

$$\mathcal{N}_k \;=\; \bigcap_{i=1}^{m-1} \ker\left[\left(A^{\wedge k}\right)^i,\, B^{\wedge k}\right]. \tag{9}$$

Hence the complexity of finding $\mathcal{N}_k$ reduces to $O(m^4)$, that is $O(n^{4k})$.

Let us note, however, that somewhat weaker assumption of diagonalizability of A does not, in general, lead to a simplification of the Shemesh formula by limiting the number of commutators that have to be computed. This is illustrated by the following simple example. Let $\{e_1, \ldots, e_4\}$ be a basis in which A and B have the following form:

$$A \;=\; \begin{bmatrix} 1 & 1 & & \\ 1 & 2 & & \\ & & 3 & \\ & & & 3 \end{bmatrix}, \qquad B \;=\; \begin{bmatrix} 2 & 1 & & \\ & 2 & 1 & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix},$$

where we have suppressed all zero entries. Note that A is diagonalizable with twofold degenerate eigenvalue 3. Its minimal polynomial has degree 3. Hence

$$\ker[A,B] \cap \ker[A^2,B] \;=\; \ker[A,B] \cap \ker[A^2,B] \cap \ker[A^3,B] \;=\; \mathrm{Span}\{e_4\}$$

but

$$\ker[A,B] \cap \ker[A^2,B] \cap \ker[A,B^2] \;=\; \{\mathbf{0}\}\,.$$

In the next subsection we will explore an alternative approach based on the so-called polynomial identities for matrix algebras.

5.2. Algebraic approach — polynomial identities

In algebra, polynomial identities are used to characterize various algebraic structures. We will limit the exposition to a necessary minimum so as to make the present text self-contained, focusing on applications to common invariant subspace problems.

Definition 3. *An algebra $\mathcal{A}$ is said to be a polynomial identity algebra (a PI-algebra for short) if there exists a polynomial $P(x_1, x_2, \ldots, x_k)$ over the ring of integers in noncommuting variables x_i such that $P(A_1, A_2, \ldots, A_k) = 0$ for all k-tuples of the elements A_i of $\mathcal{A}$.*

For example, a commutative algebra $\mathcal{A}$ is a PI-algebra with the polynomial $Q_2(x_1, x_2) = x_1x_2 - x_2x_1$. It turns out that special role is played by the so-called standard polynomials which are natural generalizations of Q_2,

$$Q_n(x_1, \ldots, x_n) \;=\; \sum_{\sigma \in S_n} \mathrm{sign}(\sigma)\, x_{\sigma(1)} \cdots x_{\sigma(n)}\,, \tag{10}$$

where the summation extends over the symmetric group S_n. Their importance is exemplified by the Amitsur-Levitzki theorem on matrix algebras $\mathcal{M}_n$.

Theorem 5 (Amitsur-Levitzki 1950). *The full algebra $\mathcal{M}_n(\mathbb{C})$ satisfies the standard polynomial identity of degree $2n$,*

$$Q_{2n}(A_1, \ldots, A_{2n}) \;\equiv\; 0 \qquad \forall\, A_1, \ldots, A_{2n} \in \mathcal{M}_n\,,$$

but it does not satisfy any polynomial identity of smaller degree.

In order to make a connection with the problem of common invariant subspaces, let us first observe that if two matrices A and B share such a subspace W, then W is also invariant for the entire algebra $\mathcal{A}(A,B) \subset \mathcal{M}_n$ generated by A and B. In what follows we shall denote this algebra by $\mathcal{A}$ for simplicity. So according to the Shemesh criterion (5), $\mathcal{A}$ restricted to $\mathcal{N}_1$ satisfies the standard polynomial identity $Q_2 \equiv 0$, that is

$$(C_1C_2 - C_2C_1)v = 0\,, \qquad \forall\, C_1, C_2 \in \mathcal{A}, \quad \forall\, v \in \mathcal{N}_1.$$

http://dx.doi.org/10.5772/56159

Following [2] let us define the family of subspaces

$$\mathcal{N}_k \;=\; \bigcap \ker\left[Q_{2k}(C_1,\ldots,C_{2k})C_{2k+1}\right]\,, \tag{11}$$

where the intersection extends over all $(2k+1)$-tuples of elements $C_i \in \mathcal{A}$. It turns out that $\mathcal{A}$ restricted to $\mathcal{N}_k$ analogously obeys the identity $Q_{2k} \equiv 0$. Of course, this is an interesting property provided that $\mathcal{N}_k$ is not just the zero space.

