Review of the linear algebra of quantum mechanics

Kris Innanen

ABSTRACT

The continuing (though gradual), development of applied geophysical algorithms and ideas in the area of quantum computing at CREWES, as well as our growing interest in characterizing inversion uncertainty with quasi-dynamical systems (inspired by the so-called Hamiltonian Monte Carlo methods), both motivate some review work on the linear algebra of quantum mechanics. The description of simple systems, in isolation and in combination, is developed, with the latter allowing simple entangled states to be discussed, which is relevant in quantum information and computing. *Laddering operators* in the description of harmonic oscillators are also given special focus: since in the above mentioned Hamiltonian MC methods, energy and data misfit are associated, laddering operators may be useful descriptors of convergence. Steps needed to increase the dimensionality of systems, and steps needed to increase the number of particles/oscillators in systems are also given attention, since both will be needed in evolving applications in optimization and geophysical quantum computing.

INTRODUCTION

In this report a review of the basic ideas of quantum mechanics is set out. It is an outgrowth of the author's decision this year to refresh his memory on the subject, which he felt was timely for two main reasons. First, our group's interest in quantum computing, and in particular in algorithms for quantum computers which will have a positive impact on geophysical computations (e.g., Moradi et al., 2018), continues this year (Monsegny, 2022). If computing technology based on quantum logic and qubits does in fact emerge, and if it does in fact change the geophysical computations we do, it will be the task of geophysicists to build on, and communicate, this new capability. If you are like me, and your interaction with quantum mechanics occurred somewhere between 20 years ago and never, a simple review might be worthwhile.

The second reason is less obvious, but in my view it makes such a review more worthwhile now, rather than a few years from now, when, quantum computing is more likely to have come into its own. A growing literature exists in which uncertainty in large inverse problems is characterized in terms of notional dynamical systems involving notional particles, or "shuttles" (e.g., Keating and Innanen, 2021; Fichtner et al., 2021). Shuttles move through model space, subject to the "gravity" of the misfit or objective function, in essentially the same way as objects move in three-dimensional physical space, i.e., obeying generalized versions of the rules of classical mechanics. Sampling the models along the resulting notional orbital paths gives insight into the uncertainty of a model estimate. The shuttles answer the question: what other models can be conceived of which fit the data equally well? Last year, some theoretical results were obtained which built on this idea (Innanen, 2021). To sample model space effectively, setting out not one, but instead full ensembles, of these objects, in the form of shuttle "gases", seems like a natural step. This led to the introduction of a Gibbs' ensemble of shuttles, and it was found that an analog to

the barometric equation for the density of shuttles was relatively easy to derive. However, it is not possible to consider extensions of dynamical rules based on statistical mechanics without considering quantum analogs as well. In fact, if a notional dynamical system is to be used to extract information about optimization problems, then there are many reasons to select a quantum, as opposed to a classical, dynamical rule set. This research is just getting underway, and it will be reported on in some detail in 2023.

I relied on some excellent pedagogical tools for this review, especially the texts of Harris (1972) and Susskind and Friedman (2015), which are approachable enough to later allow a dip or two into the more difficult but central text of Dirac (1930). The terminology probably looks closest to that of Harris, but especially in discussing ladder operators and harmonic oscillators I have diverged into a terminology that looks a little geophysical, and is chosen to assist in future reports.

There is no original work here at all, only organizational effort and discussion points which, appearing to be important to one quantitative geophysicist, might appeal to other quantitative geophysicists. The review starts with the basics of vectors and operators, bra-ket notation, eigen-decompositions, etc., and ends with discussions of combined systems and entanglement, which is relevant for quantum computing, and the dynamical behaviour of harmonic oscillators, which will play the role of the shuttles in the continuation of the analysis of inversion uncertainty.

VECTORS AND OPERATORS

Bra-ket notation

A generic vector in quantum mechanics is known as a *ket*. The object $|\alpha\rangle$, for instance, is a ket. It is labelled α to distinguish it from other kets; in the end we might produce kets with several labels, or indices; the main thing is that the label or labels are there to help make sure we don't mistake this ket for some other. The ket $|\alpha\rangle$ is considered to be an element of a Hilbert space $\mathcal{H}_a$, which generally contains complex elements and has dimension N_a.

The ket $|\alpha\rangle$ is an *abstract vector* in the same sense that in our standard notation the vector $\mathbf{a}$ is abstract: we know that it can be specified with an array of N_a numbers, but we have to add more information before we can assign a particular set of numbers to it. Still, it is helpful to remember that when a coordinate system has been specified, the ket vector $|\alpha\rangle$ will be realized as a column array

$$|\alpha\rangle = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{N_a} \end{bmatrix}, \tag{1}$$

with the α_i generally complex. A second type of vector, called a *bra*, is also invoked, and is denoted $\langle\alpha|$. Each ket in $\mathcal{H}_a$ has a counterpart bra, the full complement of which make up a second, "dual" space. In our standard notation these bras are the equivalents of $\mathbf{a}^T$, i.e., the Hermitian, or conjugate transposes of ordinary vectors. As with the kets, with a

coordinate basis chosen the bra vector is realized as a row array

$$\langle\alpha| = \left[\alpha_1^*, \alpha_2^*, \ldots, \alpha_{N_a}^*\right], \tag{2}$$

i.e., the transpose of $|\alpha\rangle$ with the complex conjugate taken for each element. As in all applications of linear algebra, inner products between vectors are critical to quantum mechanics. In standard notation, the inner product of $\mathbf{a}$ with itself would be given by $\mathbf{a}^T\mathbf{a}$; the comparable product in quantum mechanics is

$$\langle\alpha|\alpha\rangle\,. \tag{3}$$

The elegance of the bra-ket notation is evident here, with this calculation producing a complete "bracket".

As in standard vector analysis, we refer to two vectors, say $|\alpha\rangle$ and $|\alpha'\rangle$, as orthogonal if $\langle\alpha|\alpha'\rangle = 0$, and a suite of vectors $|\alpha_i\rangle$ for $i = 1, ..., N_a$ as being mutually orthonormal if

$$\langle\alpha_i|\alpha_j\rangle = \delta_{ij}, \tag{4}$$

where δ_{ij} is the Kronecker delta. An orthonormal basis of $\mathcal{H}_a$ is a suite of vectors of this kind. Any element of $\mathcal{H}_a$ can be expanded as a linear combination of these basis vectors. If $|\psi\rangle \in \mathcal{H}_a$, then

$$|\psi\rangle = \sum_{i=1}^{N_a} |\alpha_i\rangle\,\psi_i = \sum_{i=1}^{N_a} |\alpha_i\rangle\,\langle\alpha_i|\psi\rangle\,. \tag{5}$$

For reasons that don't appear for quite a while in the build up of quantum mechanics, the weights $\psi_i = \langle\alpha_i|\psi\rangle$ are referred to as the *wave function*. The definition of the wave function ψ_i also closes the loop on the realization of an abstract vector as a column array of numbers. In the α representation, one says, the ket ψ has elements

$$|\psi\rangle = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{N_a} \end{bmatrix} = \begin{bmatrix} \langle\alpha_1|\psi\rangle \\ \langle\alpha_2|\psi\rangle \\ \vdots \\ \langle\alpha_{N_a}|\psi\rangle \end{bmatrix}. \tag{6}$$

In quantum mechanics the length of a vector (a bra, or a ket) does not affect the physics, and generally all vectors are normalized to unit length, in keeping with eventual probabilistic interpretations. However, because all vector quantities in quantum mechanics are generally complex, a vector that has been normalized to unit length has still not been fully specified, since it can be multiplied by a phase term $e^{i\theta}$ without changing its length.

