The Rayleigh–Schrödinger perturbation and the linked-diagram theorem for a multi-configurational model space

Ingvar Lindgren

Department of Physics, Chalmers University of Technology, Göteborg, Sweden

Abstract. The Rayleigh–Schrödinger perturbation formalism is extended to the case of a model space, which is not necessarily degenerate. The model space defines the zero-order or model wavefunction, and the new formalism makes it possible to use a model wave-function of multi-configurational type. The effect of the states outside the model space are as usual taken into account by means of a perturbation expansion and expressed in terms of an 'effective' Hamiltonian, operating only within the model space. The extended Rayleigh–Schrödinger formalism is used to prove the linked-diagram theorem for a multi-configurational model space in a simple way. Alternatively, this problem can be handled by means of the well known formalism for degenerate perturbation, treating the splitting within the model space as due to an additional perturbation. The present approach, however, is more direct and the model space splitting is handled without summing any infinite series. The problem of convergence of the perturbation expansion is briefly discussed.

1. Introduction

The purpose of the perturbation technique in quantum mechanics is to solve the Schrödinger equation by means of successive approximations. The basic idea here is that the Hamiltonian H for the system is split up into a model Hamiltonian H_0 and a perturbation V,

$$H = H_0 + V.$$

The model Hamiltonian should represent a good approximation of the full Hamiltonian, and its eigenvalues and eigenfunctions should be easy to determine.

Two types of perturbation expansions are frequently used, namely the Brillouin–Wigner (BW) and the Rayleigh–Schrödinger (RS) schemes. The BW expansion is formally very simple, but it contains the unknown energy of the system, which means that the contributions due to the perturbation have to be calculated in a self-consistent way. The RS expansion, on the other hand, is given in terms of the known eigenvalues of the model Hamiltonian, but it has instead a number of additional terms not appearing in the BW case.

The starting point for the perturbation expansion is a zero-order or model wave-function, which is confined to a subspace of the Hilbert space for the system, usually called the model space. This subspace is defined by means of certain eigenfunctions of the model Hamiltonian, and the standard procedure is to restrict the model space to the space spanned by the eigenfunctions of a single eigenvalue of this Hamiltonian (single configuration). If the eigenvalue spectrum of the model Hamiltonian is non-degenerate, the model space will then contain a single state (closed-shell case), and it is easy to derive

the perturbation expansions. If the model space has degenerate eigenvalues (open-shell case), the perturbation theory becomes more complicated, particularly in the RS formulation. Complete derivations of the RS perturbation expansion, for a model space with a single, degenerate energy level (single open-shell configuration), were first given by Kato (1949, 1950) and by Bloch (1958). In their treatment the model function is restricted to a combination of degenerate eigenfunctions of the model Hamiltonian, and the admixture of eigenfunctions belonging to all other eigenvalues are treated by perturbation. Obviously, the convergence of the perturbation expansion can be expected to be faster, if a larger fraction of the final solution is included in the model function. In particular, it may happen that certain energy levels of the model Hamiltonian are so strongly interacting that it is difficult to treat their mixing on a perturbational basis. In the present work the RS perturbation is extended to a model space with several energy levels. This makes it possible to use a model function of the 'multi-configurational' type, so that strongly interacting levels can be included to infinite order (by solving a secular equation), and only weaker interactions have to be treated by perturbation.

The various contributions in the perturbation expansion are conveniently represented by means of diagrams of the Feynman type (Feynman 1949), and it was first shown by Brueckner (1955) that the additional terms in the RS expansion cancel in this representation in the lowest orders against so-called unlinked diagrams of the leading term. This means that the RS expansion is given by an expression analogous to the BW formula, provided only linked diagrams are considered. This is the *linked-diagram theorem*, which represents a great simplification in carrying out perturbation calculations to high order. This theorem was first proved by Goldstone (1957) in the non-degenerate case (closed-shell configuration). A more general proof has been given by Brandow (1967) by starting from the BW perturbation and expanding the energy dependence out of the denominators. Brandow was then able to show that the unlinked diagrams vanish in all orders, provided some 'folded' diagrams were added. A different proof of the linked-diagram theorem, based directly on the RS expansion for an exactly degenerate model space, has been given by Sandars (1969). The methods used by Brandow and Sandars are time-independent, while Goldstone used a time-dependent approach. Later, Goldstone's proof has been extended to degenerate systems (Johnson and Baranger 1971, Kuo *et al* 1971).