Theorem 6. *If $\mathcal{N}_k$ of (11) is nontrivial, then it is an invariant subspace for $\mathcal{A}$ and this algebra restricted to $\mathcal{N}_k$ satisfies the standard polynomial identity $Q_{2k} \equiv 0$, that is*

$$Q_{2k}(C_1,\ldots,C_{2k})v = 0\,, \qquad \forall\, C_1,\ldots,C_k \in \mathcal{A}, \quad \forall\, v \in \mathcal{N}_k.$$

Any other invariant subspace of $\mathcal{A}$ on which this algebra satisfies the identity $Q_{2k} \equiv 0$ is contained in $\mathcal{N}_k$.

The proof can be found e.g. in [2]. The usefulness of this theorem can be appreciated by noting that for subsequent values of k we obtain a filtration

$$\{0\} \;=\; \mathcal{N}_0 \;\subset\; \mathcal{N}_1 \;\subset\; \cdots \;\subset\; \mathcal{N}_n \;=\; \mathbb{C}^n\,,$$

which can yield partial answers to questions concerning invariant subspaces of specific dimension. We stress here that each of $\mathcal{N}_k$ can be constructed by a finite rational procedure. Namely, because of linearity of Q_{2k} with respect to each individual variable, to find $\mathcal{N}_k$ it suffices to make each C_i in the intersection (11) run independently through the elements of a fixed basis of $\mathcal{A}$.

The basis itself can be found by the following general procedure [1]. Consider finite products of A and B, e.g. AB^2AB (called words over $\{A,B\}$) in lexicographic order:

$$I,\; A,\; B,\; A^2,\; AB,\; BA,\; B^2,\; A^3,\; A^2B,\; \ldots$$

Words of a fixed length k form the k-th layer in this sequence. I alone forms here the zeroth layer. Let $\mathcal{A}_k$ be the subspace of $\mathcal{M}_n$ spanned collectively by the layers $0 \le j \le k$. Obviously,

$$\mathcal{A}_0 \subset \mathcal{A}_1 \subset \cdots \subset \mathcal{A}_p = \mathcal{A}_{p+1}$$

for some p, the symbol $\subset$ denoting here the proper inclusion. Then $\mathcal{A}_p = \mathcal{A}(A,B)$ and the first $p+1$ layers form the spanning set for $\mathcal{A}$.

To discuss the complexity of this procedure, note first that an obvious rough bound for p is $p \le n^2 - 1$, while there are various better estimates known in literature, see e.g. [10, 21, 22], especially when some knowledge about A and B is available. In particular, if A and B commute, then $p < n$, while the best general bound so far is that due to Pappacena [21],

$$p \;\;\le\;\; n\sqrt{\frac{2n^2}{n-1}+\frac{1}{4}}+\frac{n}{2}-2 \;\;\sim\;\; O(n^{3/2})\,.$$

However, bad news is that the k-th layer contains 2^k words, so to construct $\mathcal{A}_k$ one has to take account of about 2^{k+1} words. Then unless p turns out to be much smaller than n, we are inevitably running here into the domain of nonpolynomial time complexity. So layers are huge while the dimensions of subspaces $\mathcal{A}_k$ are small, not exceeding n^2, and consequently most of the new words from the k-th layer added in the process of forming $\mathcal{A}_k$ will turn out linearly dependent with respect to the earlier processed ones. Yet p saturating the sequence of inclusions of $\mathcal{A}_p$ may very well be comparable with n or even worse than that. The check whether the next added word increases the dimension of $\mathcal{A}_k$ can itself be done by a Gaussian elimination algorithm at polynomial cost.

Let us analyze in turn the complexity of computing $\mathcal{N}_k$ by (11) under the assumption that a basis of $\mathcal{A}$ is given. Similarly as in the case of exterior-algebraic approach described previously, the time complexity here depends critically on k. Firstly, the number of terms in the standard polynomial Q_{2k} grows very rapidly being equal $(2k)!$. Secondly, as indicated above, the intersection in (11) has to extend over all $(2k+1)$-tuples of d basis elements of $\mathcal{A}$, where $d = \dim \mathcal{A}$. Hence the number of terms to account for is d^{2k+1}, which in the worst case of $d \sim n^2$ is of the order of $O(n^{4k+2})$. For $k = 2$ it is $O(n^{10})$. We can see again that such a direct method of construction of $\mathcal{N}_k$ can be carried out in practice only for small k. It is a separate and interesting issue to explore to what extent can prior knowledge of some properties of A and B simplify the computation of $\mathcal{N}_k$. For instance, the nondegeneracy of spectra of A or B can be expected to help.

