A possible application of commutators in stationary perturbation theory

1980 J. Phys. A: Math. Gen. 13 2007

(http://iopscience.iop.org/0305-4470/13/6/023)

J. Phys. A: Math. Gen. **13** (1980) 2007–2012. Printed in Great Britain

A possible application of commutators in stationary perturbation theory

Anthea Grubb

Department of Mathematics, Brunel University, Kingston Lane, Uxbridge, Middlesex, UK

Received 12 November 1979

Abstract. For an observable which can be represented by a self-adjoint operator belonging to a rather restrictive class, it is shown that the problem of calculating the expectation value up to the nth order is equivalent to solving n commutator equations. In the first order the theory yields *naturally* the interchange theorem for *any* observable. For the calculation of the first-order corrections, the theory leads to a differential equation which is solved explicitly for a wide class of multiplication operators in the configuration space.

1. Introduction

As in the usual perturbation procedure, it is assumed that the state vector and the energy eigenvalue for a stationary state of an isolated quantum system have Taylor expansions in powers of a parameter λ in the Hamiltonian. It is found that the coefficients in the corresponding expansion of the expectation value of an observable take a particularly simple and nice form in terms of certain commutators. On careful examination, the class of observables to which the theory developed in this paper is applicable seems very restrictive. Nevertheless, the first-order theory is applicable to any observable and yields *naturally* the interchange theorem of Dalgarno (Dalgarno and Lewis 1955), something which is quite useful. This indicates that, with some further development and modification, the theory could find application in the calculation of the expectation value of any observable to any order.

2. The perturbation expansions

Operators are distinguished from other objects by putting a circumflex over symbols representing them. It is assumed that an isolated stationary quantum system has a Hamiltonian which can be written in the form $\hat{H}_0+\lambda\hat{H}_1$, where λ can be regarded as a perturbation parameter, and that the eigenvectors and the eigenvalues of the Hamiltonian have Taylor expansions in powers of λ. Let $|\psi\rangle$ be a particular eigenvector of $\hat{H}_0+\lambda\hat{H}_1$ belonging to the eigenvalue E. Then

$$|\psi\rangle = \sum_{i=0}^{\infty} \lambda^i |\psi_i\rangle \tag{2.1}$$

and

$$E = \sum_{i=0}^{\infty} \lambda^i E_i, \tag{2.2}$$

0305-4470/80/062007+06$01.50

where ψ_i and E_i satisfy

$$(\hat{H}_0 - E_0\hat{I})|\psi_0\rangle = 0 \tag{2.3a}$$

and

$$(\hat{H}_0 - E_0\hat{I})|\psi_n\rangle = \sum_{i=1}^{n} E_i|\psi_{n-i}\rangle - \hat{H}_1|\psi_{n-1}\rangle \qquad \text{for all } n > 1, \tag{2.3b}$$

$\hat{I}$ being the identity operator.

Let $\hat{A}$ be a self-adjoint operator representing some observable. Then the expectation value $\langle\hat{A}\rangle$ of $\hat{A}$ has the expansion

$$\langle\hat{A}\rangle = \sum_{n=0}^{\infty} \lambda^n \langle\hat{A}\rangle_n, \tag{2.4}$$

where

$$\langle\hat{A}\rangle_n = \sum_{i=0}^{n} \langle\psi_i|\hat{A}|\psi_{n-i}\rangle = \sum_{i=0}^{n} \langle\psi_{n-i}|\hat{A}|\psi_i\rangle. \tag{2.4a}$$