Operators

A linear operator, say A, is an object defined on $\mathcal{H}_a$ such that it enters into a product with a ket to produce another ket in $\mathcal{H}_a$, i.e., $|\alpha'\rangle = A\,|\alpha\rangle$. In so doing it satisfies a range of properties keeping it linear (which we won't review here). In standard notation, an operator

can therefore be thought of as a matrix $\mathbf{A}$. Operators A have adjoints $A^\dagger$, which allow the results of their actions on bras to be kept consistent. If $|\alpha'\rangle = A\,|\alpha\rangle$, then

$$\langle\alpha'| = \langle\alpha|\,A^\dagger. \tag{7}$$

In standard notation, the adjoint would be expressed as the Hermitian transpose of the matrix $\mathbf{A}^H$.

In the same way that, given a representation (i.e., an orthonormal basis), we can express an abstract vector as a column array, through $\langle\alpha_i|\psi\rangle = \psi_i$, the i,jth element of a matrix can be got from an abstract operator once a representation is chosen. To produce

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1N_a} \\ A_{21} & A_{22} & \dots & A_{2N_a} \\ & & \ddots & \\ A_{N_a1} & A_{N_a2} & \dots & A_{N_aN_a} \end{bmatrix}, \tag{8}$$

we calculate

$$A_{ij} = \langle\alpha_i|A|\alpha_j\rangle\,. \tag{9}$$

The operators that appear most often and play the most important roles in quantum mechanics are self-adjoint, or *Hermitian*, which means they have the property that $A = A^\dagger$.

The simplest and most useful operator is the identity operator I, defined such that for any $|\psi\rangle \in \mathcal{H}_a$, $I\,|\psi\rangle = |\psi\rangle$. If a representation is chosen, in that representation it takes the form of the unit matrix $\mathbf{I}$. Another useful example of an operator is produced by the outer product of two vectors, especially basis vectors, e.g., $|\alpha_i\rangle\,\langle\alpha_i|$. Especially helpful is the fact that the identity operator can be built up from these outer product operators. Suppose a vector $|\psi\rangle$ is being constructed in the orthonormal basis $|\alpha_i\rangle$. Because the construction is linear in ψ, we can extract it from the construction and imply the existence of an identity operator:

$$|\psi\rangle = \sum_{i=1}^{N_a} |\alpha_i\rangle\,\langle\alpha_i|\psi\rangle = \left(\sum_{i=1}^{N_a} |\alpha_i\rangle\,\langle\alpha_i|\right)|\psi\rangle\,, \tag{10}$$

or

$$\sum_{i=1}^{N_a} |\alpha_i\rangle\,\langle\alpha_i| = I. \tag{11}$$

It will be very useful at times to insert this version of "one" multiplicatively into an expression to produce new results.

Eigendecomposition

An eigenvector (sometimes "eigenket") of a linear operator A is the ket $|a\rangle$ for which

$$A\,|a\rangle = a\,|a\rangle\,, \tag{12}$$

or that ket which when acted on by A returns a scaled version of itself. If A is Hermitian, it has N_a real eigenvalues, and N_a orthonormal eigenkets, which can be used as a basis over which to expand any desired element, say $|\alpha\rangle$, of $\mathcal{H}_a$:

$$|\alpha\rangle = \sum_{a=1}^{N_a} |a\rangle \langle a|\alpha\rangle = \sum_{a=1}^{N_a} |a\rangle \alpha_a. \tag{13}$$

In this expansion, the labelling convention gives the basis as $|a\rangle$ for $a = 1, 2, ..., N_a$. However, it may be framed in other ways as well. For instance, if we associate the vectors with energy levels later, we might instead write

$$A |E_a\rangle = E_a |E_a\rangle , \;\; |\alpha\rangle = \sum_{a=1}^{N_a} |E_a\rangle \langle E_a|\alpha\rangle = \sum_{a=1}^{N_a} |E_a\rangle \alpha_a. \tag{14}$$

Some degree of ease and experience with seeing variations of this kind is useful, especially when consulting the literature or more than one text on this subject.

Commutators

A Hilbert space $\mathcal{H}_a$ will generally have more than one operator defined on it. Pairs of operators, for instance A and A', may not commute, meaning that the order in which A and A' are applied to an input ket vector changes the resulting output ket. This is measured with the *commutator* $[\cdot, \cdot]$:

$$[A, A'] = AA' - A'A, \tag{15}$$

which is the 0 operator if A and A' commute, but is some non-zero operator if they do not.

Physical interpretation

The state of a physical system is characterized by a ket, say $|\alpha\rangle$, which is an element of a Hilbert space $\mathcal{H}_a$ of dimension N_a. All kets in $\mathcal{H}_a$ are plausible states of the system. Observables are quantities which are in principle measurable in an experiment. They are represented by Hermitian operators, like A. A has N_a real eigenvalues a_i, $i = 1, ..., N_a$, and N_a orthonormal eigenvectors $|a_i\rangle$ (which, being elements of $\mathcal{H}_a$, also represent plausible states of the system). If the system is in an eigenstate $|a_i\rangle$, an experiment to determine the observable A will give outcome a_i with 100% certainty.

The eigenvalues a_i, $i = 1, ..., N_a$ form a complete list of the possible outcomes of an experiment to determine A. Measurement of A will return one of these values, even if the system is in a state (say $|\alpha\rangle$) which is not equal to an eigenvector. When the system is in such a state, an experiment to determine the observable A still returns one of a_i, $i = 1, ..., N_a$, but, which of these outcomes is returned appears only probabilistically in the theory. Because the eigenvectors form an orthonormal basis of $\mathcal{H}_a$, all kets can be decomposed into a linear combination of eigenvectors:

$$|\alpha\rangle = \sum_{i=1}^{N_a} |a_i\rangle \alpha_i, \tag{16}$$

where $\alpha_i = \langle a_i|\alpha\rangle$. If the system is in state $|\alpha\rangle$, an experiment to determine the observable A will produce outcome a_i with probability $P_i = \alpha_i^* \alpha_i$. If the state vectors are properly normalized (i.e., are of unit length), the probability satisfies $\sum_i P_i = 1$.

Although at the outset there was a probability of (say) $\alpha_j^* \alpha_j$ to measure outcome a_j, if a_j is actually measured, a repeat measurement will return outcome a_j with 100% certainty. The system is therefore said, through this measurement, to have been "prepared" in state $|a_j\rangle$. The new state, being in $\mathcal{H}_a$, can also of course be expanded in eigenkets:

$$|a_j\rangle = \sum_{i=1}^{N_a} |a_i\rangle \, \delta_{ji}, \tag{17}$$

but since it itself is an eigenket, the wave function changes from α_i, which generally is non-zero across many i's, to δ_{ij}, which is non-zero at $i = j$ only. This is the origin of the idea of a quantum measurement causing the wave function to "collapse".

Quantum systems have an operator for each observable dynamical variable, which in general implies many. These operators may or may not commute. If they do not commute, the two observables are said to interfere with each other, and be incompatible; if they do commute, the observables are said to be compatible.

Example: spin

Spin is a dynamical property assigned to microscopic particles to account for their angular momentum and their behaviour in magnetic fields. Spin can be oriented in any spatial direction, so its description involves three components. Well-defined experiments can be set up to measure one component of the spin of a particle at a time, in any of the three spatial directions x, y, or z. This implies Hermitian operators $A = \sigma_z$, $A' = \sigma_x$ and $A'' = \sigma_y$ can be built up, representing the three observables. Let us take σ_z as an example. The list of possible outcomes of the spin z measurement is two in length: +1, for spin up, and -1, for spin down. This fixes the dimension of the space of states: $N_\sigma = 2$, which in turn means that when a basis is chosen, the state ket will be a column vector containing 2 elements, and the σ_i, $i = x, y, z$ will be realized as 2×2 matrices.

