Quantum diagonalization of Hermitean matrices

Stefan Weigert

Institut de Physique, Université de Neuchâtel, Neuchâtel, Switzerland
and
Department of Mathematics, University of Hull, Hull HU6 7RX, UK

E-mail: stefan.weigert@unine.ch

Received 18 January 2001
Published 29 June 2001
Online at stacks.iop.org/JPhysA/34/5619

Abstract
To measure an observable of a quantum mechanical system leaves it in one of its eigenstates and the result of the measurement is one of its eigenvalues. This process is shown to be a computational resource: Hermitean \((N \times N)\) matrices can be diagonalized, in principle, by performing appropriate quantum mechanical measurements. To do so, one considers the given matrix as an observable of a single spin with appropriate length \(s\) which can be measured using a generalized Stern–Gerlach apparatus. Then, each run provides one eigenvalue of the observable. As the underlying working principle is the ‘collapse of the wavefunction’ associated with a measurement, the procedure is neither a digital nor an analogue calculation—it defines thus a new example of a quantum mechanical method of computation.

PACS numbers: 0367, 0365
known classical algorithm [3]. Throughout, these techniques make use of the measurement of quantum mechanical observables as an unquestioned tool. This is also true for many (but not all [4]) proposals of quantum error correction schemes [5, 6], necessary to let a potential algorithm run.

Here the purpose is to point out that the bare ‘projection’ [2] effected by a quantum mechanical measurement does possess computational power. As shown below, it can be used to solve explicitly at least one specific computational task, namely to determine eigenvalues and, a fortiori, eigenstates of Hermitian \((N \times N)\) matrices.

The diagonalization of Hermitian matrices is a recurrent problem in mathematics, physics, and related fields. Using the notation of a quantum physicist the problem reads as follows.

Given a self-adjoint operator \(\hat{A}\) acting on a Hilbert space \(\mathcal{H}\) of dimension \(N\), one needs to determine its eigenstates \(|a_n\rangle, n = 1, \ldots, N\), and its \(N\) real eigenvalues \(\lambda_n\), satisfying \(\hat{A}(|a_n\rangle) = \lambda_n|a_n\rangle\). If normalized to one, the eigenstates constitute a complete orthonormal basis of the space \(\mathcal{H}\): \(\sum_{n=1}^{N} |a_n\rangle\langle a_n| = 1\) \(\langle a_n|a_m\rangle = \delta_{nm}\). The standard solution from linear algebra [7] is to write down the eigenvalue equation with respect to a given orthonormal basis \(|k\rangle, k = 1, \ldots, N\), say. The \(N^2\) matrix elements \(A_{kl} = \langle k|\hat{A}|l\rangle\) determine the operator \(\hat{A}\) uniquely and its eigenstates are characterized by the coefficients \((A_n)_k = A_{nk}\) in the expansion \(|a_n\rangle = \sum_k A_{nk}|k\rangle\). The number \(\lambda\) is an eigenvalue of \(\hat{A}\) if the characteristic polynomial \(P_A(\lambda)\) of the matrix \(A\) vanishes, \(P_A(\lambda) = \det(A - \lambda I) = 0\), where \(I\) is the \((N \times N)\) unit matrix. Once the \(N\) roots \(\lambda_n\) of the polynomial \(P_A\) are known, the non-zero solutions of the equation

\[
(A - \lambda_n I) \hat{A} = 0 \quad n = 1, \ldots, N
\]  

provide the eigenvectors \(|a_n\rangle\) in the basis \(|k\rangle\). Analytic expressions for the eigenvalues \(\lambda_n\) in terms of the matrix elements \(A_{kl}\) exist only if \(N \leq 4\). In general, numerical methods are required to determine approximately the roots of \(P_A(\lambda)\).

The quantum diagonalization of Hermitian matrices is based on the assumption that the behaviour of a spin \(s\) (of a particle) is described correctly by non-relativistic quantum mechanics. This method exploits the ‘collapse of the wavefunction’ as a computational resource. Note that the procedure does not depend on a particular interpretation of quantum mechanics. Five steps are necessary to achieve the diagonalization of a given matrix \(A\) (supposed for simplicity not to have degenerate eigenvalues). The individual steps will be described in a condensed form first; subsequently, five comments explain various technical details.

