

The quantum way to diagonalize hermitean matrices

Stefan Weigert*

HuMP – Hull Mathematical Physics, Department of Mathematics University of Hull, Cottingham Road UK-Hull

Received 10 June 2002, accepted 19 June 2002

Published online 25 February 2003

PACS 03.67.-a, 03.65Sq

An entirely quantum mechanical approach to diagonalize hermitean matrices has been presented recently. The method is based on the measurement of quantum mechanical observables which provides the computational resource. In brief, quantum mechanics is able to directly address and output eigenvalues of hermitean matrices. Here, the example of (2×2) matrices is considered in detail. This simple case allows one to illustrate the conceptual features of the general method which is applicable to $(N \times N)$ hermitean matrices.

1 Introduction

A new attitude towards quantum theory has emerged in recent years. The focus is no longer on attempts to come to terms with counter-intuitive quantum features but to capitalize on them. In this way, surprising methods have been uncovered to solve specific problems by means which have no classical equivalent: *quantum cryptography*, for example, allows one to establish secure keys for secret transmission of information [1]; entanglement [2] is used as a tool to set up powerful *quantum algorithms* which do factor large integers much more efficiently than any presently 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 [3, 5], required to let a potential algorithm run.

The purpose of the present contribution is to study the simplest situation in which a quantum mechanical measurement, i.e. the bare ‘projection’ [2], “does” the computation. The computational task is to determine the eigenvalues of hermitean (2×2) matrices by quantum means alone. Although the answer to this problem can be given analytically it is useful to discuss this particular case since there is no conceptual difference between diagonalizing (2×2) or $(N \times N)$ hermitean matrices along these lines [6].

Before turning to the explicit example, consider briefly the traditional view on quantum mechanical measurements: a measurement is thought to confirm or reveal some information about the *state* of the system. The measured observable \hat{A} is assumed to be known entirely, including in-principle-knowledge of its eigenstates and eigenvalues. Further, the observable also defines the *scope* of the possible results of a measurement since both the only allowed outcomes are its eigenvalues and, directly after the measurement, the system necessarily resides in the corresponding eigenstate.

In the context of quantum diagonalization, however, the crucial idea is to learn something about the measured *observable* – not about the state of the system. Why is there scope for information gain at all? What can one learn from a quantum mechanical experiment if both the measured observable and the state are known?

* E-mail: S.Weigert@hull.ac.uk

It is essential to realize that the *input* required to measure an observable \hat{A} and the *output* of an experiment, in which \hat{A} is actually measured, are *not* identical. In fact, it is possible to construct an apparatus which measures an observable \hat{A} *without* explicitly knowing its eigenvalues. Then, a measurement provides partial – but explicit – information about the spectral properties of the observable \hat{A} : it delivers one of its eigenvalues. As the eigenvalues of \hat{A} are *not* known explicitly before the measurement, information is indeed gained by measuring \hat{A} . This is the idea which underlies quantum diagonalization.

The quantum diagonalization of a hermitean matrix is achieved in five steps: (1) express the matrix in a *standard form*; (2) associate a quantum mechanical *observable* with it; (3) identify an *apparatus* capable of measuring the observable; (4) *measure* the observable – this provides the *eigenvalues* of the matrix; (5) determine its *eigenstates*. The next section gives the details for (2×2) matrices. Then, the generalization to $(N \times N)$ matrices is briefly summarized.

2 Quantum diagonalization of hermitean (2×2) matrices

Suppose you want to determine the eigenvalues A_{\pm} (and, subsequently, the eigenvectors $|A_{\pm}\rangle$) of the general hermitean (2×2) matrix

$$A = \begin{bmatrix} \alpha & \beta^* \\ \beta & \gamma \end{bmatrix}, \quad \alpha, \gamma \in \mathbb{R}, \quad \beta \in \mathbb{C}. \quad (1)$$

Obviously, this problem is easily solved analytically. The eigenvalues read

$$A_{\pm} = \frac{1}{2} \left(\alpha + \gamma \pm \sqrt{(\alpha - \gamma)^2 + 4\beta\beta^*} \right), \quad (2)$$

and one can also give explicit expressions for the eigenvectors of the matrix A. The five-step procedure of quantum diagonalization is now applied to A; it will, of course, reproduce the result (2). However, all conceptual points, which also apply to the technically more cumbersome case of $(N \times N)$ matrices are conveniently illustrated by this simple example.