Suppose now that there exists an operator $\hat{U}_1$ such that

$$[\hat{H}_0, \hat{U}_1] = \hat{A}, \tag{2.5}$$

or equivalently

$$[(\hat{H}_0 - E_0\hat{I}), \hat{U}_1] = \hat{A}. \tag{2.5a}$$

Then

$$\begin{aligned}
\langle\hat{A}\rangle_n &= \sum_{i=0}^{n} \langle\psi_i|[(\hat{H}_0 - E_0\hat{I}), \hat{U}_1]|\psi_{n-i}\rangle \\
&= \sum_{i=0}^{n} (\langle\psi_i|(\hat{H}_0 - E_0\hat{I})\hat{U}_1|\psi_{n-i}\rangle - \langle\psi_{n-i}|\hat{U}_1(\hat{H}_0 - E_0\hat{I})|\psi_i\rangle) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} E_j\langle\psi_{i-j}|\hat{U}_1|\psi_{n-i}\rangle - \sum_{i=1}^{n}\sum_{j=1}^{n} E_j\langle\psi_{n-i}|\hat{U}_1|\psi_{i-j}\rangle \\
&\quad - \sum_{i=1}^{n} (\langle\psi_{i-1}|\hat{H}_1\hat{U}_1|\psi_{n-i}\rangle - \langle\psi_{n-i}|\hat{U}_1\hat{H}_1|\psi_{i-1}\rangle) \\
&= \sum_{j=1}^{n}\sum_{i=j}^{n} E_j(\langle\psi_{i-j}|\hat{U}_1|\psi_{n-i}\rangle - \langle\psi_{i-j}|\hat{U}_1|\psi_{n-i}\rangle) + \sum_{i=1}^{n} \langle\psi_{n-i}|\hat{U}_1\hat{H}_1 - \hat{H}_1\hat{U}_1|\psi_{i-1}\rangle \\
&= \sum_{i=0}^{n-1} \langle\psi_{n-1-i}|[\hat{U}_1, \hat{H}_1]|\psi_i\rangle \\
&= \langle[\hat{U}_1, \hat{H}_1]\rangle_{n-1}.
\end{aligned} \tag{2.6}$$

In arriving at the final expression in (2.6), apart from using (2.3*a*, *b*) several rearrangements of the various summands have been made.

Next suppose that there exists an operator $\hat{U}_2$ with the property that

$$[\hat{H}_0, \hat{U}_2] = [\hat{U}_1, \hat{H}_1]; \tag{2.7}$$

then, by what has just been proved,

$$\langle\hat{A}\rangle_n = \langle[\hat{U}_1, \hat{H}_1]\rangle_{n-1} = \langle[\hat{U}_2, \hat{H}_1]\rangle_{n-2}, \tag{2.8}$$

and, continuing in this way, one finally has

$$\langle \hat{A} \rangle_n = \langle [\hat{U}_n, \hat{H}_1] \rangle_0. \tag{2.9}$$

Thus, instead of solving the perturbation equations (2.3*a*, *b*), if one can solve an equal number of commutator equations

$$[\hat{H}_0, \hat{U}_1] = \hat{A} \tag{2.10a}$$

and

$$[\hat{H}_0, \hat{U}_n] = [\hat{U}_{n-1}, \hat{H}_1] \qquad \text{for all } n > 1, \tag{2.10b}$$

then $\langle \hat{A} \rangle_n$ is simply given by (2.9).

The simplicity of expression (2.9) is remarkable. Admittedly it is much more difficult to solve commutator equations (2.10*a*, *b*) than the perturbation equations (2.3*a*, *b*), but the solutions of (2.3*a*, *b*) give the expectation value of $\hat{A}$ for just one state, while the solutions of (2.10*a*, *b*) give the same for any state.

It is interesting to investigate the conditions under which solutions of (2.10*a*, *b*) exist. If $|\phi_n\rangle$ is an eigenvector of $\hat{H}_0$ belonging to eigenvalue ϵ_n, then

$$[\hat{H}_0, \hat{U}_i] = [(\hat{H}_0 - \epsilon_n \hat{I}), \hat{U}_i] \tag{2.11}$$

implies that

$$\langle \phi_n | [(\hat{H}_0 - \epsilon_n \hat{I}), \hat{U}_i] | \phi_n \rangle = \langle \phi_n | [\hat{H}_0, \hat{U}_i] | \phi_n \rangle = 0, \tag{2.12a}$$

which for $i = 1$ reduces to

$$\langle \phi_n | \hat{A} | \phi_n \rangle = 0. \tag{2.12b}$$

Condition (2.12) is rather restrictive, and there might not be many observables with operators corresponding to them satisfying this requirement. However, if $\hat{A}$ is any self-adjoint operator and if $\hat{H}_0$ has a complete set $\{|\phi_n\rangle\}$ of eigenvectors belonging to eigenvalues $\{\epsilon_n\}$, then