**Step 1. Standard form of \(A\).** Write the Hermitian \((N \times N)\) matrix \(A\) as a combination of linearly independent Hermitian multipole operators \(T_v\), \(v = 0, \ldots, N^2 - 1\),

\[
A = \sum_{v=0}^{N^2-1} a_v T_v \quad a_v = \frac{1}{N} \text{Tr}[A T_v] \in \mathbb{R}.
\]  

**Step 2. Identification of an observable.** Interpret the matrix \(A\) as an observable \(H_A\) for a single quantum spin \(S\) with quantum number \(s = (N - 1)/2\),

\[
H_A(S) = \sum_{v=0}^{N^2-1} a_v T_v(S)
\]  

using the expression of the multipoles \(T_v(S)\) in terms of the components of a spin.

**Step 3. Setting up a measuring device.** Identify and construct an apparatus \(\text{app}(H_A)\) suitable to measure the observable \(H_A\).

**Step 4. Determination of the eigenvalues.** Carry out measurements with the apparatus \(\text{app}(H_A)\) on a spin \(s\) prepared in a homogeneous mixture \(\hat{\rho} = I_s/(2s + 1)\). The output of each individual
measurement will be one of the eigenvalues \(a_n\) of the matrix \(A\). After sufficiently many repetitions, all eigenvalues will be known.

**Step 5. Determination of the eigenstates.** Calculate the eigenstates \(|a_n\rangle\) of the matrix \(A\) by means of equation (1) and the experimentally determined eigenvalues \(a_n\). Alternatively, determine the eigenstates \(|a_n\rangle\) experimentally by methods of state reconstruction.

Thus, the matrix \(A\) has been diagonalized without calculating the zeros of its characteristic polynomial by traditional means. The fourth step solves the hard part of the eigenvalue problem since it provides the eigenvalues \(a_n\) of the matrix \(A\). The comments to follow expand on the background necessary to perform steps 1–5. Emphasis will be both on the construction of a device measuring a given Hermitean operator (step 3) and on the working of a quantum mechanical measurement (step 4).

**Ad 1.** Consider a Hilbert space \(H_a\) of dimension \(N = (2s + 1)\) which carries an irreducible representation of the group \(SU(2)\) with the spin components \((S_1, S_2, S_3)\) as generators. Multipole operators \(T_{j_1j_2\ldots j_k}\) are defined as the symmetrized products \(S_{j_1}S_{j_2}\ldots S_{j_k}\), \(j_i = 1, 2, 3, \) and \(\sigma = 0, 1, \ldots, 2s\), after subtracting off the trace, except for \(T_0 \equiv T^{(0)} = 1\), the \((N \times N)\) unit matrix. The index \(\sigma\) labels \((2s + 1)\) classes with \((2\sigma + 1)\) elements transforming among themselves under rotations. Explicitly, the lowest multipoles read

\[
T^{(1)}_j = S_j \quad T^{(2)}_{j_1j_2} = \frac{1}{2} (S_{j_1}S_j + S_jS_{j_1}) - \frac{\delta_{j_1j_2}}{3} S_{j_1}S_{j_2}
\]

For the sake of brevity, a collective index \(\nu \equiv (\sigma; j_1, \ldots, j_k)\) is introduced now, taking on the values \(\nu = 0, 1, \ldots, N^2 - 1\). The \(N^2\) self-adjoint multipole operators \(T_\nu = T^\dagger_\nu\) form a basis in the space of Hermitean operators acting on the \(N\)-dimensional Hilbert space \(H_a\) [8]. Two multipoles are orthogonal with respect to a scalar product defined as the trace of their product:

\[
(1/N) \text{Tr}[T_\nu T_\nu] = \delta_{\nu\nu'}.
\]

**Ad 2.** Since the multipoles are expressed explicitly as a function of the spin components not exceeding the power \(2s\), it is justified to consider them and, *a fortiori*, the quantity \(H_A\) as an observable for a spin \(s\).