The Pauli matrices were designed to produce operators which encapsulate the results of spin experiments. An apparatus to determine one of the components of spin can be oriented in any desired direction $\mathbf{n} = [n_x, n_y, n_z]^T$ in three-dimensional space*. When one of the components of spin of a system is observed with this apparatus, it returns either -1 or +1. Subsequent measurements of the same component of spin always return the same answer: the system has been "prepared" in this state.

*Notice we are using standard vector notation $\mathbf{n}$ here, rather than the Dirac notation. This is done deliberately, as a reminder that the 3D space of experiments is *not* a space of states, and the state space is typically *not* 3D. In fact, the spin example here is a good example to keep in mind, since the space of states is 2D, even though it describes experiments which occur – naturally – in 3D physical space. The state space is a calculation space only.

Consider the following eigen-system:

$$\begin{aligned}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} &= +1\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} &= -1\begin{bmatrix} 0 \\ 1 \end{bmatrix}.\end{aligned} \tag{18}$$

Notice that it has several features making it appropriate for describing a system of spin. The dimension (2) is correct, the matrix is Hermitian, the eigenvectors form a complete orthonormal set of the 2D plane, and the eigenvalues of the matrix are ± 1, which are the two possible outcomes of the measurement of a component of spin. Let us further suppose (this is arbitrary) that the particular component of spin we are discussing is in the z direction. The above construction suggests that we assume a system

$$\begin{aligned} A\left|a_1\right\rangle &= a_1\left|a_1\right\rangle \rightarrow \sigma_z\left|u\right\rangle = +1\left|u\right\rangle, \\ A\left|a_2\right\rangle &= a_2\left|a_2\right\rangle \rightarrow \sigma_z\left|d\right\rangle = -1\left|d\right\rangle. \end{aligned} \tag{19}$$

Now we are ready to assess the certainty of experimental outcomes for arbitrary states. If the spin is in state $|\alpha\rangle$, we expand it in the orthonormal basis

$$\left|\alpha\right\rangle = \sum_{i=1}^{N_a}\left|a_i\right\rangle\left\langle\alpha|a_i\right\rangle = \left|u\right\rangle\left\langle\alpha|u\right\rangle + \left|d\right\rangle\left\langle\alpha|d\right\rangle = \left|u\right\rangle\alpha_u + \left|d\right\rangle\alpha_d. \tag{20}$$

Then, we predict that if we were to measure the z component of spin, we would measure either +1, doing so with probability $P_u = \alpha_u^*\alpha_u$, or -1, doing so with probability $P_d = \alpha_d^*\alpha_d$.

The next step is to fold in the other two possible observables, namely the spin components in the x and y directions. The spin components are examples of dynamical variables which mutually interfere, so we should notice a few features emerging: the operators should not commute, for instance. Also, states which have been prepared to give a certain experimental outcome for one spin component will be altered away from this prepared state by the action of measuring a different component.

Let us describe this second feature in more detail. If an experiment is carried out to measure σ_z, the outcome will be one of +1 or -1; once one of these outcomes is observed, this same outcome will result from any number of repeated measurements, with 100% certainty. The state has, in other words, been prepared in one or other of the up and down eigenstates. Suppose the spin has been prepared in the +1 direction. If at this point the apparatus is rotated to measure σ_x, classically one would expect that the outcome would be 0, since we know in advance that the system has been prepared with spin in an orthogonal direction. However, that is not what is observed. What is observed is an outcome of +1 or -1, each with a probability of 50%. That is, *something* comes out to be zero, but it is the average spin over many realizations, and never any particular outcome. Once a +1 or -1 has been observed, the system is again said to have been prepared. If we then re-orient the apparatus to once again measure the z component, we find that the previous preparation has been destroyed: we again measure +1 or -1 with 50% probability. Importantly for what is

to follow, if the system had been prepared in a direction $\mathbf{n}$ somewhere between the z and the x component directions, and then the apparatus was oriented to measure the x component of spin, some degree of uncertainty about the outcome would remain. The outcomes +1 and -1 would not be observed with a 50% probability, but rather with a probability that approaches 100% as $\mathbf{n}$ tilts closer to the x axis. As long as $\mathbf{n}$ has a component in the z direction, some uncertainty remains in the x-component measurement.

To set up an eigen-system appropriate for the operator σ_x, let us set out some features it must have. We now have a valid basis for the state space, $|u\rangle$ and $|d\rangle$, so we don't have to discuss the operator and its eigenvectors as abstract vectors. We can use $|u\rangle$ and $|d\rangle$ as a basis, and go right to a discussion of column vectors and matrices[†]. We don't know much about the eigen-system needed for σ_x in the u-d representation, but we can say immediately that it is of the form

$$\begin{aligned} \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot \\ \cdot \end{bmatrix} &= +1 \begin{bmatrix} \cdot \\ \cdot \end{bmatrix}, \\ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot \\ \cdot \end{bmatrix} &= -1 \begin{bmatrix} \cdot \\ \cdot \end{bmatrix}, \end{aligned} \tag{21}$$

because we know it is a 2×2 system, and we know that all spin measurements are constrained to return +1 or -1 as the only possible outcomes. Let us next deduce the eigenvectors. They must correspond to states in which the x component of spin will be observed to be either +1 or -1 with 100% certainty. It follows that these eigenvectors must be orthogonal to the σ_z eigenvectors, since our experimental evidence (see above) is that if the system has a component in the direction of a σ_z eigenstate, the x component is not measured with certainty. This motivates

$$\begin{aligned} \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} &= +1 \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} &= -1 \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \end{aligned} \tag{22}$$

with the $\sqrt{2}$ factors included to keep the eigenvectors normalized. This is enough to finish filling in the construction. The σ_x matrix must be

$$\begin{aligned} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} &= +1 \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} &= -1 \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \end{aligned} \tag{23}$$

By the same reasoning it can be deduced that for σ_y in the u-d representation the eigen-

[†]It is, however, our responsibility to remember that we have therefore chosen to express our results in the u-d representation.

system is

$$\begin{aligned}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} &= +1 \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\ \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} &= -1 \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix}.\end{aligned} \tag{24}$$

This defines the Pauli spin matrices,

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \; \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \; \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{25}$$

as expressed in the u-d representation. To get a scheme where an appropriate spin operator can be developed for an apparatus which has been given an arbitrary orientation, the spin *vector operator*, a 3-component vector whose elements are the Pauli spin matrices, is defined:

$$\sigma = [\sigma_x, \sigma_y, \sigma_z]^T = \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right]^T. \tag{26}$$

Then, if the measurement apparatus is oriented in the $\mathbf{n} = [n_x, n_y, n_z]^T$ direction, the appropriate operator to describe the dynamical variable the apparatus will observe is

$$\sigma_n = \sigma^T \mathbf{n} = \begin{bmatrix} n_z & n_x - i n_y \\ n_x + i n_y & -n_z \end{bmatrix}. \tag{27}$$

One can set about determining the eigen-decomposition of this system, and proceed to make predictions similar to those we described for σ_z, but for the general orientation.