Modern applications of the perturbation technique are often based on the effective-operator formalism, which has been developed particularly in connection with nuclear problems (Eden and Francis 1955, Bloch and Horowitz 1958, Feshbach 1958, 1962, see also the review article by Barrett and Kirson 1973, where numerous references to recent works are to be found). The basic idea here is that the effect of the true operator, operating in the entire Hilbert space, can be reproduced by means of an 'effective' operator, operating only within the model space. Also the effective operators are conveniently represented by means of diagrams, and there is a corresponding linked-diagram theorem.

In the present work an extended form of the RS perturbation formalism is used to prove the linked-diagram theorem for a model space with several energy levels. A time-independent approach is used, and the work can be considered as an extension of that of Sandars (1969). The result is expressed in terms of a 'wave operator', which is the same for the entire model space and which can be used to evaluate wavefunctions and effective operators.

As mentioned, there are other ways of handling the problem of non-degeneracy within the model space. Brandow (1967) starts from a formally degenerate model space and

considers the splitting within this space as due to an additional perturbation. Under certain circumstances the effect of this perturbation can be summed to all orders. A different technique has been used by Balian and DeDominicis (1971). Like Brandow these authors start from the BW formalism, and they construct a closed form of the transformation to an energy-independent effective Hamiltonian. The present approach leads to essentially the same results, but it is more direct, since it starts from the energy-independent RS expansion. In particular, it does not depend on any summation of infinite series, which may have some consequences regarding the convergence of the perturbation expansions. This problem has recently been investigated in some detail by Weidenmuller *et al* (Schucan and Weidenmuller 1972, 1973, Hoffman *et al* 1973, 1974).

2. Perturbation expansions and effective operators

2.1. Projection and wave operators

Our basic problem is to solve the Schrödinger equation

$$H\Psi^\alpha = E^\alpha \Psi^\alpha. \tag{1}$$

H is the Hamilton operator for the system, and it is split up into two hermitian parts, a model Hamiltonian H_0 and a perturbation V:

$$H = H_0 + V. \tag{2}$$

We need not specify H_0 at this moment, but we assume that a complete set of orthonormal eigenfunctions $\{\phi^a\}$ and corresponding eigenvalues are available

$$H_0\phi^a = E_0^a\phi^a \qquad \langle \phi^a|\phi^b \rangle = \delta_{a,b}. \tag{3}$$

The eigenvalues E_0^a of the model Hamiltonian may be degenerate. (If the energy spectrum is partly continuous, it is for simplicity assume that some 'box of normalization' is used, so that all states can be treated formally in the same way.)

Some of the basis functions $\{\phi^a\}$ will define a model space D, and the remaining part of the Hilbert space is called the orthogonal space. It is assumed that all eigenstates of H_0 belonging to the same eigenvalue are in the same subspace. The projection operators associated with the two spaces are

$$P = \sum_{a \in D} |\phi^a\rangle\langle\phi^a| = \sum_{a \in D} |a\rangle\langle a|$$
$$Q = 1 - P = \sum_{r \notin D} |\phi^r\rangle\langle\phi^r| = \sum_{r \notin D} |r\rangle\langle r|. \tag{4}$$

The projection operators satisfy a number of well known relations, such as

$$P = P^+ = PP \qquad PQ = QP = 0$$
$$[P, H_0] = [Q, H_0] = 0. \tag{5}$$

If the model space has d dimensions, it can be shown that there is normally a one-to-one correspondence between d eigenfunctions of the full Hamiltonian Ψ^α and their projections onto the model space or model functions

$$\Psi_0^\alpha = P\Psi^\alpha \qquad \alpha = 1, 2, \ldots, d \tag{6}$$

(Kuo *et al* 1971). This can be made very plausible, if we consider the limiting case, when

the perturbation is turned off adiabatically. d eigenfunctions of H will then go over into distinct (and orthogonal) functions in the model space (and the remaining eigenfunctions into the orthogonal space). With no perturbation the model functions are, of course, identical to the corresponding true wavefunctions. When the perturbation is turned on again, the model functions can change, due to mixing within the model space. But the model functions are still expected to be distinct (except possibly in very accidental cases), and, therefore, there is a one-to-one correspondence between them and the original eigenfunctions of the full Hamiltonian. This makes it possible to define an operator Ω which transforms all d model functions back into these eigenfunctions