**Ad 3.** It is natural to expect that every self-adjoint operator \(\hat{B}\) comes along with an apparatus \(\text{app}(\hat{B})\) capable of measuring it [9]. For particle systems, setting up such a device remains a challenging task for an experimenter while the situation is more favourable for spin systems. Swift and Wright [8] have shown how to devise, in principle, a *generalized Stern–Gerlach apparatus* which measures any observable \(H_A(S)\) as a traditional Stern–Gerlach apparatus measures the spin component \(n \cdot S\) along the direction \(n\). The construction requires that arbitrary static electric and magnetic fields, consistent with Maxwell’s equations, can be created in the laboratory. To construct an apparatus \(\text{app}(H_A)\) means to identify a spin Hamiltonian \(H(r, S)\) which splits an incoming beam of particles with spin \(s\) into subbeams corresponding to the eigenvalues \(a_n\). The most general Hamiltonian acting on the Hilbert space \(H_a\) of a spin \(s\) reads

\[
H(r, S) = \sum_{i=0}^{N^2-1} \Phi_\nu(r)T_\nu
\]

with traceless multipoles (except for \(\nu = 0\)), and totally symmetric expansion coefficients \(\Phi_\nu(r)\) (\(\equiv \Phi_{\nu}^{(s)}(r)\)) which vary in space. Tune the electric and magnetic fields in such a way that the coefficients \(\Phi_\nu(r)\) and their first derivatives with respect to some spatial direction,
As shown explicitly in [8], this is always possible with realistic fields satisfying Maxwell’s equations. Then, the Hamiltonian in (5) has two important properties. (i) At the origin, \( r = 0 \), it coincides with the matrix \( H_A \) in equation (3). (ii) Suppose that a beam of particles with spin \( s \) enters the generalized Stern–Gerlach apparatus \( \text{app}(H_A) \) just described. At its centre, particles in an eigenstate \( |a_n⟩ \), say, will experience a force in the \( r_1 \) direction given (up to second order in distance from the centre) by

\[
F_i(r)|_{r=0} = -\frac{\partial}{\partial r_1} \langle a_n|H(r,S)|a_n⟩ = -a_n \quad n = 1, \ldots, 2s + 1.
\]

Consequently, this apparatus can spatially separate particles with a spin projected onto one of the eigenstates \(|a_n⟩\) of the operator \( H_A \), with a separation proportional to the eigenvalue \( a_n \). The working principle is entirely analogous to that of a familiar Stern–Gerlach apparatus for a spin \( 1/2 \) (see [8] for more details).

**Ad 4.** The ‘projection postulate’ of quantum mechanics describes the effect of measuring an observable \( \hat{B} \) on a system \( S \) by means of an apparatus \( \text{app}(\hat{B}) \). If the system is prepared initially in a state with density matrix \( \hat{ρ} \), the impact of measuring \( \hat{B} \) is:

\[
\text{app}(\hat{B}) : \quad \hat{ρ} \xrightarrow{p_n} \left( b_n; \hat{ρ}_n \right) \quad p_n = \text{Tr}[\hat{ρ}\hat{ρ}_n].
\]

In words, the action of the apparatus is, with probability \( p_n \), to throw the system prepared in state \( \hat{ρ} \) into an eigenstate \( \hat{ρ}_n = |b_n⟩⟨b_n| \) of the observable \( \hat{B} \); the outcome of the measurement will be the associated eigenvalue \( b_n \). As a matter of fact, the notion of ‘collapse’ or ‘projection’ can be avoided if one characterizes the process indirectly by referring to ‘repeatable measurements’ [10].

As is well known, the outcome of an individual measurement cannot be predicted due to the probabilistic character of quantum mechanics, and it is necessary to repeat the experiment a number of times until all values \( a_n \) have been obtained. If the spin \( s \) is prepared initially in a homogeneous mixture, \( \hat{ρ} = 1/(2s + 1) \), the \( (2s + 1) \) possible outcomes will occur with equal probability. The probability not to have obtained one specific value \( a_n \) after \( N_0 \gg N \) measurements equals \( (2s/2s + 1)^{N_0} \simeq \exp[-N_0/2s] \), decreasing exponentially with \( N_0 \).

