Introduction to Quantum Mechanics

William Gordon Ritter[1]
[1]*Harvard University Department of Physics*
17 Oxford St., Cambridge, MA 02138

(Dated: January 11, 2005)

I. THE HARMONIC OSCILLATOR

Define operators

$$a^* = \frac{1}{\sqrt{2}}(X - D) \quad \text{and} \quad a = \frac{1}{\sqrt{2}}(X + D),$$

where X is multiplication by x and D is the derivative with respect to x. It follows that $a + a^* = \sqrt{2}X$, so $a^* = \sqrt{2}X - a$. Define f_0 to be such that $af_0 = 0$. Any solution of this equation is a constant times $f_0 = e^{-x^2/2}$. Define $f_n = (a^*)^n f_0$, which implies $af_n = nf_{n-1}$ and

$$\begin{aligned} f_{n+1} &= a^* f_n = (\sqrt{2}X - a)f_n \\ &\Rightarrow \quad f_{n+1}(x) = \sqrt{2}\, x f_n(x) - n f_{n-1}(x). \end{aligned}$$

This is known as a *recurrence relation*. Its most natural representation is in terms of a generating function. Suppose there exists $F(x, z)$ such that

$$F(x, z) = \sum_{n=0}^{\infty} \frac{f_n(x)}{n!} z^n \,.$$

We may then calculate the derivative with respect to z,

$$\begin{aligned} D_z F &= \sum_{n=1} \frac{f_n}{(n-1)!} z^{n-1} = \sum_{n=0} \frac{f_{n+1}}{n!} z^n \\ &= \sum_{n=0} \sqrt{2} x \frac{f_n}{n!} z^n - \sum_{n=1} n \frac{f_{n-1}}{n!} z^n \\ &= \sqrt{2} x F - z \sum_{n=1} \frac{f_{n-1}}{(n-1)!} z^{n-1} = (\sqrt{2}x - z)F. \end{aligned}$$

The separable differential equation $D_z F = (\sqrt{2}x - z)F$ has solution $F(x, z) = e^{\sqrt{2}xz - z^2/2}$.

In quantum mechanics, momenta are represented as eigenvalues of the operator $P = -iD$. The energy H is the sum of the kinetic energy, $P^2/(2m)$, and the potential energy, given by a function $V(x)$. For the quadratic potential $V(x) = \frac{1}{2}(x^2 - 1)$, we have

$$H \equiv \frac{1}{2}P^2 + V(X) = \frac{1}{2}(-D^2 + X^2 - 1) = a^* a.$$

Theorem 1. *$Hf_n = nf_n$ for all integers $n \geq 0$.*

Proof. For any operators A, B, define $\mathscr{D}_A B = [A, B]$ and note that $\mathscr{D}_A$ satisfies the product rule for derivatives, $\mathscr{D}_A(BC) = (\mathscr{D}_A B)C + B\mathscr{D}_A C$. We now prove that $[a, (a^*)^n] = n(a^*)^{n-1}$. This is obvious for $n = 0$, and for $n = 1$ it is the statement that $\mathscr{D}_a a^* = 1$, which we have checked previously. We therefore use induction, assuming that $\mathscr{D}(a^*)^{n-1} = (n-1)(a^*)^{n-2}$. By the product rule,

$$\begin{aligned} \mathscr{D}_a(a^*\,(a^*)^{n-1}) &= (\mathscr{D}_a a^*)(a^*)^{n-1} + a^*\mathscr{D}_a(a^*)^{n-1} \\ &= (a^*)^{n-1} + a^*(n-1)(a^*)^{n-2} = n(a^*)^{n-1}\,. \end{aligned}$$

Then

$$\begin{aligned} Hf_n &= a^*a(a^*)^n f_0 = (a^*)^{n+1}\underbrace{af_0}_{=0} + a^*[a, (a^*)^n]f_0 \\ &= a^*n(a^*)^{n-1}f_0 = nf_n\,. \qquad \square \end{aligned}$$

Eigenvalues of H are identified with possible energies of the system; this analysis shows that the oscillator energies are discrete. Aside from their mathematical interest as an orthogonal basis for the space $\{\psi \mid \int_{-\infty}^{\infty} |\psi(x)|^2\,dx < \infty\}$, the eigenfunctions f_n are also physically meaningful, since $|f_n(x)|^2$ represents the probability distribution for a particle's position, if the particle is in the n^{th} excited state.[1] Graphs of the first few are shown in Fig. 1.