1. *Standard form of A:* Any hermitean (2×2) matrix A can be written as a unique linear combination of the three Pauli matrices $\sigma_j = \sigma_j^{\dagger}, j = 1, 2, 3$, and the unit matrix $\sigma_0 = I_2$,

$$A = (a_0\sigma_0 + \mathbf{a} \cdot \boldsymbol{\sigma}), \quad \mathbf{a}_j = \frac{1}{2} \text{Tr} [A\sigma_j] \in \mathbb{R}, \quad (3)$$

where

$$\mathbf{a}_0 = \frac{1}{2}(\alpha + \gamma), \quad \mathbf{a}_1 = \frac{1}{2}(\beta + \beta^*), \quad \mathbf{a}_2 = \frac{1}{2i}(\beta^* - \beta), \quad \mathbf{a}_3 = \frac{1}{2}(\alpha - \gamma). \quad (4)$$

From a general point of view, this corresponds to an expansion of A in multipole operators (cf. below).

2. *Identification of an observable:* The most general Hamiltonian of a spin-1/2 in a homogeneous magnetic field \mathbf{B}_0 is linear in the components of the spin $\mathbf{S} = \hbar\boldsymbol{\sigma}/2$. Therefore, any matrix A has an interpretation as a Hamiltonian operator of a quantum spin subjected to an appropriately chosen magnetic field,

$$A = aI_2 - g\mu_B\mathbf{B}_0 \cdot \mathbf{S} \equiv H_A(\mathbf{S}), \quad a = a_0, \quad \mathbf{B}_0 = \frac{-2}{g\mu_B\hbar}\mathbf{a}. \quad (5)$$

The part aI_2 shifts the energy globally by a fixed amount. Let E_{\pm} denote the eigenvalues of the second part of this expression, $-g\mu_B\mathbf{B}_0 \cdot \mathbf{S} \equiv H_A^0(\mathbf{S})$; then, the eigenvalues A_{\pm} of the matrix A are given by

$$A_{\pm} = a + E_{\pm}. \quad (6)$$

Consequently, the door is now open to determine the eigenvalues of A experimentally, i.e. through *measuring* the eigenvalues E_{\pm} of the Hamiltonian $H_A^0(S)$. In the next step it is shown how to devise an apparatus which measures this operator.

3. *Setting up a measuring device:* The apparatus is required to measure the eigenvalues of the operator $H_A^0(S)$. In the case of a (2×2) matrix this is just a familiar Stern-Gerlach apparatus, appropriately oriented in space. However, as the method will be applied to $(N \times N)$ hermitean matrices later on, it is important to go through the construction of the measuring device in detail.

Consider the spatially varying Hamiltonian

$$H^0(\mathbf{r}, S) = aI_2 - g\mu_B \mathbf{B}(\mathbf{r}) \cdot S \equiv aI_2 + H_A^0(\mathbf{r}, S), \tag{7}$$

which describes the interaction of a spin with an inhomogeneous magnetic field $\mathbf{B}(\mathbf{r})$. In order that $H^0(\mathbf{r}, S)$ measure the observable $H_A^0(S)$, the field needs to satisfy the conditions

$$\mathbf{B}(0) = \mathbf{B}_0 \quad \text{and} \quad \nabla \cdot \mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{B}(\mathbf{r}) = 0. \tag{8}$$