$$\Psi^\alpha = \Omega \Psi_0^\alpha \qquad \alpha = 1, 2, \ldots, d. \tag{7a}$$

This operator is often called the wave operator (Møller 1945, 1946) or the model operator (Eden and Francis 1955), and it should be observed that it is the same for *all d* states. The wave operator is analogous to the time-development operator $U(0, -\infty)$ used in the time-dependent approach. In order to specify Ω completely, we assume that it gives a null result when operating on the orthogonal space, ie

$$\Omega Q = 0. \tag{7b}$$

Since $P + Q = 1$, it follows directly from (7b) that

$$\Omega P = \Omega. \tag{8a}$$

Operating on (7a) with P from the left gives

$$\Psi_0^\alpha = P\Omega \Psi_0^\alpha \qquad \alpha = 1, 2, \ldots, d$$

which means that $P\Omega$ leaves any function in the model space unchanged. But since Ω destroys any component in the orthogonal space according to (7b), we always have the relation

$$P\Omega = P. \tag{8b}$$

The projection operator P transforms a d-dimensional subspace (spanned by Ψ^α) of the Hilbert space into another subspace of the same dimensionality (the model space, spanned by Ψ_0^α), and the wave operator Ω transforms the model space back to the former subspace. This does not mean, however, that Ω is the inverse of P in the ordinary sense. In figure 1(a), (b) we have illustrated the projection and wave operators in a very
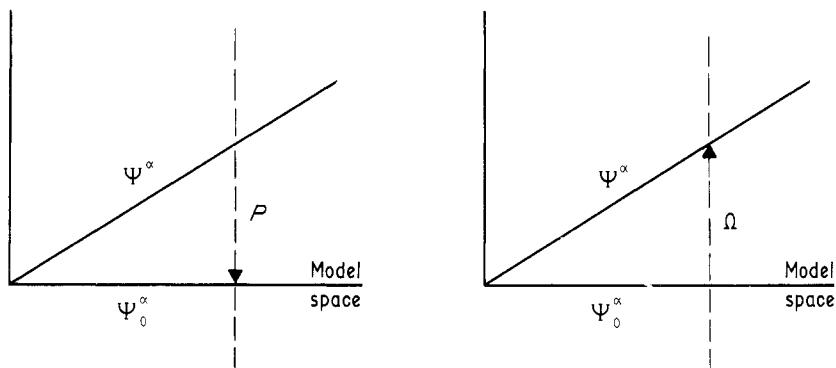

Figure 1. Simple illustrations of the projection and wave operators.

simple way. The horizontal axis here represents the model space and the vertical axis the orthogonal space. The projection operator P projects out of any function in the combined space its components in the model space. The wave operator transforms any function with the projection Ψ_0^α into Ψ^α, independent of the component in the orthogonal space. Therefore P has no effect in ΩP and, similarly, Ω has no effect in $P\Omega$ in accordance with (8a, b).

2.2. Effective operators

By means of the wave operators (7a) we can write the Schrödinger equation (1)

$$H\Omega\Psi_0^\alpha = E^\alpha\Omega\Psi_0^\alpha \qquad \alpha = 1, 2, \ldots, d \qquad (9)$$

for the states Ψ^α that correspond to the d model states. It should be observed that the wave operator Ω is the same for all these states. If we operate with P from the left on this equation, we obtain

$$PH\Omega\Psi_0^\alpha = E^\alpha\Psi_0^\alpha \qquad \alpha = 1, 2, \ldots, d. \qquad (10)$$

We then find an operator

$$H_{\mathrm{eff}} = PH\Omega = PH_0P + PV\Omega \qquad (11)$$

which satisfies the eigenvalue equation

$$H_{\mathrm{eff}}\Psi_0^\alpha = E^\alpha\Psi_0^\alpha \qquad \alpha = 1, 2, \ldots, d. \qquad (12)$$

This 'effective Hamiltonian', which is the same for the entire model space, generates d of the exact eigenvalues of the full Hamiltonian, although it operates only within the model space. The eigenfunctions of the effective Hamiltonian are the projections of the corresponding true wavefunctions, ie the model functions. These functions are not necessarily orthogonal and therefore the effective Hamiltonian is generally non-hermitian. (Its eigenvalues are, of course, always real, since they correspond to true energies.) We shall return to the question of the non-hermiticity in § 2.5.

