AIP | The Journal of Chemical Physics

A Practical Method for the Solution of Certain Problems in Quantum Mechanics by Successive Removal of Terms from the Hamiltonian by Contact Transformations of the Dynamical Variables Part I. General Theory

L. H. Thomas

Citation: The Journal of Chemical Physics **10**, 532 (1942); doi: 10.1063/1.1723760
View online: http://dx.doi.org/10.1063/1.1723760
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/10/8?ver=pdfcov
Published by the AIP Publishing

A Practical Method for the Solution of Certain Problems in Quantum Mechanics by Successive Removal of Terms from the Hamiltonian by Contact Transformations of the Dynamical Variables

Part I. General Theory

L. H. THOMAS
Mendenhall Laboratory of Physics, The Ohio State University, Columbus, Ohio

(Received December 18, 1941)

Problems in quantum mechanics may be solved by canonical change of the representation in terms of which the dynamical variables are expressed. They may also be solved by contact transformations of the dynamical variables. The usual notation is modified slightly to make it more convenient for developing these two methods. Their parallel development shows their essential equivalence but formal difference.

INTRODUCTION

THE motion of a molecule is complicated and requires for its description a number of dynamical variables which may be chosen in various ways.[1] If we choose variables, such as the normal coordinates for small oscillations, which vary with the time nearly independently, a zero-order approximation to the motion is obtained by neglecting entirely the interaction between these variables.[2] If the electric moment and polarizability of the molecule are expressed in terms of them, the absorbtion, emission, and scattering of light may be calculated.[3] The gross structure of the spectrum of the molecule can be explained in this approximation, but to explain the details, the interactions must be taken into account. This is conveniently done by perturbation theory or successive approximation, going to the first order, second order, or even to higher orders, in the small interactions.[4] This more accurate work is necessary also if we wish to obtain the forces between the atomic nuclei in the molecule from the observed spectrum with even moderate precision.[5]

In the customary method perturbation theory is applied to the representation of the motion in terms of the stationary states of the zero-order approximation.[6] When there are many variables this becomes cumbersome, and is liable to error because of the large number of separate terms to be considered. We have found in a number of problems that the work is considerably abbreviated by applying transformations directly to the normal coordinates; and this liability to error is reduced.[7] While this is one of the earliest methods of quantum mechanics and is explained in many textbooks, it was found less useful than the customary method in dealing with atomic spectra and it has been little used.[8] The transformations are parallel to the contact transformations used in dynamical astronomy to remove terms from the Hamiltonian function.

1. NOTATION

The usual notations are a little clumsy for dealing with contact transformations, and we modify them slightly as follows:

A symbol in square brackets as

$$[s] \tag{1.01}$$

[1] H. Kronig, *Band Spectra and Molecular Structure* (Cambridge, 1930); H. Sponer, *Molekulspectren II* (Springer, 1936); G. Herzberg, *Molecular Spectra and Molecular Structure I*, (Prentice Hall, 1939); D. M. Dennison, Rev. Mod. Phys. **3**, 280 (1931); **12**, 175 (1940).

[2] H. Kronig, reference 1, p. 31; H. Sponer, reference 1, p. 42.

[3] H. Kronig, reference 1, Chapter 3; G. Herzberg, reference 1, Chapter 3.

[4] Born, Heisenberg, and Jordan, Zeits. f. Physik **35**, 557 (1925).

[5] D. M. Dennison, reference 1, p. 177.

[6] A. Ruark and H. C. Urey, *Atoms, Molecules and Quanta* (McGraw-Hill, 1930), p. 597; P. A. M. Dirac, *Quantum Mechanics*, second edition, (Oxford, 1934), Chapter 8; E. C. Kemble, *Fundamental Principles of Quantum Mechanics* (McGraw-Hill, 1937), Chapter 11.

[7] Shaffer, Nielsen, and Thomas, Phys. Rev. **56**, 895 (1939); W. H. Shaffer, J. Chem. Phys. **9**, 607 (1941); S. Silver and W. H. Shaffer, J. Chem. Phys. **9**, 599 (1941).

[8] A. Ruark and H. C. Urey, reference 6, p. 595; P. A. M. Dirac, reference 6, p. 110; E. C. Kemble, reference 6, p. 373.

will stand for a matrix with elements

$$(m|s|u). \tag{1.02}$$

In this symbol m and u are variables or sets of variables designating the row and column of the matrix; they may be given by several variables taking integral values, or continuous ranges of values, or both. In general there need be no correlation implied between the values of m and the values of u, but most usually u is given by corresponding variables to those giving m taking the same values, and we then conveniently denote them by m and n or by n and n'.