Then, at the centre of the apparatus, the operator $H_A^0(\mathbf{r}, S)$ coincides with the observable to be measured, $H_A^0(0, S) = H_A^0(S)$, and the magnetic field is consistent with Maxwell's equations. Consider the field [7]

$$\mathbf{B}(\mathbf{r}) = (1 + \mathbf{k} \cdot \mathbf{r})\mathbf{B}_0 + (\mathbf{B}_0 \cdot \mathbf{r})\mathbf{k}, \tag{9}$$

which is consistent with Eqs. (8) if the vector \mathbf{k} is perpendicular to \mathbf{B}_0 . Diagonalize the operator $H_A^0(\mathbf{r}, S)$ to first order of $\nabla|\mathbf{B}|/|\mathbf{B}|$. This leads to \mathbf{r} -dependent eigenvalues

$$E_{\pm}(\mathbf{r}) = \pm \frac{\hbar}{2}(1 + \mathbf{k} \cdot \mathbf{r})B_0, \tag{10}$$

which imply the existence of a state-dependent force

$$\mathbf{F}_{\pm}(\mathbf{r}) = -\nabla E_{\pm}(\mathbf{r}) = \mp \frac{\hbar}{2}B_0\mathbf{k}. \tag{11}$$

Consequently, particles on their way through the apparatus will be deflected deterministically from a straight line once the projection to an eigenstate $|E_{\pm}\rangle$ has occurred. In this way, the eigenvalues of the observable can be accessed experimentally.

4. *Determination of the eigenvalues:* If a measurement of the operator $H_A^0(S)$ is performed on the state $\hat{\rho} = I_2/2$, then it is thrown with probability $1/2$ into one of the eigenstates $|E_{\pm}\rangle$ with density matrix $\hat{\rho}_{\pm}$,

$$\text{app}(H_A^0) : \hat{\rho} = I_2/2 \xrightarrow{p_{\pm}} (E_{\pm}; \hat{\rho}_{\pm}), \quad p_{\pm} = \text{Tr} [\hat{\rho}\hat{\rho}_{\pm}] = \frac{1}{2}. \tag{12}$$

Repeating the measurement a few times, both eigenvalues will have been found soon. The probability not to obtain one of the two values after N_0 identical runs of the experiment equals $1/2N_0$, and hence goes to zero exponentially with the number of runs. Subsequently, the sought-after eigenvalues A_{\pm} of the matrix A are known due to the relation (6), and the major step in the diagonalization of the matrix has been achieved in a quantum way.

5. *Determination of the eigenstates:* Once both eigenvalues A_{\pm} are known, it is straightforward to determine the associated eigenstates by a simple calculation. Optionally, one continues along an experimental line. One exploits the fact that the apparatus $\text{app}(H_A^0)$ prepares an ensemble of eigenstates

of A with density matrix $\hat{\rho}_+$ if the other subbeam (containing $\hat{\rho}_-$) is blocked – and *vice versa*. The Bloch representation of the density matrix $\hat{\rho}_+$, say, can be parameterized by *expectation values*,

$$\hat{\rho}_+ = \frac{1}{2}(\sigma_0 + \langle \boldsymbol{\sigma} \rangle_+ \cdot \boldsymbol{\sigma}), \quad \langle \hat{\sigma}_j \rangle_+ = \text{Tr} [\hat{\rho}_+ \sigma_j] \equiv \langle E_+ | \sigma_j | E_+ \rangle. \quad (13)$$

Hence, the components of the vector $\langle \boldsymbol{\sigma} \rangle_+$ (and therefore $\hat{\rho}_+$) are easily obtained by means of a second, appropriately oriented Stern-Gerlach apparatus, which amounts to a reconstruction of the density matrix $\hat{\rho}_+$.

3 Quantum diagonalization of hermitean ($N \times N$) matrices

Five steps are necessary to diagonalize a hermitean ($N \times N$) matrix A by quantum means.