It is often convenient to use the intermediate normalization, which means that the model function—and not the true wavefunction—is normalized, ie

$$\langle\Psi_0^\alpha|\Psi_0^\alpha\rangle = \langle\Psi_0^\alpha|\Psi^\alpha\rangle = 1.$$

The exact energy is then given by

$$E^\alpha = \langle\Psi_0^\alpha|H_{\mathrm{eff}}|\Psi_0^\alpha\rangle. \qquad (13)$$

The wave operator can also be used to construct effective operators corresponding to other quantities than the energy, but we shall not be further concerned with this problem here. (A brief discussion is given in § 3.4.)

2.3. Brillouin–Wigner expansion

By operating with Q from the left on the Schrödinger equation

$$(E^\alpha - H_0)\Psi^\alpha = V\Psi^\alpha \qquad (14a)$$

we obtain

$$(E^\alpha - H_0)Q\Psi^\alpha = QV\Psi^\alpha. \qquad (14b)$$

This equation can be solved for $Q\Psi^\alpha$ by introducing a resolvent T^α (see eg Messiah 1965, Löwdin 1966, 1968), defined by

$$T^\alpha(E^\alpha - H_0) = Q \qquad T^\alpha Q = T^\alpha. \tag{15}$$

This operator is the inverse of $(E^\alpha - H_0)$ in the orthogonal space (and zero in the model space) and is often written

$$T^\alpha = \frac{Q}{E^\alpha - H_0}. \tag{16a}$$

It can also be given a more explicit form using the 'spectral resolution' (Löwdin 1968). (15) gives directly

$$T^\alpha(E^\alpha - H_0)|r\rangle = T^\alpha(E^\alpha - E_0^r)|r\rangle = Q|r\rangle$$

where $|r\rangle$ is an eigenfunction of H_0(3). This shows that T^α has the same eigenfunctions as H_0 in the orthogonal space and that the corresponding eigenvalues are $(E^\alpha - E_0^r)^{-1}$. We then get the expansion

$$T^\alpha = \sum_{r \notin D} \frac{|r\rangle\langle r|}{E^\alpha - E_0^r}. \tag{16b}$$

By operating with T^α from the left on (14b) we get (using 15)

$$Q\Psi^\alpha = T^\alpha V \Psi^\alpha$$

or

$$\Psi^\alpha = \Psi_0^\alpha + T^\alpha V \Psi^\alpha. \tag{17}$$

This gives immediately the Brillouin–Wigner (BW) expansion

$$\Psi^\alpha = (1 + T^\alpha V + T^\alpha V T^\alpha + \ldots)\Psi_0^\alpha. \tag{18}$$

By means of the expansion (16b) of the resolvent this leads directly to the well known expansion in matrix form.

From (17) it follows that we can introduce an operator X^α, defined by

$$X^\alpha = P + T^\alpha V X^\alpha \tag{19}$$

which converts the model function Ψ_0^α into the correspondent true wavefunction Ψ^α,

$$\Psi^\alpha = X^\alpha \Psi_0^\alpha.$$

We can also define an 'effective interaction' W^α satisfying the equation

$$W^\alpha = V X^\alpha = VP + V T^\alpha W^\alpha \tag{20}$$

or

$$W^\alpha = VP + V T^\alpha VP + V T^\alpha V T^\alpha VP + \ldots.$$

We then obtain an 'effective Hamiltonian'

$$H_{\text{eff}}^\alpha = PH_0P + P W^\alpha \tag{21a}$$

which satisfies the eigenvalue equation

$$H_{\text{eff}}^\alpha \Psi_0^\alpha = E^\alpha \Psi_0^\alpha. \tag{21b}$$

This equation, however, differs from the previous effective-operator equation (11, 12)

in one important respect. Since T^α in (16) depends on the exact energy E^α of the state considered, the effective Hamiltonian H^α_{eff} also becomes energy dependent. Therefore, the equation (21) can only be solved in an iterative way for one energy level at a time. The effective Hamiltonian discussed in the previous section, on the other hand, is exactly energy independent and the same for all states corresponding to the model states. Although it is possible to eliminate the energy dependence of the BW operators by means of series expansions, (Brandow 1967), we shall find it more convenient to start directly from a perturbation expansion of RS type.