In the discussion of consequences of Theorem 6 the following two corollaries can be immediately formulated:

1. *If W is an invariant subspace of $\mathcal{A}$ such that* $\dim W \leq k$, *then it is necessarily contained in* $\mathcal{N}_k$.
2. *$\mathcal{A}$ has a nontrivial invariant subspace with dimension not exceeding k iff* $\mathcal{N}_k \neq \{0\}$.

While this constitutes some improvement over the previous exterior-algebraic treatment of the existence of k-dimensional common invariant subspaces, the very question for a fixed value of k cannot be fully answered on the basis of Theorem 6 alone. Let us mention here only, without going into details which prove to be quite technical in this case, some more results addressing this issue. In [9] the complete solution for $k = 2$ is given and it is indicated that in the case of semisimple algebras $\mathcal{A}$ there is a complete rational solution for of the problem for any $1 < k < n$. In [3], the following theorem is proved.

Theorem 7. *Let $\mathcal{A} = \mathcal{A}(A, B)$ be a semisimple algebra. Then $\mathcal{A}$ has an irreducible*[3] *invariant subspace of dimension k iff* $\dim \mathcal{N}_{k-1} < \dim \mathcal{N}_k$.

Moreover, this result is further extended to arbitrary algebras by means of restricting the analysis to the so-called socle of $\mathcal{A}$, which is the maximal invariant subspace Λ of $\mathcal{A}$ such that the restriction $\mathcal{A}|_\Lambda$ is a semisimple algebra. Hence one can use Theorem 7 for $\mathcal{A}|_\Lambda$. Then, since Λ can be shown to contain all irreducible invariant subspaces of $\mathcal{A}$, the solution turns out to be valid also for the original algebra $\mathcal{A}$.

[3] W is an irreducible invariant subspace of $\mathcal{A}$ if the restriction of $\mathcal{A}$ to W coincides with entire $\mathcal{M}_k$, where k is the dimension of W.

http://dx.doi.org/10.5772/56159

It should be noted that for a finite-dimensional algebra $\mathcal{A}$ checking it for semisimplicity as well as the construction of the socle of $\mathcal{A}$ can all be done by finite rational procedures. They can be reduced to a Gaussian elimination on a $d \times d$ matrix, where $d = \dim \mathcal{A}$. Here again we assume that some basis of $\mathcal{A}$ is given, for otherwise we run into the intractable problem of constructing it.

Finally, let us point the reader to yet another approach [4] discussing a solution of the common invariant subspace problem in the language of algebraic geometry and Gröbner bases.

5.3. The application of Burnside's theorem

Let us begin with the formulation of the theorem.

Theorem 8 (Burnside). *Any subalgebra $\mathcal{A}$ of $\mathcal{M}_n(\mathbb{C})$ whose only invariant subspaces are $\{0\}$ and $\mathbb{C}^n$ is necessarily equal to $\mathcal{M}_n(\mathbb{C})$.*

This result can be used to verify whether a given set of operators generates the whole matrix algebra $\mathcal{M}_n$, so it has natural application in analyzing sufficiency of various sets of observables. Let us also note that the question of irreducibility of a quantum operation Φ is equivalent to saying that the collection of Kraus operators for Φ (1) generates $\mathcal{M}_n$.

When $\mathcal{A} = \mathcal{A}(A, B)$, then Shemesh criterion is the tool that can be used directly to verify the assumption in Burnside's theorem: if $\mathcal{N} = \{0\}$ then $\mathcal{A}(A, B) = \mathcal{M}_n$. Suppose in turn that the algebra $\mathcal{A}$ is generated by more than two operators, $\mathcal{A} = \mathcal{A}(A_1, \ldots, A_p)$. We can adopt the following strategy.

1. Compute Shemesh kernels $\mathcal{N}(A_i, A_j)$ for all pairs of operators.
2. Find the intersection $\Lambda_1 = \bigcap_{i,j} \mathcal{N}(A_i, A_j)$. If $\Lambda_1 = \{0\}$, then $\mathcal{A} = \mathcal{M}_n$, otherwise continue to step 3.
3. Replace the operators A_i with their restrictions to Λ_1, $A_i := A_i|_{\Lambda_1}$ and carry on steps 1 and 2 to obtain Λ_2. If $\Lambda_2 = \Lambda_1$, then Λ_2 is the nontrivial invariant subspace of $\mathcal{A}$ and consequently $\mathcal{A} \neq \mathcal{M}_n$. Otherwise iterate 3 with Λ_2 in place of Λ_1 to obtain Λ_3 and so on.