$$\hat{A}' = \hat{A} - \sum_{i=1}^{\infty} \langle \phi_i | \hat{A} | \phi_i \rangle |\phi_i\rangle\langle\phi_i| \tag{2.13}$$

clearly satisfies requirement (2.12*b*). In cases where the set $\{|\phi_n\rangle\}$ is known, $\hat{A}'$, in principle, is completely determined in terms of $\hat{A}$ and known operators. It is easy to see that the Hermitian part of $\hat{U}_i$, that is, $\frac{1}{2}(\hat{U}_i + \hat{U}_i^*)$, commutes with $\hat{H}_0$, and that each $\hat{U}_i$ is undetermined by an additive multiple of any operator which commutes with $\hat{H}_0$. This indeterminacy can be used to require that $\hat{U}_i$ is anti-Hermitian as the essential part of $\hat{U}_i$ is indeed so.

3. Applications in the first order: the interchange theorem

Suppose $|\psi_0\rangle$ is the nth eigenvector $|\phi_n\rangle$ of $\hat{H}_0$. In order to calculate $\langle \hat{A} \rangle_1$, we require $\langle \phi_n | [\hat{U}_1, \hat{H}_1] | \phi_n \rangle$ only, and all that is required, assuming $|\phi_n\rangle$ is non-degenerate, is the restriction of $[\hat{U}_1, \hat{H}_1]$ to the one-dimensional subspace spanned by $|\phi_n\rangle$. In this subspace

$$[\hat{H}_0, \hat{U}_1] = \hat{A} - \sum_{i=1}^{\infty} \langle \phi_i | \hat{A} | \phi_i \rangle |\phi_i\rangle\langle\phi_i| \tag{3.1}$$

takes the simple form

$$[\hat{H}_0, \hat{U}_1]|\phi_n\rangle = \hat{A}|\phi_n\rangle - \langle\phi_n|\hat{A}|\phi_n\rangle|\phi_n\rangle, \tag{3.2}$$

which, of course, satisfies

$$\langle\phi_n|[\hat{H}_0, \hat{U}_1]|\phi_n\rangle = \langle\phi_n|\hat{A}|\phi_n\rangle - \langle\phi_n|\hat{A}|\phi_n\rangle = 0. \tag{3.3}$$

In this one-dimensional subspace $\hat{U}_1$ is anti-Hermitian, which for real $|\phi_n\rangle$ implies that

$$\langle\phi_n|\hat{U}_1|\phi_n\rangle = 0, \tag{3.4}$$

and since 0 is real, $\hat{U}_1$ in this subspace is also Hermitian. This explains why Killingbeck (1972), seeking Hermitian solutions of (3.2) in a version of the theory which started and ended at the first order, was able to obtain correct results.

Since

$$[\hat{H}_0, \hat{U}_1] = [\hat{H}_0 - \epsilon_n\hat{I}, U_1], \tag{3.5}$$

(3.4) reduces to

$$(\hat{H}_0 - \epsilon_n\hat{I})\hat{U}_1|\phi_n\rangle = \hat{A}|\phi_n\rangle - \langle\phi_n|\hat{A}|\phi_n\rangle|\phi_n\rangle. \tag{3.6}$$

In other words, $-\hat{U}_1|\phi_n\rangle$ is the first-order correction $|\phi'_n(\hat{A})\rangle$ to $|\phi_n\rangle$ if $\hat{H}_0$ is perturbed by the hypothetical potential $\hat{A}$, and remembering that $\hat{U}_1$ is anti-Hermitian, one gets

$$\langle\hat{A}'\rangle_1 = \langle\phi_n|\hat{U}_1\hat{H}_1 - \hat{H}_1\hat{U}_1|\phi_n\rangle = \langle\phi_n|\hat{H}_1|\phi'_n(\hat{A})\rangle + \langle\phi'_n(\hat{A})|\hat{H}_1|\phi_n\rangle. \tag{3.7a}$$