Example: particle position

It is useful to also include the dynamical observable associated with the position of a particle constrained to move in one spatial dimension (say x) in these examples. This is because the Hilbert space is so much different for particle position. Because position is a continuous variable, the state space must be infinite in dimension, and elements of that space must be functions of x, as opposed to discrete vectors. Kets are therefore realized as functions when a basis is chosen, eigen-kets are eigen-functions, and operators acting on kets are based not on discrete inner products of matrix row elements with vector column elements, but of integrals of functions over a continuous interval in x, etc. As geophysicists, for whom model vectors are sometimes more easily represented as continuous functions, and for whom Fourier transforms, which involve continuous inner products, are both quite familiar, this is thankfully not too onerous a leap.

Let us describe the motion of a particle of mass m in 1 dimension; the position x of the particle on the interval $-\infty, \infty$ is the dynamical variable we consider. In the spin case above, we started out by choosing a base representation (u-d representation), within which the basis kets were unit vectors, i.e., were equal to $[1, 0]^T$ and $[0, 1]^T$ as a matter of definition. Subsequently, if a state were equal to one of the basis vectors, the state was

known to be up or down and experimental outcomes of +1 or -1 were certain to be observed. Taking our cue from that starting point, let us choose "position along the x axis" as our base representation. It follows that the basis kets will be the 1D continuous equivalents of unit vectors, which are Dirac delta functions. Thus, the basis ket representing the outcome of position being measured at the point x' along the x axis is

$$|x'\rangle \rightarrow \delta(x - x'). \tag{28}$$

The basis kets $|x'\rangle$, $x' \in (-\infty, \infty)$ form a complete, orthonormal basis of the continuous space of 1D functions. In general we expect our system to be described by a vector in the space spanned by these basis kets $|x'\rangle$. Let an arbitrary state vector be given by

$$|\alpha\rangle \rightarrow \alpha(x). \tag{29}$$

This $\alpha(x)$ is in fact the wave function, in the sense defined earlier.

The position of a particle is observable, so the formalism requires there to be an abstract operator X associated with it. This operator can now be specified. Consider the construction

$$xf(x, x') = x'f(x, x'), \tag{30}$$

where x' is a fixed value on the interval $x = -\infty, \infty$. Manipulating, we have

$$(x - x')f(x, x') = 0. \tag{31}$$

This equation motivates the Dirac delta function, because in order for it to be satisfied, f must be equal to zero everywhere when $x \neq x'$, but it is un-defined when $x = x'$. Thus we write

$$x\delta(x - x') = x'\delta(x - x'), \tag{32}$$

as the basic eigen-system to work with. By observing its role in this relationship, we conclude that, in the x representation, (1) to apply the operator X is to multiply by x, and (2) the eigenfunctions of X are the delta functions. The expansion of a general state function in the delta function basis is

$$\alpha(x') = \int dx\delta(x - x')\alpha(x). \tag{33}$$

In what follows we will focus on systems like this one, involving systems with continuous positions.

AVERAGES

Expectation values

Suppose a system is in a state given by the vector $|\alpha\rangle$, which inhabits a Hilbert space $\mathcal{H}_a$ of dimension N_a. Let a_i, $i = 1, ..., N_a$ be the possible outcomes of observable A with

eigenvectors $|a_i\rangle$. Take the inner product of $A|\alpha\rangle$ with $|\alpha\rangle$, and express it explicitly in terms of the expansion of $|\alpha\rangle$ in these eigenvectors:

$$\langle\alpha|A|\alpha\rangle = \left(\sum_{i=1}^{N_a} \alpha_i^* \langle a_i|\right) A \left(\sum_{j=1}^{N_a} \alpha_j |a_i\rangle\right) = \sum_{i=1}^{N_a} P_i a_i \equiv \langle A\rangle \,, \tag{34}$$

where $P_i = \alpha_i^* \alpha_i$ is the probability of measuring A to have value a_i, and $\langle a_i|a_j\rangle = \delta_{ij}$ eliminates the sum over j. The last identification comes about from recognizing the last sum of all the possible outcomes of the experiment to determine A, weighted by the probability of that outcome occurring, as the standard calculation of an average. This is generally referred to as the expectation value of A.

Pure versus mixed states and density matrices

In addition to the uncertainty already discussed within the quantum mechanics framework, which pertains to measurements we make even when we are 100% certain about the state vector, we may also be uncertain about the state vector itself. To accommodate this additional kind of uncertainty, we introduce the idea of *pure* versus *mixed* states. So far, all state vectors we have discussed have been pure. The mixed state and the density matrix associated with it emerges from a re-consideration of the averaging operation above. Take an operator $A \in \mathcal{H}_a$. Given a basis $|a_i\rangle$, $i = 1, ..., N_a$, the operator has a matrix representation

$$A_{ij} = \langle a_i|A|a_j\rangle \,. \tag{35}$$

The *trace* of A is the sum of the diagonal elements in any matrix representation of A (a quantity which is invariant under changes of basis):

$$\mathrm{Tr}A = \sum_{i=1}^{N_a} \langle a_i|A|a_i\rangle \,. \tag{36}$$

This gives us a way of re-expressing averages. Consider the operator $|\alpha\rangle\langle\alpha|A$. Its trace is

$$\mathrm{Tr}\,|\alpha\rangle\langle\alpha|A = \sum_{i=1}^{N_a} \langle a_i| \left(|\alpha\rangle\langle\alpha|A\right)|a_i\rangle = \langle\alpha|A\left(\sum_{i=1}^{N_a}|a_i\rangle\langle a_i|\right)|\alpha\rangle = \langle\alpha|A|\alpha\rangle \,, \tag{37}$$

using the identity in (11). The left-hand side is an alternative way of writing the expectation value of the observable $\langle A\rangle$ for a system in state $|\alpha\rangle$. This form is useful when we wish to incorporate uncertainty in the state vector. For instance, if we were maximally uncertain if the state was $|\alpha\rangle$ or $|\alpha'\rangle$, we could fold this into the expectation $\langle A\rangle$ by changing the operator:

$$\langle A\rangle = \mathrm{Tr}\left(\frac{1}{2}|\alpha\rangle\langle\alpha| + \frac{1}{2}|\alpha'\rangle\langle\alpha'|\right)A = \mathrm{Tr}\rho A, \tag{38}$$

where ρ is an operator which apparently could include a weighted sum of the influence of all possible states, weighted by their probability. This ρ, which in general is

$$\rho = \sum_{i=1}^{N} P_i |\alpha_i\rangle\langle\alpha_i| \,, \tag{39}$$

for N possible states $|\alpha_i\rangle$, each of which have probability P_i of being the true current state of the system, is the *density matrix*. Any ρ that is distributed over more than one state vector is a sign of the presence of a mixed as opposed to a pure state.

ENERGY AND MOMENTUM

Unitary transformations

Several very important operators, some connected with observables, are developed from the idea of a *unitary transformation*. A unitary transformation is based on a linear operator (generally not Hermitian), designed such that it enacts a smooth change to a state vector, while "preserving information". This last feature is guaranteed by constraining the operator (call it U) to be unitary, which means, mathematically, that its inverse and its Hermitian transpose are equal, i.e., $U^\dagger U = I$.

Let us consider two of these, one connected to time, and the other to space. In each case, we will need to assume that the state vectors are functions of time and/or space, such that it is meaningful to discuss changes with respect to these variables.