**Ad 5.** It would be convenient to ‘read out’ directly the quantum state \( \hat{ρ}_n \) obtained from a single measurement with result \( a_n \). However, due to the no-cloning theorem [11, 12], an unknown state cannot be determined if only one copy is available. Upon repeating the measurement a large number of times and keeping only those states with the same eigenvalue \( a_n \), one produces an ensemble of systems prepared in the state \( \hat{ρ}_n \). This is sufficient to experimentally reconstruct an unknown state since a density matrix \( \hat{ρ} \) can be written as

\[
\hat{ρ} = \frac{1}{N} \sum_{\mu=1}^{N^2} P_\mu \hat{Q}^\mu \quad N = 2s + 1
\]

where the coefficient \( P_\mu = \langle n_\mu|\hat{ρ}|n_\mu⟩ \) is the probability to find the system in a coherent spin state \( |n_\mu⟩ \). As shown in [13], the operators [13] \( \hat{Q}^\mu, \mu = 1, \ldots, N^2 \), form a basis for Hermitian operators, similar to but different from the multipoles \( T_\nu \). Thus, equation (9) parametrizes \( \hat{ρ} \) by expectation values \( P_\mu \) which can be measured by a standard Stern–Gerlach apparatus.

Let us turn to the discussion of the idea underlying quantum diagonalization. Traditionally, a measurement is thought to confirm or reveal some information about the state of the system.
The probabilities $p_n$ provide information about the state of the system conditioned by the selected observable. Thus, a measurement reveals (or confirms) properties of the state $\hat{\rho}$ of the system while the observable $\hat{B}$ at hand is assumed to be known, including its eigenstates and eigenvalues. To put it differently, the observable defines the scope of the possible results of a measurement: the only possible outcomes are its eigenvalues $b_n$, and, directly after the measurement the system necessarily resides in the corresponding state $|b_n\rangle$.

In the present context, however, the idea is to extract information about the measured observable—not about the state $\hat{\rho}$. Why is this possible at all? It is fundamental to realize that the input needed to set up a measurement of $\hat{A}$ and to actually measure it is different from the output of the experiment. In order to measure the observable $\hat{A}$, the construction of an apparatus $\text{app}(\hat{A})$ is sufficient—and its construction is indeed possible without knowing eigenvalues and eigenstates of $\hat{A}$ beforehand. After a measurement, however, partial information about the spectral properties of the observable $\hat{A}$ is available according to (8). This is due to the constraints (i) that the possible outcomes of measuring $\hat{A}$ are its eigenvalues and (ii) that the system subsequently will reside in the corresponding eigenstate. Thus, if the eigenstates and eigenvalues of $\hat{A}$ are not known initially, information about them indeed emerges by measuring $\hat{A}$.

The quantum mechanical diagonalization appears to be neither an analogue nor a digital calculation. It is not based on the representation of a mathematical equation in terms of a physical system which then would 'simulate' it, even though the outcome will be read on an analogue scale. Further, no 'software program' is executed which would implement a diagonalization algorithm. One might best describe the measuring device $\text{app}(H_A)$ as a 'special purpose machine' based on the projection postulate.

For the time being, the method introduced here is important from a conceptual but not necessarily from a technological point of view. On the one hand, the diagonalization of matrices is not a hard problem such as factorization of large integer numbers; on the other, the actual implementation in the laboratory is challenging just as it is to set up a full-fledged quantum computer. It must be stressed, however, that there is no physical principle which would forbid to build such machines.

Further, quantum diagonalization is expected to be fruitful from a conceptual point of view since it provides a different perspective on the projection postulate [16]. It shows that—in an unexpected way—standard quantum mechanics attributes computational power to the measurement of an observable. The fact that one can use a measurement to perform calculations might turn into an argument in favour of the 'reality' of the quantum mechanical projection postulate.

What is the relation of quantum diagonalization to other research in quantum computing? Various well-established quantum algorithms can be cast in a form which asks for the determination of eigenvalues of unitary operators [14] representing the action of a network on some initial state. The result of the calculation is encoded in the phase of the output state. Subsequently, measurements are necessary to read off the result but their role in the algorithm is different: the calculation has already been performed—and the measurement itself is not the working principle. Further, it has been pointed out in [15] how to use an ideal quantum computer to effectively simulate another quantum system. As indicated briefly, such a device would be a useful tool to measure arbitrary observables, in a spirit somewhat similar to the method presented here.

In sum, the basic ingredient of quantum diagonalization, the ‘collapse’ of the wave function projecting any state onto a randomly selected eigenstate of the measured observable, has been shown to possess computational power. Generalizations of this approach are expected to include the diagonalization of unitary matrices and the determination of roots of polynomials.
References


(Weigert St 1999 Preprint quant-ph/99030103)


(Zalka C 1996 Preprint quant-ph/9603026 v2)