Clearly we have

$$\Lambda_1 \supset \Lambda_2 \supset \cdots \supset \{0\},$$

so either all the inclusions above are proper and after a finite number of iterations we must end up with $\Lambda_t = \{0\}$, or $\Lambda_t = \Lambda_{t+1} \neq \{0\}$ for some t. Hence this procedure terminates. Let us estimate its complexity. There are $\binom{p}{2} \sim p^2$ kernels to compute in step one, so its cost is bounded by $O(p^2 n^5)$. The construction of Λ_1 can be realized iteratively with the use of Gauss elimination at the total cost of at most $O(p^2 n^3)$. Finally, the number of iterations of step 3 is bounded by the dimension of $\mathcal{A}$, that is by n^2. Consequently the upper bound on the complexity of the entire procedure is $O(p^2 n^7)$.

6. Conclusions

We have seen that some rational computational procedures, while very useful for quantum information theoretic analyses, have nonpolynomial time complexity which in principle disqualifies them from practical applications. The polynomial complexity bounds obtained for procedures using the Shemesh criterion may also look somewhat pessimistic, yet they are certainly crude and we believe there is plenty of room for improvement if one uses some extra knowledge about the operators taking part in the computation. There is an apparent need for efficient algorithms for the construction of bases of finite-dimensional algebras — without such methods many of the procedures discussed here cannot be carried out efficiently. It is possible that some efficient Monte Carlo methods could be designed for such a class of problems. Such situation is not uncommon in computational algebra, as many of its problems belong to the BPP class. We hope to address some of these issues in future research.

Author details

Miłosz Michalski

Institute of Physics, Nicolaus Copernicus University, Toruń, Poland

References

[1] Alpin Yu. A., L. Elsner, K. D. Ikramov, Linear Algebra & Appl. **306** (2000), 165–182.

[2] Alpin Yu. A., A. George, K. D. Ikramov, Linear Algebra & Appl. **312** (2000), 115–123.

[3] Alpin Yu. A., K. D. Ikramov, J. Math. Sci. **114**:6 (2003), 1757–1764.

[4] Arapura D., C. Peterson, Linear Algebra & Appl. **384** (2004), 1–7.

[5] Chistov A., H. Fournier, L. Gurvits, P. Koiran, Found. Computational Math. **3**:4 (2003), 421–427.

[6] Choi M. D., D. W. Kribs, K. Życzkowski, Rep. Math. Phys. **58** (2006), 77–91.

[7] Cook S. A., *The complexity of theorem proving procedures*, Proc. 3rd Annual ACM Symposium on the Theory of Computing, ACM, New York, 1971, 151–158.

[8] Glynn D. G., European J. of Combinatorics **31**(2010), 1887–1891.

[9] George A., K. D. Ikramov, Linear Algebra & Appl. **287** (1999), 171–179.

[10] Gerstenhaber M., Ann. Math. **73** (1961), 324–348.

[11] Gorini V., A. Kossakowski, E. C. G. Sudarshan, J. Math. Phys. **17** (1976), 821–825.

[12] Jamiołkowski A., Rep. Math. Phys. **46**:3 (2000), 469–482.

[13] Jamiołkowski A., *On sufficient algebraic conditions for identification of quantum states*, in: "Quantum Bio-Informatics IV", L. Accardi, W. Freudenberg, M. Ohya (Eds.), World Scientific, Singapore, 2011, 185–197.

[14] Jamiołkowski A., Int. J. Geometric Methods in Modern Physics **9**:2 (2012), 1260014.

[15] Jamiołkowski A., *On applications of PI-algebras in the analysis of quantum channels*, this volume.

[16] Jamiołkowski A., private communication.

[17] Knill E., R. Laflamme, Phys. Rev. A **55** (1997), 900–911.

[18] Lidar D. A., I. L. Chuang, K. B. Whaley, Phys. Rev. Lett. **81** (1998), 2594–2597.

[19] Lidar A. A., D. Bacon, K. B. Whaley, Phys. Rev. Lett. **82** (1999), 4556–4559.

[20] Lindblad G., Comm. Math. Phys. **48** (1976), 119–130.

[21] Pappacena C. J., J Algebra **197** (1997), 535–545.

[22] Paz A., Linear & Multilinear Algebra **15** (1984), 161–170.

[23] Ryser H. J., *Combinatorial Mathematics*, The Carus Mathematical Monographs 14, Wiley, N.Y., 1965.

[24] Shemesh D., Linear Algebra Appl. **62** (1984) 11–18.

[25] Valiant L. G., Theoret. Comp. Sci. **8**:2 1979, 189–201.

[26] See e.g. `http://en.wikipedia.org/wiki/Wang_tile`

[27] See e.g. `http://en.wikipedia.org/wiki/Sylvester_matrix`

IntechOpen