The physical systems in which we are interested are described in terms of dynamical variables, ξ, η, $\cdots$, which combine according to a non-commutative algebra we suppose known, given by kinematical conditions which usually comprise commutation relations: e.g., a coordinate q and its canonically conjugate momentum p satisfying

$$qp-pq=i\hbar \tag{1.10}$$

(h, as usual will stand for Planck's constant, 6.62×10^{-27} erg-sec.; $\hbar$ stands for $h/2\pi$).

The dynamical variables can be represented accordingly in various ways by Hermitian matrices,

$$[a|\xi|a],\ [a|\eta|a],\ \cdots, \tag{1.11}$$

with elements

$$(a_m|\xi|a_n),\ (a_m|\eta|a_n),\ \cdots. \tag{1.12}$$

In these symbols m and n are the variables or sets of variables designating the row and column of the matrix; a designates the representation; and ξ, η, $\cdots$, designate the dynamical variable which the matrix represents. Since $[a|\xi|a]$ is Hermitian, $(a_m|\xi|a_n)$ and $(a_n|\xi|a_m)$ are conjugate complex numbers.

E.g., if $q\surd(2\mu\omega/\hbar)$ and $p\surd(2/\mu\omega\hbar)$ are represented by the matrices

$$\begin{matrix} 0 & i & 0 & 0 & 0 & \cdot\cdot \\ -i & 0 & i\surd2 & 0 & 0 & \cdot\cdot \\ 0 & -i\surd2 & 0 & i\surd3 & 0 & \cdot\cdot \\ 0 & 0 & -i\surd3 & 0 & i\surd4 & \cdot\cdot \\ 0 & 0 & 0 & -i\surd4 & 0 & \cdot\cdot \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \end{matrix}$$

and

$$\begin{matrix} 0 & 1 & 0 & 0 & 0 & \cdot\cdot \\ 1 & 0 & \surd2 & 0 & 0 & \cdot\cdot \\ 0 & \surd2 & 0 & \surd3 & 0 & \cdot\cdot \\ 0 & 0 & \surd3 & 0 & \surd4 & \cdot\cdot \\ 0 & 0 & 0 & \surd4 & 0 & \cdot\cdot \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \end{matrix}$$

and if we call the representation E, then $[E|q\surd(2\mu\omega/\hbar)|E]$ and $[E|p\surd(2/\mu\omega\hbar)|E]$ are these matrices, and if we number the rows from the top by $m=0, 1, 2, \cdots$, and the columns from the left by $n=0, 1, 2, \cdots$,

$$\begin{aligned}(E_m|q\surd(2\mu\omega/\hbar)|E_n) &= i\surd(\tfrac{1}{2}m) \qquad n=m+1 \\ &\qquad\qquad\qquad m, n=0, 1, 2, \cdots, \\ &= -i\surd(\tfrac{1}{2}m) \quad m=n+1 \\ &= 0 \qquad\qquad\quad \text{otherwise.}\end{aligned}$$

m may be given by several variables taking integral values, or continuous ranges of values, or both; but n must be given by corresponding variables taking the same values; and the summation or integration giving the product of two matrices in the representation, which corresponds to the product of the dynamical variables in the algebra, must be specified or implied. We write

$$[a|\xi\eta|a]=[a|\xi|a][a|\eta|a] \tag{1.21}$$

and

$$(a_m|\xi\eta|a_o)=\textstyle\sum_n (a_m|\xi|a_n)(a_n|\eta|a_o), \tag{1.22}$$

with the understanding that $\sum_n$ is to be interpreted as required. We abbreviate the unit matrix $[a|1|a]$ to $[a|a]$. When there is no material ambiguity as to what representation is implied, $(a_m|\xi|a_n)$ may be abbreviated to $(m|\xi|n)$, and since m and n take the same values it is often convenient to replace them by n and n', especially when they are given explicitly by sets of variables; e.g., $(j, m|L|j', m')$. When convenient we designate the representation by the symbol of a dynamical variable or by those of a set of commuting dynamical variables, α, β, $\cdots$, say, represented by diagonal matrices for which the elements vanish when $n\neq m$. We may then use the diagonal elements, α', α'', $\cdots$; β', β'', $\cdots$; $\cdots$; of these matrices, the characteristic values of the dynamical variables, to designate the rows and columns of the representation, and write