Let $U(dt)$ be a "time development" operator, which acts on a state vector $|\alpha(t)\rangle$ and gives the vector an infinitesimal time later, i.e.,

$$U(\delta t)\,|\alpha(t)\rangle = |\alpha(t+\delta t)\rangle\,. \tag{40}$$

The fact that it is "smooth" enters in our ability to expand U linearly about I as follows:

$$U(dt) = I - \frac{i}{\hbar}\delta t H + ..., \tag{41}$$

where the additional terms are of second order or higher in δt, and where H is another linear operator. Here $i/\hbar$ is a constant introduced for conventional reasons. Rearranging these equations, we obtain in the limit

$$\lim_{\delta t\to 0}\frac{|\alpha(t+\delta t)\rangle - |\alpha(t)\rangle}{\delta t} = \frac{d\,|\alpha(t)\rangle}{dt} = -\frac{i}{\hbar}H\,|\alpha(t)\rangle\,. \tag{42}$$

A unitary transformation in time implies that variations of state vectors in time involve an operator H; furthermore if U is unitary,

$$U^\dagger U = \left(I + \frac{i}{\hbar}\delta t H^\dagger\right)\left(I - \frac{i}{\hbar}\delta t H\right) = I + \frac{i}{\hbar}\left(H^\dagger - H\right) + ... = I, \tag{43}$$

it must be true that $H^\dagger = H$, i.e., the operator H must be Hermitian, which suggests that it corresponds to an observable in the system described by the kets $|\alpha(t)\rangle$. Exactly the same development applied to smooth variations in x (actually, by convention, $-x$) implies a Hermitian operator P where

$$\frac{d\,|\alpha(x)\rangle}{dx} = \frac{i}{\hbar}P\,|\alpha(x)\rangle\,. \tag{44}$$

Energy and momentum operators

Return to the equation involving the operator H. The letter H is used because this operator is the Hamiltonian. In classical mechanics, the Hamiltonian is identified with the total energy of the system, and that interpretation applies here as well. This emerges gradually from the analysis, and is more or less obvious depending on how one looks at the theory; for now it is most effective to think of this as a definition of energy in quantum mechanics. We have established that it is Hermitian, which means it has an eigen-decomposition

$$H\left|E_j\right\rangle = E_j\left|E_j\right\rangle \tag{45}$$

for which the $|E_j\rangle$, $j = 1, ..., N_E$ form an orthonormal basis. If we take the time-variation equation, expand the $|\alpha(t)\rangle$ in this basis, and let $a_j(t) = \langle E_j|\alpha(t)\rangle$,

$$\left[\frac{\partial}{\partial t} + \frac{i}{\hbar}H\right]\sum_{j=1}^{N_a}\left|E_j\right\rangle\left\langle E_j|\alpha(t)\right\rangle = \sum_{j=1}^{N_a}\left|E_j\right\rangle\left[\frac{\partial\alpha_j(t)}{\partial t} + \frac{i}{\hbar}E_j\alpha_j(t)\right] = 0. \tag{46}$$

Since the $|E_j\rangle$ are orthonormal, the only way this can sum to zero is if each coefficient is 0, thus:

$$\frac{\partial\alpha_j(t)}{\partial t} + i\frac{E_j}{\hbar}\alpha_j(t) = 0. \tag{47}$$

This can be integrated to produce an expression for the α_j wave functions:

$$\alpha_j(t) = \alpha_j(0)e^{-i\frac{E_j}{\hbar}t}. \tag{48}$$

A similar development of P, based on the eigendecomposition $P\left|\alpha_j(x)\right\rangle = p_j\left|\alpha_j(x)\right\rangle$ results in

$$\frac{\partial\alpha_j(x)}{\partial x} - i\frac{p_j}{\hbar}\alpha_j(x) = 0, \tag{49}$$

and

$$\alpha_j(x) = \alpha_j(0)e^{i\frac{p_j}{\hbar}x}. \tag{50}$$

We can also re-arrange the H and P equations to get explicit expressions for the H and P operators:

$$P = -i\hbar\frac{\partial}{\partial x},\ H = i\hbar\frac{\partial}{\partial t}. \tag{51}$$

Time evolution of averages and conservation

We have now that the average value of an observable A for a system in state $|\alpha\rangle$ is $\langle A\rangle = \langle\alpha|A|\alpha\rangle$. Suppose the state varies in time, i.e., $|\alpha\rangle = |\alpha(t)\rangle$. The time derivative of the average $\langle A\rangle$ is

$$\frac{\partial\left\langle A\right\rangle}{\partial t} = \left\langle\frac{\partial\alpha}{\partial t}\right|A|\alpha\rangle + \langle\alpha|A\left|\frac{\partial\alpha}{\partial t}\right\rangle = -\frac{i}{\hbar}\left\langle\alpha|HA|\alpha\right\rangle + \frac{i}{\hbar}\left\langle\alpha|AH|\alpha\right\rangle = \frac{i}{\hbar}\left\langle[A,H]\right\rangle. \tag{52}$$

The time-variation of the average value of an observable is seen to be determined by the commutator of that observable with the Hamiltonian. From this, we can conclude that an observable is conserved if it commutes with the Hamiltonian operator.

Position and momentum

Having defined P and X on the same interval and in the same representation (i.e., x on $-\infty$,∞), we can examine their commutator. In this domain it is convenient to do so by having the operators act on a wave function:

$$[X,P]\alpha(x) = x\left(-i\hbar\frac{\partial}{\partial x}\right)\alpha(x) + i\hbar\frac{\partial}{\partial x}x\alpha(x) = i\hbar\alpha(x), \tag{53}$$

implying that

$$[X,P] = i\hbar. \tag{54}$$

ENTANGLED SYSTEMS

Suppose that we have two independent systems, one described with states $|\alpha\rangle$ and the other with states $|\beta\rangle$. The systems are sufficiently different that we cannot assume $|\alpha\rangle$ and $|\beta\rangle$ are elements of the same Hilbert space, indeed we must have two, $\mathcal{H}_a$ of dimension N_a and $\mathcal{H}_b$ of dimension N_b. Because completely isolated systems need to be describable in the theory, but so must two systems which have been brought together and allowed to interact, the linear algebra formalism must have a way of combining the systems. We cannot add or multiply elements of $\mathcal{H}_a$ with elements of $\mathcal{H}_b$, but we can combine elements of both into objects which be added and multiplied. This is done with tensor products.

Tensor product states

The *tensor product* is defined as the space $\mathcal{H}_c = \mathcal{H}_a \otimes \mathcal{H}_b$, which has dimension $N_c = N_a N_b$. Elements of the space are *tensor product states* $|\gamma\rangle = |\alpha\rangle \otimes |\beta\rangle$. To see what the $|\gamma\rangle$ kets look like "in real life", suppose $N_a = 3$, $N_b = 2$, and bases are chosen in $\mathcal{H}_a$ and $\mathcal{H}_b$, such that the kets can be expressed as column vectors

$$|\alpha\rangle = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}, \quad |\beta\rangle = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}. \tag{55}$$

Then the tensor product state can be written

$$|\gamma\rangle = \begin{bmatrix} \alpha_1\beta_1 \\ \alpha_1\beta_2 \\ \alpha_2\beta_1 \\ \alpha_2\beta_2 \\ \alpha_3\beta_1 \\ \alpha_3\beta_2 \end{bmatrix}. \tag{56}$$

We can bring the basis kets used to create the α_i and β_i in to be more explicit:

$$|\gamma\rangle = |\alpha\rangle \otimes |\beta\rangle = \left(\sum_{j=1}^{N_a} \alpha_j |a_j\rangle\right) \otimes \left(\sum_{k=1}^{N_b} \beta_k |b_k\rangle\right) = \sum_{j,k} \alpha_j\beta_k |a_j\rangle \otimes |b_k\rangle . \tag{57}$$

This implies an expansion

$$|\gamma\rangle = \sum_{I=1}^{N_c} \gamma_I \, |c_I\rangle \, , \tag{58}$$

where $|c_I\rangle = |a_j\rangle \otimes |b_k\rangle$, the coefficients are

$$\gamma_I = \alpha_j \beta_k, \tag{59}$$

and the index $I = I(j,k)$ is

$$I(j,k) = k + (j-1)N_b. \tag{60}$$

Notice that if we remove (59) and (60) as restrictions on the expansion in (58), more general elements of $\mathcal{H}_c$ are built. Our assemblages of $|\alpha\rangle$ and $|\beta\rangle$ are actually very "special case" elements of $\mathcal{H}_c$. Most $|\gamma\rangle$ kets cannot be decomposed into combinations $|\alpha\rangle \otimes |\beta\rangle$. The quantum interpretation is as follows. If two systems are isolated and independent, and bringing them together does not change this, their individual states combine as a tensor product state and are so treated in in a combined theory based on $\mathcal{H}_c$. Combined systems which are not decomposable in this way correspond to entangled states – they cannot be treated as isolated, independent systems.