1. *Standard form of A:* Write the hermitean ($N \times N$) matrix A as a combination of linearly independent hermitean *multipole* operators $T_\nu, \nu = 0, \dots, N^2 - 1$,

$$A = \sum_{\nu=0}^{N^2-1} \mathbf{a}_\nu T_\nu, \quad \mathbf{a}_\nu = \frac{1}{N} \text{Tr} [A T_\nu] \in \mathbb{R}. \quad (14)$$

Multipole operators $T_{j_1 j_2 \dots j_a}$ act in a Hilbert space \mathcal{H}_s 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. They are defined as the symmetrized products $S_{j_1} S_{j_2} \dots S_{j_a}, j_i = 1, 2, 3$, and $a = 0, 1, \dots, 2s$, after subtracting off the trace, except for $T_0 \equiv T^{(0)} = 1$, the ($N \times N$) unit matrix. The index a labels $(2s+1)$ classes with $(2a+1)$ elements each, transforming among themselves under rotations. Explicitly, the lowest multipoles read

$$T_j^{(1)} = S_j, \quad T_{j_1 j_2}^{(2)} = \frac{1}{2} (S_{j_1} S_{j_2} + S_{j_2} S_{j_1}) - \frac{\delta_{j_1 j_2}}{3} S_{j_1} S_{j_2}. \quad (15)$$

For the sake of brevity, a collective index $\nu \equiv (a; j_1, \dots, j_k)$ has been used in (14), taking on the values $\nu = 0, 1, \dots, N^2 - 1$. The N^2 self-adjoint multipole operators $T_\nu = T_\nu^\dagger$ form a basis in the space of hermitean operators acting on the N -dimensional Hilbert space \mathcal{H}_s [7]. Two multipoles are orthogonal with respect to a scalar product defined as the trace of their product: $(1/N) \text{Tr} [T_\nu T_{\nu'}] = \delta_{\nu\nu'}$.

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

$$A = \sum_{\nu=0}^{N^2-1} \mathbf{a}_\nu T_\nu(S) \equiv H_A(S), \quad (16)$$

using the expression of the multipoles $T_\nu(S)$ in terms of the components of a spin. 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 $H_A(S)$ as an *observable* for a spin s .

3. *Setting up a measuring device:* Swift and Wright [7] have shown how to devise, in principle, a *generalized* Stern-Gerlach apparatus which measures any observable $H_A(S)$ – just as a traditional Stern-Gerlach apparatus measures the spin component $\mathbf{n} \cdot S$ along the direction \mathbf{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

$H(\mathbf{r}, S)$ 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 \mathcal{H}_s of a spin s reads

$$H(\mathbf{r}, S) = \sum_{\nu=0}^{N^2-1} \Phi_\nu(\mathbf{r}) T_\nu, \tag{17}$$

with traceless multipoles (except for $\nu = 0$), and totally symmetric expansion coefficients $\Phi_\nu(\mathbf{r}) (\equiv \Phi_{j_1 j_2 \dots j_k}^{(k)}(\mathbf{r}))$. These functions, which vary in space, generalize the magnetic field $\mathbf{B}(\mathbf{r})$ in (7). Tune the corresponding electric and magnetic fields in such a way that the coefficients $\Phi_\nu(\mathbf{r})$ and their first derivatives with respect to some spatial direction, r_1 , say, satisfy

$$\Phi_\nu(\mathbf{r})|_{\mathbf{r}=0} = a_\nu, \quad \text{and} \quad \left. \frac{\partial \Phi_\nu(\mathbf{r})}{\partial r_1} \right|_{\mathbf{r}=0} = a_\nu. \tag{18}$$

As shown explicitly in [7], this is always possible with realistic fields satisfying Maxwell’s equations. Then, the Hamiltonian in (17) has two important properties:

- (i) At the origin, $\mathbf{r} = 0$, it coincides with the matrix $H(0, S) = H_A(S)$.
- (ii) Suppose that a beam of particles with spin s enters the generalized Stern-Gerlach apparatus $\text{app}(H_A)$ just described. At its center, particles in an eigenstate $|A_n\rangle$, say, will experience a force in the r_1 direction given (up to second order in distance from the center) by

$$F_1(\mathbf{r})|_{\mathbf{r}=0} = - \left. \frac{\partial \langle A_n | H(\mathbf{r}, S) | A_n \rangle}{\partial r_1} \right|_{\mathbf{r}=0} = -A_n, \quad n = 1, \dots, 2s + 1. \tag{19}$$

Consequently, particles with a spin residing in different eigenstates $|A_n\rangle$ of the operator H_A will be separated spatially by this apparatus. From a conceptual point of view, the procedure is equivalent to the method outlined above for a spin $1/2$.