$$[\alpha, \beta, \cdots|\xi|\alpha, \beta, \cdots]$$

for the matrix representing ξ in this representation, and

$$(\alpha', \beta', \cdots | \xi | \alpha'', \beta'', \cdots)$$

for an arbitrary element of that matrix. The use of quantum numbers combines these notations, dropping the primes when there is no ambiguity. E.g., if q and p are represented as above, and if

$$E = \tfrac{1}{2}(p^2/\mu + \mu\omega^2 q^2)$$

$E/\hbar\omega$ is represented by the matrix

$$\begin{matrix} \tfrac{1}{2} & 0 & 0 & 0 & \cdot\cdot \\ 0 & \tfrac{3}{2} & 0 & 0 & \cdot\cdot \\ 0 & 0 & \tfrac{5}{2} & 0 & \cdot\cdot \\ 0 & 0 & 0 & \tfrac{7}{2} & \cdot\cdot \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \end{matrix}$$

and if we put $E = (n + \tfrac{1}{2})\hbar\omega$, we may write

$$(n-1 | q\sqrt{(2\mu\omega/\hbar)} | n) = i\sqrt{(n/2)}, \quad n = 0, 1, 2, \cdots$$
$$(n | q\sqrt{(2\mu\omega/\hbar)} | n-1) = -i\sqrt{(n/2)},$$

other elements vanishing (though the phases of the elements are not determined by

$$\tfrac{1}{2}(p^2/\mu + \mu\omega^2 q^2) = (n + \tfrac{1}{2})\hbar\omega).$$

If we have two representations, a and a^*, for which the variables designating the rows and columns take the same values, and if dynamical variables ξ, η, are represented in repesentation a by the matrices $[a|\xi|a]$, $[a|\eta|a]$, $\cdots$, we may define (different) dynamical variables ξ^*, η^*, $\cdots$, represented by the same matrices in the representation a^*, so that

$$\begin{aligned} (a_m^* | \xi^* | a_n^*) &= (a_m | \xi | a_n), \\ (a_m^* | \eta^* | a_n^*) &= (a_m | \eta | a_n) \\ \cdots \quad & \quad \cdots \end{aligned} \tag{1.32}$$

or

$$\begin{aligned} [a^* | \xi^* | a^*] &= [a | \xi | a], \\ [a^* | \eta^* | a^*] &= [a | \eta | a] \\ \cdots \quad & \quad \cdots . \end{aligned} \tag{1.31}$$

The dynamical variables ξ^*, η^*, $\cdots$, then satisfy kinematical conditions of the same form as those satisfied by ξ, η, $\cdots$. Conversely, if we have two sets of dynamical variables ξ, η, $\cdots$, and ξ^*, η^*, $\cdots$, satisfying kinematical conditions of the same form, and if ξ, η, $\cdots$, are represented in a representation a by the matrices $[a|\xi|a]$, $[a|\eta|a]$, $\cdots$, we can define a representation a^* in which ξ^*, η^*, $\cdots$, are represented by the same matrices, so that

$$\begin{aligned} (a_m^* | \xi^* | a_n^*) &= (a_m | \xi | a_n) \\ (a_m^* | \eta^* | a_n^*) &= (a_m | \eta | a_n), \\ \cdots \quad & \quad \cdots \end{aligned} \tag{1.32}$$

or

$$\begin{aligned} [a^* | \xi^* | a^*] &= [a | \xi | a], \\ [a^* | \eta^* | a^*] &= [a | \eta | a] \\ \cdots \quad & \quad \cdots . \end{aligned} \tag{1.31}$$

Such a symbol as $F(\xi, \eta, \cdots)$ will designate a function of the dynamical variables ξ, η, $\cdots$; $F(\xi, \eta, \cdots)$ will be itself a dynamical variable and may be defined in general by a matrix representing it in some representation of the dynamical variables ξ, η, $\cdots$, which matrix must be invariant for any change of representation that leaves the matrices representing the dynamical variables ξ, η, $\cdots$, invariant; the most usual functions can be regarded as convergent power series in the dynamical variables, the order of factors in the terms being, however, important.