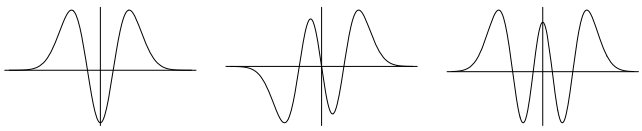

FIG. 1 Plots of f_2, f_3, and f_4.

II. MATRIX REPRESENTATIONS OF SU_2

Given a set of commutation relations $[A, B] = C$ etc, an n-dimensional *matrix representation* of the algebra is a set of $n \times n$ matrices A, B, C, etc. which satisfy the given relations. The description of spin in quantum mechanics uses matrix representations of the following algebra:

$$[h, x] = 2x, \quad [h, y] = -2y, \quad [x, y] = h\,. \tag{1}$$

[1] The last statement is only true if we divide f_n by a constant equal to the total area under its graph, so it becomes a valid probability distribution function.

This algebra is often called $\mathfrak{sl}_2$ due to its fundamental representation in terms of 2×2 matrices:

$$x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{2}$$

The choice of the coefficients $2, -2$, in equation (1) is unimportant, in the sense that we can multiply x, y, and h by three arbitrary scale factors and obtain results similar to what will follow here. The choice of coefficients in (1) is standard in the mathematics literature (though not the standard in physics) and causes certain other calculations to work out simply; for example, the matrix elements in (2) are all 1 or -1.

In the following, we assume that x, y, and h are operators which act on a finite-dimensional vector space V and which satisfy the commutation relations (1). From this (minimal) set of assumptions, we will obtain much information. If V is one-dimensional, then all matrices act as multiplication by a single number, and we cannot have nontrivial commutation relations, so we assume $\dim V > 1$.

Suppose that $v \in V$ is an eigenvector of h with eigenvalue λ, so $hv = \lambda v$. Then using (1),

$$2xv = [h, x]v = hxv - xhv = h(xv) - \lambda xv \, .$$

Therefore $h(xv) = (2 + \lambda)xv$. In other words, xv is another eigenvector, with eigenvalue $\lambda + 2$. To organize these calculations, let V_ρ denote all eigenvectors in V with eigenvalue ρ. We have just shown that $x(V_\lambda) \subset V_{\lambda+2}$, and a similar calculation shows that $y(V_\lambda) \subset V_{\lambda-2}$, so we can view x and y as raising and lowering operators for eigenvalues of h, similar to the way a^* and a are raising and lowering operators for the eigenvalues of the oscillator Hamiltonian $H = a^*a$.

Let X and A be any two noncommuting operators. We will need the following formula for the commutator of X with a power of A (which may easily be proved by induction)

$$[X, A^n] = \sum_{k=1}^{n} A^{k-1}[X, A]A^{n-k} \, . \tag{3}$$

Since V is finite-dimensional, h can have at most finitely many different eigenvalues; call them $\lambda_1, \ldots, \lambda_M$. In general, the eigenvalues λ_i could be complex, but we will see this is not the case. Choose λ_{max} to be the eigenvalue with the largest real part. Let v_0 be a nonzero eigenvector with eigenvalue λ_{max} (in this case, v_0 is called a *maximal vector*).