It is often useful to adopt a new system of vectors like the $|\gamma\rangle$, without continually writing in the $\otimes$ symbol, but to do so without hiding the $|\alpha\rangle$ and $|\beta\rangle$ origins. This is particularly true when basis vectors are being combined. So, rather than use $|c_I\rangle$, above, we would instead write the basis vector as $|a_j b_k\rangle$.

Operators

Operators generalize similarly. If A and B are operators in $\mathcal{H}_a$ and $\mathcal{H}_b$ respectively, then there is an operator $C \in \mathcal{H}_c = \mathcal{H}_a \otimes \mathcal{H}_b$ which is $C = A \otimes B$. If with bases chosen for both $\mathcal{H}_a$ and $\mathcal{H}_b$ the operators A and B have component / matrix form

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \; B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, \tag{61}$$

then C has matrix form

$$C = \begin{bmatrix} A_{11}B_{11} & A_{11}B_{12} & A_{12}B_{11} & A_{12}B_{12} & A_{13}B_{11} & A_{13}B_{12} \\ A_{11}B_{21} & A_{11}B_{22} & A_{12}B_{21} & A_{12}B_{22} & A_{13}B_{21} & A_{13}B_{22} \\ A_{21}B_{11} & A_{21}B_{12} & A_{22}B_{11} & A_{22}B_{12} & A_{23}B_{11} & A_{23}B_{12} \\ A_{21}B_{21} & A_{21}B_{22} & A_{22}B_{21} & A_{22}B_{22} & A_{23}B_{21} & A_{23}B_{22} \\ A_{31}B_{11} & A_{31}B_{12} & A_{32}B_{11} & A_{32}B_{12} & A_{33}B_{11} & A_{33}B_{12} \\ A_{31}B_{21} & A_{31}B_{22} & A_{32}B_{21} & A_{32}B_{22} & A_{33}B_{21} & A_{33}B_{22} \end{bmatrix}, \tag{62}$$

wherein each 2×2 subset of C is one element of A multiplied by B in its entirety. Operators A, previously defined on $\mathcal{H}_a$, and designed to be enacted on the "a part" of a vector in $\mathcal{H}_c$ are

still meaningful in this new context, but they must be augmented as $C_a = A \otimes I$. Likewise, operators which act only on the "b part" of a vector in $\mathcal{H}_c$ are of the form $C_b = I \otimes B$.

Like with vectors in $\mathcal{H}_c$, elements of matrices in $\mathcal{H}_c$ are indexed with combinations of the indices used within $\mathcal{H}_a$ and $\mathcal{H}_b$. It is useful for indices labelling the N_c dimensions of $\mathcal{H}_c$ to "remember" their relationships to the indices labelling the original N_a and N_b labels. Recall that once a basis has been established in $\mathcal{H}_a$, say $|a_i\rangle$, $i = 1, ..., N_a$, an operator can be expressed as a matrix, with elements

$$A_{ij} = \langle a_i | A | a_j \rangle \, . \tag{63}$$

This is suggestive that a similar matrix element expression is possible for C:

$$C_{jk,lm} = \langle a_j b_k | C | a_l b_m \rangle \, . \tag{64}$$

When using terminology like this, it is important to remember that although four indices are invoked, this remains a matrix, i.e., an array with rows and columns. The j, k pair is a label for one row, and the l, m pair is a label for one column. This can be made clearer by using the capital indices

$$C_{I,J} = \langle a b_I | C | a b_J \rangle \, , \tag{65}$$

where $I = I(j,k) = k + (j-1)N_b$ and $J = J(l,m) = m + (l-1)N_b$.

Density matrices and entangled states

Suppose we have a system in state $|\gamma\rangle$, which is in $\mathcal{H}_c = \mathcal{H}_a \otimes \mathcal{H}_b$. The expectation value of an observable C is

$$\langle \gamma | C | \gamma \rangle = \left(\sum_{j=1}^{N_a} \sum_{k=1}^{N_b} \gamma_{jk}^* \langle a_j b_k | \right) C \left(\sum_{l=1}^{N_a} \sum_{m=1}^{N_b} \gamma_{lm} | a_l b_m \rangle \right) = \sum_{j=1}^{N_a} \sum_{k=1}^{N_b} \gamma_{jk}^* \gamma_{jk} c_{jk}, \tag{66}$$

using $\langle a_j b_k | a_l b_m \rangle = \delta_{jk,lm}$, and where c_{jk} is the jkth eigenvalue of C. Again, the pairs j, k and l, m act as single indices in all of these quantities. This is the combined-system version of the expectation calculation in (34).

Consider the special case of a pure (though possibly entangled) state $|\gamma\rangle$ being investigated for its behaviour regarding system A only, i.e., in which the operator $C = C_a = A \otimes I$. Then

$$\langle \gamma | C_a | \gamma \rangle = \sum_{j=1}^{N_a} \sum_{k=1}^{N_b} \sum_{l=1}^{N_a} \sum_{m=1}^{N_b} \gamma_{jk}^* \gamma_{lm} A_{jl} \delta_{km} = \sum_{j=1}^{N_a} \sum_{l=1}^{N_a} \rho_{jl} A_{jl}, \tag{67}$$

where

$$\rho_{jl} = \sum_{k=1}^{N_b} \gamma_{jk}^* \gamma_{lk}. \tag{68}$$

We observe that the average in such an observation – although it concerns only the "a" part of the system, and although $|\gamma\rangle$ is pure, naturally generates a density matrix, as if the state were mixed. This is the character of an entangled state: the presence of the second system acts to produce uncertainty in the first, even for pure combined states.