If the dynamical variables ξ^*, η^*, $\cdots$, satisfy kinematical conditions of the same form as those satisfied by ξ, η, $\cdots$; and if $F(\xi, \eta, \cdots)$ denotes a function of the dynamical variables ξ, η, $\cdots$; then $F(\xi^*, \eta^*, \cdots)$ will denote the same function of the dynamical variables ξ^*, η^*, $\cdots$; e.g., if

$$F(\xi, \eta, \cdots) = a\xi + b\eta + \cdots,$$
$$F(\xi^*, \eta^*, \cdots) = a\xi^* + b\eta^* + \cdots.$$

If we require the function of ξ^*, η^*, $\cdots$, that is equal to $F(\xi, \eta, \cdots)$, we shall denote it by $F^*(\xi^*, \eta^*, \cdots)$ so that we shall have

$$F^*(\xi^*, \eta^*, \cdots) = F(\xi, \eta, \cdots) \tag{1.4}$$

and these may represent the same physical quantity, such as energy, electric moment, etc. E.g.,

$$\text{if } \left.\begin{aligned} q &= q^* \cos\epsilon + p^* \sin\epsilon \\ p &= -q^* \sin\epsilon + p^* \cos\epsilon \end{aligned}\right\} \text{(so that } qp - pq = i\hbar$$

and $q^*p^* - p^*q^* = i\hbar$ follow from each other), and if $F(q, p) = a(q^2 - p^2) + b(qp + pq)$, then

$$\begin{aligned} F(q, p) = F^*(q^*, p^*) = (a \cos 2\epsilon - b \sin 2\epsilon)(q^{*2} + p^{*2}) \\ + (a \sin 2\epsilon + b \cos 2\epsilon)(q^*p^* + p^*q^*). \end{aligned}$$

These notations are believed to be in current use;[9] except that it is not usual to write $[a|\xi|a]$ for the matrix representing the dynamical vari-

[9] P. A. M. Dirac, reference 6, Chapter 3, and p. 109; E. C. Kemble, reference 6, p. 268.

able ξ in the representation a, distinguished on the one hand from its elements and on the other from matrices representing ξ in other representations; and that the idea of a function of dynamical variables is often confined to a function of a set of commuting dynamical variables.

2. CHANGE OF REPRESENTATION AND TRANSFORMATION OF DYNAMICAL VARIABLES[10]

We change from one representation a with rows and columns denoted by m, n, to another representation b with rows and columns denoted by u, v, by using transformation matrices $[b|a]$, $[a|b]$, such that for any dynamical variable ξ,

$$\begin{aligned}[b|\xi|b]&=[b|a][a|\xi|a][a|b],\\ [a|\xi|a]&=[a|b][b|\xi|b][b|a].\end{aligned} \tag{2.01}$$

$[b|a)$ and $[a|b]$ must be such that $(b_u|a_m)$ and $(a_m|b_u)$ are complex conjugates and that

$$\begin{aligned}[b|a][a|b]&=[b|b],\\ [a|b][b|a]&=[a|a].\end{aligned} \tag{2.02}$$

In general there need be no correlation implied between the values of m or n and the values of u or v. If a one-to-one correspondence is given between these values so that the methods of summation or integration correspond, we may conveniently denote them by the same symbol m or n, and denote the representations by a and a^*; and we shall call the change of representation a canonical change of representation. (This specialization is not usually made explicitly.) The transformation matrix is then unitary and can be expressed in the form

$$[a|a^*]=e^{i[s]} \tag{2.03}$$

so that

$$[a^*|a]=e^{-i[s]}$$

where $[s]$ is a Hermitian matrix with elements $(m|s|n)$ such that $(m|s|n)$ and $(n|s|m)$ are complex conjugates.

We shall regard

$$[a^*|\xi|a^*]=e^{-i[s]}[a|\xi|a]e^{i[s]} \tag{2.04}$$

or

$$[a|\xi|a]=e^{i[s]}[a^*|\xi|a^*]e^{-i[s]}$$

as the standard form for a canonical change of representation.