But, since $\psi_1 = |\phi'_n(\hat{H}_1)\rangle$, one also gets

$$\begin{aligned}\langle\hat{A}'\rangle_1 &= \langle\phi'_n(\hat{H}_1)|\hat{A}'|\phi_n\rangle + \langle\phi_n|\hat{A}'|\phi'_n(\hat{H}_1)\rangle \\ &= \langle\phi'_n(\hat{H}_1)|\hat{A}|\phi_n\rangle + \langle\phi_n|\hat{A}|\phi'_n(\hat{H}_1)\rangle - \langle\phi_n|\hat{A}|\phi_n\rangle(\langle\phi'_n(\hat{H}_1)|\phi_n\rangle + \langle\phi_n|\phi'_n(\hat{H}_1)\rangle) \\ &= \langle\phi'_n(\hat{H}_1)|\hat{A}|\phi_n\rangle + \langle\phi_n|\hat{A}|\phi'_n(\hat{H}_1)\rangle \\ &= \langle A\rangle_1,\end{aligned} \tag{3.7}$$

where we have used

$$\langle\phi'_n(\hat{H}_1)|\phi_n\rangle + \langle\phi_n|\phi'_n(\hat{H}_1)\rangle = 0, \tag{3.8}$$

which follows from the normalisation of the perturbed state vector (Sharma 1968). Thus (3.7) not only shows that $\hat{A}'$ and $\hat{A}$ have the same first-order term, but the interchange theorem is contained in it. Here we have an alternative, almost inadvertent, derivation of this well-known theorem. It is probable that Killingbeck (1972) could not see the equivalence of his result to the interchange theorem because he was requiring his $\hat{U}_1$ to be Hermitian.

4. Explicit solution of (3.6) for a certain class of multiplicative operators for hydrogenic states in atoms

In atomic problems which can be reduced to the hydrogenic form, one works in the configuration space, where

$$\hat{H}_0 = -\tfrac{1}{2}\nabla^2 - 1/r, \qquad \phi_n = R(r)\Theta(\theta)\Phi(\phi). \tag{4.1}$$

Assuming $\hat{A}'$ to be a multiplicative operator having the form $f(r)$, one looks for an

operator $\hat{U}_1$ which one assumes to be a multiplicative operator. Then

$$[\hat{H}_0, \hat{U}_1]\phi_n = -(\Theta\Phi/2)[(2/r)U_1'R + 2U_1'R' + U_1''R]. \tag{4.2}$$

Hence $\hat{U}_1$ is the solution of the differential equation

$$(2/r)U_1'R + 2U_1'R' + U_1'' = -2f(r)R(r) \tag{4.3}$$

or

$$U_1'' + 2(R'/R + 1/r)U_1' = -2f(r), \tag{4.3a}$$

where, while treating $\hat{U}_1$ as a multiplicative operator, which is just a function of r, we have dropped the circumflex over it.

This differential equation has an integrating factor

$$\exp\left(2\int (R'/R + 1/r)\,\mathrm{d}r\right) = \exp[2\ln(rR)] = (rR)^2. \tag{4.4}$$

Hence

$$r^2R^2U_1' = c_1 - 2\int_0^r f(s)s^2R^2(s)\,\mathrm{d}s \tag{4.5}$$

and

$$U_1' = \frac{c_1}{r^2R^2} - \frac{2}{r^2R^2}\int_0^r f(s)s^2R^2(s)\,\mathrm{d}s, \tag{4.6}$$

and, integrating again,

$$U_1 = c_2 + c_1\int_0^r \frac{\mathrm{d}s}{s^2R^2(s)} - 2\int_0^r \frac{\mathrm{d}s}{s^2R^2(s)}\int_0^s f(t)t^2R^2(t)\,\mathrm{d}t. \tag{4.7}$$

Since the integral in the second term acting on $R(r)$ clearly gives a non-normalisable function, one must take $c_1 = 0$, giving

$$U_1 = c_2 - 2\int_0^r \frac{\mathrm{d}s}{s^2R^2(s)}\int_0^s f(t)t^2R^2(t)\,\mathrm{d}t. \tag{4.8}$$