HARMONIC OSCILLATORS

We will focus on harmonic oscillators in this review; a harmonic oscillator in 1D with equilibrium point x_0 is a particle with kinetic energy and Hookean potential energy terms in the Hamiltonian:

$$H = \frac{p_x^2}{2m} + \frac{1}{2}k'(x - x_0)^2 = \frac{1}{2}p_x^2 + \frac{1}{2}\left(kx - l_0\right)^2, \tag{69}$$

where for convenience we write $k = \sqrt{k'}$ and $l_0 = kx_0$. The same substitution procedure gives

$$H = \frac{1}{2}P_x^2 + \frac{1}{2}Q_x^2, \tag{70}$$

where $Q_x = kX - L_0$; an expression with two important dynamical variables P_x and X satisfying

$$[P_x, Q_x] = i\hbar k. \tag{71}$$

The time-independent Schrödinger equation is associated with the eigen-decomposition of the Hamiltonian operator, i.e.,

$$H\left|n\right\rangle = E_n\left|n\right\rangle. \tag{72}$$

The special form H takes for harmonic oscillators allows for it to be discussed in terms of what are termed *ladder operators*, which are also known as *creation* and *annihilation*, and *raising* and *lowering* operators. We observe that the H developed in (70) can be factored to produce:

$$H = \frac{1}{2}\big(P_x + iQ_x\big)\big(P_x - iQ_x\big) + \frac{\hbar}{2}k, \tag{73}$$

where we have used the commutator in (**??**). Let

$$L \equiv \frac{i}{\sqrt{2\hbar k}}\big(P_x - iQ_x\big) \tag{74}$$

be the *lowering*, or *annihilation* operator, and

$$L^\dagger \equiv -\frac{i}{\sqrt{2\hbar k}}\big(P_x + iQ_x\big) \tag{75}$$

be the *raising*, or *creation* operator, and then further define

$$N \equiv L^\dagger L, \tag{76}$$

such that

$$H = \hbar k\left(N + \frac{1}{2}\right), \text{ and } N = \frac{1}{\hbar k}H - \frac{1}{2}. \tag{77}$$

One can quickly confirm that the three quantities L, $L^\dagger$ and N satisfy

$$[L, L^\dagger] = 1,\ [L^\dagger, L] = -1,\ [L, N] = L,\ \text{and}\ [L^\dagger, N] = -L^\dagger. \tag{78}$$

The usefulness of the ladder operators comes about as follows. Since N and H are distinct only via trivial linear operations, the eigenvectors $|n\rangle$ in $H|n\rangle = E_n|n\rangle$ are also the eigenvectors of N:

$$N|n\rangle = n|n\rangle\,. \tag{79}$$

Operating on both sides with $L^\dagger$, and applying the commutator rule $L^\dagger N = NL^\dagger - L^\dagger$, we get

$$N\big(L^\dagger|n\rangle\big) = (n+1)\big(L^\dagger|n\rangle\big). \tag{80}$$

Evidently, acting on $|n\rangle$ with $L^\dagger$ gives a new eigenvector of N, with eigenvalue greater than that of $|n\rangle$ by 1. This can sensibly be called $|n+1\rangle$. Thus a recipe for a calculation of eigenvectors and eigenvalues by "laddering" emerges:

$$|n+1\rangle = L^\dagger|n\rangle\,,\ \ |n-1\rangle = L|n\rangle\,. \tag{81}$$

From this analysis also comes the insight that the eigenvalues of H are separated from each other by integral values. Specifically, using (**??**),

$$N|n\rangle = n|n\rangle \rightarrow H|n\rangle = \hbar k\left(n + \frac{1}{2}\right)|n\rangle\,, \tag{82}$$

meaning that the energies accessible to a harmonic oscillator are $E_n = \hbar k\,(n + 1/2)$. We add to this that n are positive integers or zero. Because the eigenvalues of N are integers, i.e., numbers, there is already some motivation for us to refer to N as the number operator. The number relates to the energy level the harmonic oscillator is currently in. Later, however, when there may be more than one (in fact unspecified) harmonic oscillators in a system, there will be more obvious reasons to refer to N this way. So, like the *creation* and *annihilation* operators, the name *number* operator only makes complete sense from later uses.

Degeneracy

A situation like the 1D harmonic oscillator is non-degenerate, as is any system where a single eigenvector belongs to each eigenvalue. This permits eigenvalues to uniquely label eigenvectors, a convention we have already been using (i.e., eigenvalue n is associated with eigenvector $|n\rangle$). In degenerate systems, more than one eigenvector has the same eigenvalue. If an operator like H has non-degenerate eigenvalues, it is said to constitute a *complete set of commuting observables*, or CSCO; if eigenvalues are degenerate, it is not. In some situations, operators which do not constitute CSCOs can be combined with other operators in commutators which are CSCOs.

Ground state of the harmonic oscillator

To proceed using the ladder operators requires one solution to be available, after which the others can be derived. Consider the case $n = 0$. Evidently $E_0 = \hbar k/2$, and

$$H\,|0\rangle = \left(\frac{P_x^2}{2} + \frac{Q_x^2}{2}\right)|0\rangle = \frac{\hbar k}{2}\,|0\rangle\,. \tag{83}$$

In the x representation let $|0\rangle$ have the wave function $\alpha_0(x)$. The operators are likewise $P_x^2 \to -\hbar^2\partial^2/\partial x^2$ and $Q_x^2 \to (kx - l_0)^2$, so the equation requiring solution is

$$-\frac{\hbar^2}{2}\frac{\partial^2}{\partial x^2}a_0(x) = \left[\frac{\hbar k}{2} - \frac{1}{2}(kx - l_0)^2\right]a_0(x). \tag{84}$$

One may quickly confirm that this equation is satisfied by

$$a_0(x) = a_0 \exp\left[-\frac{k}{2\hbar}\big(x - l_0/k\big)^2\right]. \tag{85}$$

Excited states

The ladder operators in the x representation are

$$L = \frac{i}{\sqrt{2\hbar k}}\big(P_x - iQ_x\big) \to \frac{1}{\sqrt{2\hbar k}}\left(\hbar\frac{\partial}{\partial x} + (kx - l_0)\right) \tag{86}$$

$$L^\dagger = -\frac{i}{\sqrt{2\hbar k}}\big(P_x + iQ_x\big) \to \frac{1}{\sqrt{2\hbar k}}\left(-\hbar\frac{\partial}{\partial x} + (kx - l_0)\right). \tag{87}$$

If we apply the raising operator $L^\dagger$ to the ground state solution, we produce the first excited state:

$$L^\dagger\,|0\rangle \to \frac{1}{\sqrt{2\hbar k}}\left(-\hbar\frac{\partial}{\partial x} + (kx - l_0)\right)a_0(x) = \sqrt{\frac{2k}{\hbar}}\left(x - \frac{l_0}{k}\right)a_0(x). \tag{88}$$

Meanwhile, if we apply the lowering operator to the ground state solution, we produce nil:

$$L\,|0\rangle \to \frac{1}{\sqrt{2\hbar k}}\left(\hbar\frac{\partial}{\partial x} + (kx - l_0)\right)a_0(x) = 0. \tag{89}$$

This process continues:

$$a_2(x) = L^\dagger(x)a_1(x) = L^\dagger(x)L^\dagger(x)a_1(x), \tag{90}$$

and then

$$a_n(x) = L^\dagger(x)a_{n-1}(x) = L^\dagger(x)...L^\dagger(x)a_0(x) = (L^\dagger)^n(x)a_0(x). \tag{91}$$

The states themselves can be plotted by plotting the wave functions $a_i(x)$, which can