We transform from one set of dynamical variables ξ, η, $\cdots$, to another set α, β, $\cdots$, by defining each set as functions of the other. In general the new set will satisfy kinematical conditions of different forms from those satisfied by the old. If they satisfy kinematical conditions of the same form, we may conveniently denote them by ξ^*, η^*, $\cdots$; and if ξ, η, $\cdots$ are represented by matrices in a representation a, then ξ^*, η^*, $\cdots$ will be represented by the same matrices in another representation a^*. The change from a to a^* is canonical and

$$\begin{aligned}[a|\xi^*|a]&=[a|a^*][a^*|\xi^*|a^*][a^*|a]\\ &=[a|a^*][a|\xi|a][a^*|a].\end{aligned} \tag{2.05}$$

Thus if U is a unitary function of ξ, η, $\cdots$, which has matrix elements in the representation a given by[11]

$$(a_m|U|a_n)=(a_m|a_n^*),$$

so that

$$[a|U|a]=[a|a^*], \tag{2.06}$$

we have

$$\begin{aligned}\xi^*&=U\xi U^{-1},\\ \eta^*&=U\eta U^{-1}\end{aligned} \tag{2.071}$$

and indeed, for any function $F(\xi, \eta, \cdots)$,

$$F(\xi^*, \eta^*, \cdots)=UF(\xi, \eta, \cdots)U^{-1} \tag{2.072}$$

(contrast this with

$$F^*(\xi^*, \eta^*, \cdots)=F(\xi, \eta, \cdots) \tag{1.4}$$

by definition). $(a_n|U^{-1}|a_m)$ and $(a_m|U|a_n)$ are complex conjugates and $UU^{-1}=U^{-1}U=1$. Such a transformation from dynamical variables ξ, η, $\cdots$, to dynamical variables ξ^*, η^*, $\cdots$, satisfying the same kinematical conditions is called a contact transformation of the dynamical variables.

U being unitary, it can be expressed in the form

$$U=e^{iS/\hbar}, \tag{2.08}$$

where S is a real function of the dynamical variables (i.e., the matrices representing it are Hermitian). If

$$[a|a^*]=e^{i[s]}, \tag{2.03}$$

then

$$(a_m|S|a_n)=\hbar(m|s|n)$$

or

$$[a|S|a]=\hbar[s]. \tag{2.09}$$

[10] P. A. M. Dirac, reference 6, Chapter 3; E. C. Kemble, reference 6, p. 394.

[11] P. A. M. Dirac, reference 6, p. 109.

We shall write

$$S=S(\xi,\eta,\cdots)=S^*(\xi^*,\eta^*,\cdots) \quad (2.10)$$

and shall regard the form

$$\xi^*=e^{iS/\hbar}\xi e^{-iS/\hbar},\quad \eta^*=e^{iS/\hbar}\eta e^{-iS/\hbar},\quad \cdots, \quad (2.11)$$

or

$$\xi=e^{-iS/\hbar}\xi^* e^{iS/\hbar},\quad \eta=e^{-iS/\hbar}\eta^* e^{iS/\hbar},\quad \cdots,$$

as the standard form for a contact transformation. A function of the original dynamical variables ξ, η, $\cdots$, is then transformed as follows:

$$F^*(\xi^*,\eta^*,\cdots)=F(\xi,\eta,\cdots)$$
$$=e^{-iS/\hbar}F(\xi^*,\eta^*,\cdots)e^{iS/\hbar}. \quad (2.12)$$

E.g., if $qp-pq=i\hbar$, $S=cp$ gives

$$q^*=e^{icp/\hbar}qe^{-icp/\hbar}=q+c,$$
$$p^*=e^{icp/\hbar}pe^{-icp/\hbar}=p$$

and $F^*(p^*,q^*)=F(p,q)=F(p^*,q^*-c)$.[12]

$$[a^*|F|a^*]=e^{-i[s]}[a|F|a]e^{i[s]} \quad \text{(from (2.04))}$$

and

$$F^*(\xi^*,\eta^*,\cdots)=e^{-iS/\hbar}F(\xi^*,\eta^*,\cdots)e^{iS/\hbar} \quad (2.12)$$

are actually just different ways of writing the same result. The formula

$$e^{-(i/\hbar)K}Fe^{(i/\hbar)K}\equiv F+\frac{i}{\hbar}(FK-KF)$$
$$+\frac{1}{2}\left(\frac{i}{\hbar}\right)^2((FK-KF)K-K(FK-KF))+\cdots$$
$$\equiv F+[K,F]+\tfrac{1}{2}[K,[K,F]]+\cdots$$
$$+\frac{1}{r!}[K,[\cdots,[K,F]\cdots]]+\cdots, \quad (2.13)$$