One determines c_2 by the requirement that U_1 is anti-Hermitian in the one-dimensional subspace spanned by the real function ϕ_n, that is

$$\int_0^\infty U_1(r)R^2(r)r^2\,\mathrm{d}r = 0, \tag{4.9}$$

giving

$$c_2 = \left(2\Big/\int_0^\infty R^2(r)r^2\,\mathrm{d}r\right)\int_0^\infty R^2(r)r^2\,\mathrm{d}r\int_0^r \frac{\mathrm{d}s}{s^2R^2(s)}\int_0^s f(t)t^2R^2(t)\,\mathrm{d}t. \tag{4.10}$$

Killingbeck (1972), starting with simple forms of U_1, evaluated the corresponding $f(r)$, whereas the problem really is to find U_1 for a given $f(r)$. Killingbeck's indirect method gave the solution for a few $f(r)$; the method of this paper is completely general, and Killingbeck's few solutions can easily be seen to be particular examples of the integral in (4.8).

5. Concluding remarks

It has been seen that the theory developed in this paper, in spite of the rather rigid restriction on $\hat{A}$, is useful for the determination of the expectation value of any observable, at least in the first order, where it is equivalent to the interchange theorem. Since the second-order term in the expansion of $\hat{A}$ is the first-order term in the expansion of $[\hat{U}_1, \hat{H}_1]$, and $\hat{U}_1$ has already been obtained, it may seem at first sight that by repeating the kind of work done for the first-order case one can calculate $\langle \hat{A} \rangle_n$ for any n. Unfortunately $\hat{U}_1$ obtained in § 4 is a local solution valid only in the one-dimensional subspace spanned by the particular eigenvector, and this solution cannot be assumed to hold globally. This point is illustrated by taking $R(r)$ to be $r^l \, \mathrm{e}^{-\xi r}$. For

$$f(r) = \xi - (l+1)/r, \tag{5.1}$$

equation (4.8) yields (see also Killingbeck 1972), if we take $c_2 = 0$ (from the point of view of a commutator an additive constant is inconsequent),

$$U_1 = r. \tag{5.2}$$

Now the radial part of $-\nabla^2/2$ for hydrogenic states with azimuthal quantum number l has the form

$$-\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} + \frac{l(l+1)}{2r^2}. \tag{5.3}$$

Therefore

$$[\hat{H}_0, \hat{U}_1] = \left[-\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} + \frac{l(l+1)}{2r^2}, r\right] = -\frac{\mathrm{d}}{\mathrm{d}r} - \frac{1}{r}. \tag{5.4}$$

The restrictions of $-\mathrm{d}/\mathrm{d}r - 1/r$ and $\xi - (l+1)/r$ to the one-dimensional subspace spanned by $r^l \, \mathrm{e}^{-\xi r}$ are identical, but globally they are very different operators. In order to extend the theory to higher orders, one needs not only the restriction of $\hat{U}_1$ to the particular one-dimensional subspace, but also its form globally. Hence considerable further work is required before one can use this theory in higher-order calculations.

Before concluding, it should be pointed out that Kauffman (1972) has found a relation between properties of $\hat{A}$, $\hat{H}$ and $[\hat{A}, \hat{H}]$ for closed operators $\hat{A}$, Lavine (1971, 1972) has studied applications of commutators to scattering theory, and Coulson (1965) has studied the use of commutators in solving Schrödinger's equation.

References

Coulson C A 1965 *Q. J. Math., Oxford* (2) **16** 279–87
Dalgarno A and Lewis J T 1955 *Proc. R. Soc.* A **233** 70–4
Kauffman R M 1972 *Proc. Am. Math. Soc.* **33** 95–100
Killingbeck J 1972 *J. Phys. B: Atom. Molec. Phys.* **5** 1655–61
Lavine R B 1971 *Commun. Math. Phys.* **20** 301–23
—— 1972 *Indiana Univ. Math. J.* **21** 643–56
Sharma C S 1968 *Proc. R. Soc.* A **304** 513–30