Theorem 2. *λ_{max} is an integer.*

Proof. Since x is a raising operator, $xv_0 = 0$. Then

$$\begin{aligned} xy^n v_0 &= \underbrace{y^n x v_0}_{=0} + [x, y^n]v_0 = \sum_{k=1}^{n} y^{k-1} h y^{n-k} v_0 \\ &= \sum_{k=1}^{n} y^{k-1}(\lambda_{max} - 2(n-k)) y^{n-k} v_0 \\ &= \Big(\sum_{k=1}^{n} (\lambda_{max} - 2n + 2k)\Big) y^{n-1} v_0 \\ &= \Big(n(\lambda_{max} - 2n) + 2\frac{n(n+1)}{2}\Big) y^{n-1} v_0 \\ &= n(\lambda_{max} - n + 1) y^{n-1} v_0. \end{aligned} \tag{4}$$

Since the latter formula is valid for all n, it is in particular valid for the largest n such that $y^{n-1}v_0 \neq 0$. Since that n was the largest (call it N), we have $y^N v_0 = 0$ and hence the left side of equation (4) is zero. Therefore,

$$N(\lambda_{max} - N + 1) y^{N-1} v_0 = 0\,.$$

If $N = 0$ then $v_0 = 0$, a contradiction. We conclude that $\lambda_{max} = N - 1$. □

In the process, we have proven that the dimension of the space in which the operators act is $N = \lambda_{max} + 1$.

Theorem 3. *Consider a matrix representation with no invariant subspace. If* $[Q, X] = 0$ *for all* X *in the algebra, and if* $Ker(Q) \neq \{0\}$ *then* $Q = 0$.

Proof. Let $v \neq 0$ be an element of $Ker(Q)$, so $Qv = 0$. Then $QXv = XQv = 0$, so X preserves $Ker(Q)$. Then $Ker(Q)$ is a nontrivial invariant subspace, but this is a contradiction. □

In the situation of Theorem 3, we infer that if $[A, X] = 0$ for all X in the algebra, then $A = \lambda I$ for some λ, where I is the identity matrix. This result is called *Schur's Lemma*; the proof is to apply Theorem 3 with $Q = A - \lambda I$ where λ is an eigenvalue of A.

III. ANGULAR MOMENTUM

In quantum mechanics, the momentum 3-vector $\boldsymbol{p}$ is represented by the operator $-i\nabla$, where ∇ is the gradient. The angular momentum $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ then has three components, which are operators satisfying $[L_1, L_2] = iL_3$. If we define $x = L_1 + iL_2$, $y = L_1 - iL_2$, and $h = 2L_3$, then we have the commutation relations (1). As the eigenvalues of h are integers separated by two, the eigenvalues of L_3 must be half-integers separated by one. Thus the representation with highest L_3 eigenvalue given by ℓ must have dimension $2\ell + 1$ (note: 2ℓ is λ_{max} for h).

Note that $L^2 = \boldsymbol{L} \cdot \boldsymbol{L}$ commutes with L_1, L_2, L_3 and hence, by Schur's Lemma, $L^2 = \lambda I$ in this representation, so every vector is an eigenvector of L^2.

In this context, the eigenvalue equation for L_3 and for L^2 are differential equations, whose solutions are the spherical harmonics. The latter are special functions which, up to a constant factor, take the form

$$e^{im\phi} P_l^m(\cos\theta), \quad -l \leq m \leq l,$$

in spherical coordinates and which are responsible for the odd-looking shape of electron orbitals.

In quantum mechanics, the possible outcomes of a measurement are identified with the possible eigenvalues of an operator. In this spirit, possible measurements of the z-component of angular momentum correspond to allowed eigenvalues of L_3. Theorem 2 and the surrounding discussion then imply that these measurement outcomes are not arbitrary; the highest one, ℓ, must be a half-integer and there are $2\ell + 1$ eigenvectors, and L_- lowers the eigenvalue by one each time, so the possible outcomes are $m \in \{\ell, \ell - 1, \ldots, -\ell\}$. The angular dependence of the corresponding wave function is proportional to $e^{im\phi} P_l^m(\cos\theta)$. Moreover, higher ℓ correspond to higher energy, so the distinct values of ℓ give the distinct orbitals in an atom, in order of increasing excitation energy.