where the Poisson bracket $[K,F]$ is defined by

$$[K,F]=-\frac{i}{\hbar}(KF-FK), \quad (2.14)$$

is useful for evaluating these expressions. If $F(\xi,\eta,\cdots)$ and $S^*(\xi^*,\eta^*,\cdots)$ are known as power series, this formula gives

$$F^*(\xi^*,\eta^*,\cdots)=F(\xi^*,\eta^*,\cdots)$$
$$+[S(\xi^*,\eta^*,\cdots),F(\xi^*,\eta^*,\cdots)]$$
$$+\tfrac{1}{2}[S(\xi^*,\eta^*,\cdots),[S(\xi^*,\eta^*,\cdots),$$
$$\times F(\xi^*,\eta^*,\cdots)]]+\cdots. \quad (2.15)$$

3. THE MECHANICAL PROBLEM. PERTURBATION THEORY

We suppose that the Hamiltonian function H of a system, and the various physical properties in which we are interested, such as electric moment $\mathbf{e}$ and polarizability Φ, are given in terms of dynamical variables the kinematical relations between which are known. Our object is to find matrices representing H, $\mathbf{e}$, Φ, etc., in a representation b in which H has only diagonal elements: to find values for u, v, and for E_u, such that

$$(b_u|H|b_v)=E_u \quad v=u \quad (3.1)$$
$$=0 \quad \text{otherwise},$$

and to find $(b_u|\mathbf{e}|b_v)$, $(b_u|\Phi|b_v)$, $\cdots$ etc. We divide this into a zero-order problem, and perturbation theory.[13]

The zero-order problem is to transform to dynamical variables ξ, η, $\cdots$, such as components of angular momentum and normal coordinates and momenta, in terms of which H has nearly a form H_0 for which we know such a representation, a. H is conveniently expressed as a power series in ϵ

$$H(\xi,\eta,\cdots)=H_0(\xi,\eta,\cdots)+\epsilon H_1(\xi,\eta,\cdots)$$
$$+\epsilon^2H_2(\xi,\eta,\cdots)+\cdots, \quad (3.2)$$

where ϵ is small (e.g., $\epsilon=h/\nu I$ where ν is a vibration frequency, I a moment of inertia), and where values for m, n, and E_{0m} are known, so

$$(a_m|H_0|a_n)=E_{0m} \quad m=n \quad (3.3)$$
$$=0 \quad \text{otherwise}.$$

We must also be able to find the matrix components in the representation a of functions of ξ, η, $\cdots$, in particular $(a_m|H_1|a_n)$, $(a_m|H_2|a_n)$, $\cdots$, $(a_m|\mathbf{e}|a_n)$, $(a_m|\Phi|a_n)$, $\cdots$. If we apply the canonical change of representation for which

$[a|a^*]=e^{i\epsilon[s]}$ to (from (2.03))

$$[a|H|a]=[a|H_0|a]+\epsilon[a|H_1|a]+\epsilon^2[a|H_2|a]+\cdots, \text{ (from (3.2))}$$

[12] P. A. M. Dirac, reference 6, p. 94.
[13] P. A. M. Dirac, reference 6, Chapter 8; E. C. Kemble, reference 6, p. 395; A. Ruark and H. C. Urey, reference 6, p. 596.

changing it to

$$[a^*|H|a^*]=[a^*|H_0^*|a^*]+\epsilon[a^*|H_1^*|a^*]+\epsilon^2[a^*|H_2^*|a^*]+\cdots$$

we obtain, expanding in powers of ϵ,[14]

$$[a^*|H_0^*|a^*]=[a|H_0|a], \tag{3.41}$$
$$[a^*|H_1^*|a^*]=[a|H_1|a]+i([a|H_0|a][s]-[s][a|H_0|a]), \tag{3.42}$$
$$\begin{aligned}[a^*|H_2^*|a^*]&=[a|H_2|a]+i([a|H_1|a][s]-[s][a|H_1|a])\\ &\quad+\tfrac{1}{2}i^2(([a|H_0|a][s]-[s][a|H_0|a])[s]-[s]([a|H_0|a][s]-[s][a|H_0|a]))\\ &=[a|H_2|a]+\tfrac{1}{2}i([a|H_1+H_1^*|a][s]-[s][a|H_1+H_1^*|a]),\end{aligned} \tag{3.43}$$
$$\cdots \quad \cdots \quad \cdots \quad \cdots \quad .^{15}$$