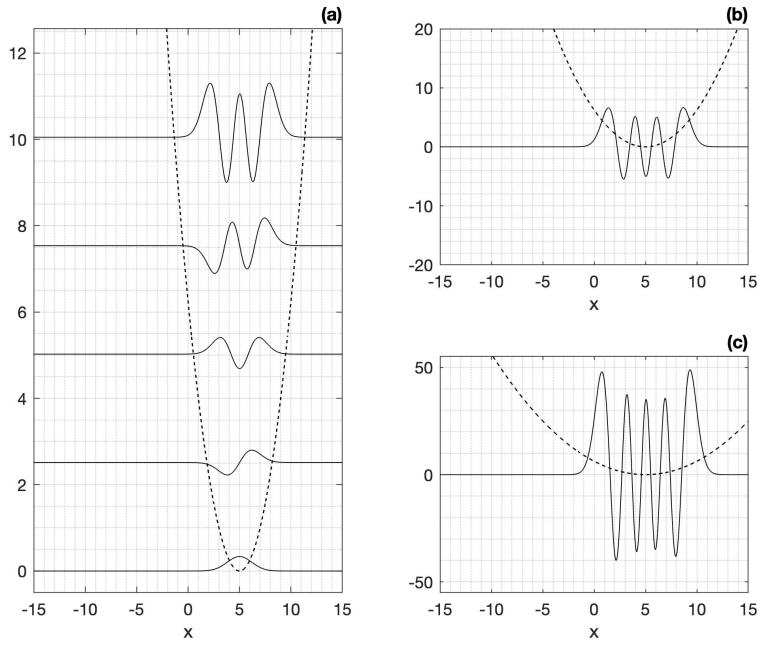

FIG. 1. (a) ground and first four excited states of the 1D harmonic oscillator. (b)-(c) 6th and 8th excited states.

be generated analytically, or (as we do here) numerically by producing a matrix which approximates the raising operator and applying it repeatedly to a discretized version of $a_0(x)$. In Figure 1a ground state and the first four excited states are plotted; in (b) and (c) the 6th and 9th excited states are plotted respectively. Of note, as the energy rises, the wave function has increasing energy in regions away from the equilibrium point of the oscillator.

Typically we envision particles (e.g., harmonic oscillators) as being able to move in up to 3 dimensions, but because we have in mind for later applications notional, high dimensional physical spaces, we will let the dimensions be arbitrary in number $M > 0$. That is, we will involve spatial directions x_j where $j = 1, ..., M$. Each direction in which an oscillator can move contributes energy to the system. If the restoring force in direction x_j is $-(k_j x_j - l_j)$, this suggests a Hamiltonian of

$$H = \sum_{j=1}^{M} H_j, \tag{92}$$

where

$$H_j = \frac{1}{2}P_j^2 + \frac{1}{2}Q_j^2, \tag{93}$$

and where $P_j = -i\hbar\partial/\partial x_j$ and $Q_j = k_j X_j - L_j$. We refer to this as a *separable* problem, primarily because of the conditions it puts on the elastic potential energy; if the elastic potential energy contribution of direction $j = j'$ had depended on the position of the oscillator in any other direction $j \neq j'$, H would not have been separable into H_j's, and would instead have been a more complicated quadratic form. However, because all applications we will study can be transformed into a separable form, this level of generality will do for now.

Because each H_j is the Hamiltonian of a 1D harmonic oscillator, identical to the case we treated above, we can assume for this analysis that we have access to eigenvalues and vectors

$$H_j \left|n_j\right\rangle = E_{n_j} \left|n_j\right\rangle, \; j = 1, ..., M. \tag{94}$$

The eigen-decomposition we seek is that of the full problem, i.e., H. In a separable problem of this kind, the full problem can be considered as an assemblage of 1D oscillators, where the assemblage is formed in the manner we combined systems in the discussion of entanglement. Consider the product

$$\left|n_1 n_2 ... n_M\right\rangle = \left|n_1\right\rangle \otimes \left|n_2\right\rangle \otimes ... \otimes \left|n_M\right\rangle. \tag{95}$$

Each H_j acts on the jth sub-vector within such combinations, with the other vectors acting as spectators, i.e., $H_j \left|n_1 ... n_M\right\rangle = E_{n_j} \left|n_1 ... n_M\right\rangle$. The full H being the sum of the H_j's implies then that the vector $\left|n_1 ... n_M\right\rangle$ must be an eigenvector satisfying

$$H \left|n_1 ... n_M\right\rangle = \left(H_1 + ... + H_M\right) \left|n_1 ... n_M\right\rangle = \left(E_{n_1} + ... + E_{n_M}\right) \left|n_1 ... n_M\right\rangle, \tag{96}$$

and the allowable energies of the system are

$$E_n = E_{n_1} + ... + E_{n_M} = \hbar k_1 \left(n_1 + \frac{1}{2}\right) + ... + \hbar k_M \left(n_M + \frac{1}{2}\right), \tag{97}$$

where the n_j are positive integers or zero. Since, unless the k_j are very specially chosen, this E_n can be got through many combinations of n_j values, in general it is degenerate. The "fold" of its degeneracy, i.e., the number of combinations of n_j which produce E_n, is determined by the k_j values.

In determining the detailed form of the stationary states, in the x_j representation, ladder operators are applicable in the multidimensional case as well as in 1D. The generalization is

$$L_j \equiv \frac{i}{\sqrt{2\hbar k_j}}\left(P_j - iQ_j\right), \; L_j^\dagger \equiv -\frac{i}{\sqrt{2\hbar k_j}}\left(P_j + iQ_j\right), \tag{98}$$

which, with $N_j = L_j^\dagger L_j$, satisfy the commutation rules

$$[L_j, L_k^\dagger] = \delta_{jk}, \; [L_j, N_k] = \delta_{jk} L_j, \; [L_j^\dagger, N_k] = -\delta_{jk} L_j^\dagger. \tag{99}$$

These rules lead to expressions for any desired $\left|n_1 ... n_M\right\rangle$, given a starting point like the ground state $\left|0...0\right\rangle$, through raising and lowering:

$$\left|n_1 ... n_M\right\rangle = \frac{1}{\sqrt{n_1! ... n_M!}} (L_1)^{n_1} ... (L_M)^{n_M} \left|0...0\right\rangle. \tag{100}$$

CONCLUSIONS

Our group's continuing (though gradual) development of quantum algorithms with particular application in applied geophysics, as well as our growing interest in characterizing inversion uncertainty with quasi-dynamical systems (inspired by the so-called Hamiltonian Monte Carlo methods), both motivate some review work on the linear algebra of quantum mechanics. The description of simple systems, in isolation and in combination, is developed, with the latter allowing simple entangled states to be discussed, which is relevant in quantum information and computing. *Laddering operators* in the description of harmonic oscillators are also given special focus: since in the above mentioned Hamiltonian MC methods, energy and data misfit are associated, laddering operators may be useful descriptors of convergence. Steps needed to increase the dimensionality of systems, and steps needed to increase the number of particles/oscillators in systems are also given attention, since both will be needed in evolving applications in optimization and geophysical quantum computing.

ACKNOWLEDGEMENTS

The sponsors of CREWES are gratefully thanked for continued support. This work was funded by CREWES industrial sponsors, NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19, and in part by an NSERC-DG.

REFERENCES

Dirac, P. A. M., 1930, The principles of quantum mechanics: Oxford University Press, 1st edn.

Fichtner, A., Zunino, A., Gebraad, L., and Boehm, C., 2021, Autotuning Hamiltonian Monte Carlo for efficient generalized nullspace exploration: Geophysical Journal International, **227**, 941–968.

Harris, E. G., 1972, A pedestrian approach to quantum field theory: Dover Books, 1st edn.

Innanen, K. A., 2021, The statistical mechanics of model space shuttles: CREWES Research Report, **33**.

Keating, S. D., and Innanen, K. A., 2021, Null-space shuttles for targeted uncertainty analysis in full-waveform inversion: Geophysics, **86**, No. 1, R63–R76.

Monsegny, J., 2022, Geophysical inverse problems with quantum computing: CREWES Research Report, **34**.

Moradi, S., Trad, D. O., and Innanen, K. A., 2018, On the use of quantum computation in exploration seismology: CREWES Research Report, **30**.

Susskind, L., and Friedman, A., 2015, Quantum mechanics: the theoretical minimum: Basic Books, 1st edn.