- 4. *Determination of the eigenvalues:* Once the apparatus $\text{app}(H_A)$ has been set up, one needs to carry out measurements on (particles carrying) a spin s prepared in the homogeneous mixture $\hat{\rho} = I_N/N$. The output of each individual measurement will be one of the eigenvalues A_n of the matrix A , according to the ‘projection postulate:’

$$\text{app}(\hat{A}) : \hat{\rho} = I_N/N \xrightarrow{p_n} (A_n; \hat{\rho}_n), \quad p_n = \text{Tr} [\hat{\rho} \hat{\rho}_n] = \frac{1}{N}. \tag{20}$$

In words, the action of the apparatus is, with probability $p_n = 1/N$, to throw the system prepared in state $\hat{\rho}$ into an eigenstate $\hat{\rho}_n \equiv |A_n\rangle\langle A_n|$ of the observable \hat{A} ; the *outcome* of the measurement is given by the associated eigenvalue A_n .

After sufficiently many repetitions, all eigenvalues will be known, although the outcome of an *individual* measurement cannot be predicted due to the probabilistic character of quantum mechanics. It is necessary to repeat the experiment a number of times until *all* values A_n have been obtained. Since the spin s has been prepared initially in a homogeneous mixture, $\hat{\rho} = I_N/N$, the $(2s + 1)$ possible outcomes 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 .

- 5. *Determination of the eigenstates:* As before it is possible to either calculate the eigenstates $|A_n\rangle$ of the matrix A on the basis of the known eigenvalues, or to determine them *experimentally* by methods of state reconstruction (see [8] for details).

4 Summary and outlook

As a result of the five steps just described, a hermitean ($N \times N$) matrix A has been diagonalized *without* calculating the zeroes 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 – information which cannot be obtained in closed form if $N \geq 5$. One might best describe the measuring device $\text{app}(H_A)$ as a ‘special purpose machine’ which is based on the ‘collapse of the wave function’ as computational resource. However, one could avoid to use the notion of ‘collapse’ or ‘projection’ by characterizing the process indirectly using the concept of ‘repeatable measurements’ described in [9].

By construction, the quantum mechanical diagonalization is not based on the representation of a mathematical equation in terms of a physical system which then would ‘simulate’ it. Similarly, no ‘software program’ runs which would implement an diagonalization algorithm. Therefore, the method resembles neither an analog nor a digital calculation.

It is worthwhile to point out that the quantum mechanical approach to the diagonalization of hermitean matrices is based on the assumption that the behaviour of a spin s is described correctly by non-relativistic quantum mechanics. Note, finally, that quantum diagonalization does *not* depend on a particular interpretation of quantum mechanics.

References

- [1] C. H. Bennet and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, in: Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing, 175 (1984).
- [2] E. Schrödinger, Die gegenwärtige Situation in der Quantenmechanik, *Naturwissenschaften* **23**, 823 (1935).
- [3] P. W. Shor, Scheme for reducing decoherence in quantum computer memory, *Phys. Rev. A* **52**, R2493 (1995).
- [4] D. Aharonov and M. Ben-Or, Fault-Tolerant Quantum Computation With Constant Error Rate, *quant-ph/9906129*
- [5] A. M. Steane, Error correcting codes in quantum theory, *Phys. Rev. Lett.* **77**, 793 (1996).
- [6] St. Weigert, Quantum diagonalization of hermitean matrices, *J. Phys. A* **34**, 5916 (2001).
- [7] A.R. Swift and R. Wright, Generalized Stern-Gerlach experiments and the observability of arbitrary operators, *J. Math. Phys.* **21**, 77 (1977).
- [8] St. Weigert, Quantum time evolution in terms of nonredundant probabilities, *Phys. Rev. Lett.* **84**, 802 (2000); *quant-ph/99030103*.
- [9] A. Peres, *Quantum theory: concepts and methods* (Kluwer, 1995).