If we apply the equivalent contact transformation from ξ, η, $\cdots$, to ξ^*, η^*, $\cdots$, given by $U=e^{i\epsilon S/\hbar}$ to (from (2.08)

$$H(\xi,\eta,\cdots)=H_0(\xi,\eta,\cdots)+\epsilon H_1(\xi,\eta,\cdots)+\epsilon^2H_2(\xi,\eta,\cdots)+\cdots \text{ (from (3.2))}$$

transforming it to

$$H^*(\xi^*,\eta^*,\cdots)=F_0^*(\xi^*,\eta^*,\cdots)+\epsilon F_1^*(\xi^*,\eta^*,\cdots)+\epsilon^2F_2^*(\xi^*,\eta^*,\cdots)+\cdots$$ [16]

we obtain, expanding in powers of ϵ,

$$F_0^*(\xi^*,\eta^*,\cdots)=H_0(\xi^*,\eta^*,\cdots), \tag{3.51}$$
$$F_1^*(\xi^*,\eta^*,\cdots)=H_1(\xi^*,\eta^*,\cdots)+[S(\xi^*,\eta^*,\cdots),H_0(\xi^*,\eta^*,\cdots)], \tag{3.52}$$
$$\begin{aligned}F_2^*(\xi^*,\eta^*,\cdots)&=H_2(\xi^*,\eta^*,\cdots)+[S(\xi^*,\eta^*,\cdots),H_1(\xi^*,\eta^*,\cdots)]\\ &\quad+\tfrac{1}{2}[S(\xi^*,\eta^*,\cdots),[S(\xi^*,\eta^*,\cdots),H_0(\xi^*,\eta^*,\cdots)]]\\ &=H_2(\xi^*,\eta^*,\cdots)+\tfrac{1}{2}[S(\xi^*,\eta^*,\cdots),H_1(\xi^*,\eta^*,\cdots)+F_1^*(\xi^*,\eta^*,\cdots)].\end{aligned} \tag{3.53}$$
$$\cdots \quad \cdots \quad \cdots \quad \cdots$$

These results we shall write

$$H_0^*=H_0, \tag{3.51}$$
$$H_1^*=H_1+[S,H_0], \tag{3.52}$$
$$H_2^*=H_2+\tfrac{1}{2}[S,H_1+H_1^*], \tag{3.53}$$
$$H_3^*=H_3+[S,H_2]+\tfrac{1}{6}[S,[S,2H_2+H_2^*]], \tag{3.54}$$
$$\cdots \quad \cdots \quad \cdots \quad \cdots$$

We remove terms from H by taking suitable forms for s or S. If the zero-order approximation is not degenerate, i.e., has no coincident energy values, we remove first terms of order ϵ; then terms up to ϵ^3, then to ϵ^7, and so on; only terms that are functions of H_0 remain, i.e., that are diagonal in the representation a^*, a^{**}, $\cdots$, that we have reached. If the zero-order approximation is degenerate, and if H_1, H_2, $\cdots$ contain terms that remove the degeneracy, these terms also remain, and we must solve a zero-order perturbation problem, changing the representation or transforming the dynamical variables suitably, to approach the representation b.

If this process converged we should approach the desired representation b, and could apply the same transformations to **e** and Φ. In any case, if ϵ is small, the first few transformations must give results with errors of order ϵ^2, ϵ^4, ϵ^8, $\cdots$, in the elements of **e** and Φ as well as in E_m.

Successive canonical changes of representation are especially useful when we want results for only a few values of m or n; contact transformations where the functions can be expanded in power series in the dynamical variables. We shall give tables of forms for S which give various values of $[S, H_0]$ for standard forms of H_0, and of the resulting higher order terms that are introduced into H.

[14] Born, Heisenberg, and Jordon, Zeits. f. Physik **35**, 565 (1925).

[15] Often applied: e.g., J. H. Van Vleck, Phys. Rev. **33**, 484 (1929); O. M. Jordahl, Phys. Rev. **43**, 87 (1934).

[16] Note that

$$H_0^*(\xi^*,\eta^*,\cdots)=H_0(\xi^*,\eta^*,\cdots)+\epsilon[S(\xi^*,\eta^*,\cdots),H_0(\xi^*,\eta^*,\cdots)]+\cdots\neq F_0^*(\xi^*,\eta^*,\cdots).$$