2.4. Generalized Rayleigh–Schrödinger expansion

2.4.1. General. In order to obtain a perturbation expansion that is independent of the energy, we operate on the Schrödinger equation (14a) by the wave operator (7a, 8a) $\Omega = \Omega P$. This gives

$$(E^\alpha - \Omega H_0 P)\Psi^\alpha = \Omega V \Psi^\alpha. \tag{22}$$

We can now eliminate E^α by subtracting this equation from the Schrödinger equation,

$$(\Omega H_0 - H_0 \Omega)\Psi^\alpha_0 = (V\Omega - \Omega V \Omega)\Psi^\alpha_0. \tag{23}$$

This equation holds for all d model functions of the model space, and the wave operator Ω is the same in all these cases. Since the operators on both sides of this equation give a null result when operating on the orthogonal space, it follows that the operator relation

$$[\Omega, H_0] = V\Omega - \Omega V \Omega \tag{24}$$

holds in the entire space. The equation is the basic formula for the generalized Rayleigh–Schrödinger perturbation introduced here, and it represents the main result of the present work. We shall demonstrate below that this leads to an energy-independent perturbation expansion, which is valid also for a model space containing several (degenerate) energy levels, ie of multi-configurational type. Before we do so we shall show that it yields the familiar RS expansion, when all model functions have the same energy (single configuration).

2.4.2. Completely degenerate model space. We assume now that all states of the model space have the same unperturbed energy E_0, ie

$$H_0 \Psi^\alpha_0 = E_0 \Psi^\alpha_0 \qquad \alpha = 1, 2, \ldots, d. \tag{25}$$

This gives

$$[\Omega, H_0]\Psi^\alpha_0 = (E_0 - H_0)\Omega\Psi^\alpha_0 \qquad \alpha = 1, 2, \ldots, d \tag{26}$$

and instead of (24)

$$(E_0 - H_0)\Omega = V\Omega - \Omega V \Omega. \tag{27}$$

By introducing a resolvent R in analogy with (15)–(16)

$$R(E_0 - H_0) = Q \qquad RQ = R$$

$$R = \frac{Q}{E_0 - H_0} = \sum_{r \notin D} \frac{|r\rangle\langle r|}{E_0 - E^r_0} \tag{28}$$

we can transform this equation into

$$Q\Omega = R(V\Omega - \Omega V\Omega)$$

or since $P\Omega = P$ according to (8b),

$$\Omega = P + R(V\Omega - \Omega V\Omega). \tag{29}$$

This expression was first derived by Bloch (1958, equation (31)). (Our wave operator Ω is identical to the U operator used by Bloch.)

In order to obtain an expansion of Ω we set

$$\Omega = \Omega^{(0)} + \Omega^{(1)} + \Omega^{(2)} + \ldots \tag{30}$$

where $\Omega^{(n)}$ represents the part of Ω that contains n interactions of the perturbation V. Obviously, we have

$$\Omega^{(0)} = P \tag{31}$$

and the following terms are given by the recursion formula

$$(E_0 - H_0)\Omega^{(n)} = QV\Omega^{(n-1)} - \sum_{m=1}^{n-1} \Omega^{(n-m)} V\Omega^{(m-1)} \tag{32a}$$

or

$$\Omega^{(n)} = RV\Omega^{(n-1)} - R \sum_{m=1}^{n-1} \Omega^{(n-m)} \Omega V^{(m-1)}. \tag{32b}$$

This leads to the well known Rayleigh–Schrödinger expansion (Bloch 1958, Messiah 1965)

$$(E_0 - H_0)\Omega^{(1)} = QVP$$

$$(E_0 - H_0)\Omega^{(2)} = QV\Omega^{(1)} - \Omega^{(1)} VP \tag{33}$$

$$(E_0 - H_0)\Omega^{(3)} = QV\Omega^{(2)} - \Omega^{(1)} V\Omega^{(1)} - \Omega^{(2)} VP$$

etc, or

$$\Omega^{(1)} = RVP$$

$$\Omega^{(2)} = RVRVP - R^2 VPVP \tag{34}$$

$$\Omega^{(3)} = RVRVRV - RVR^2 VPVP - R^2 VPVRVP - R^2 VRVPVP - R^3 VPVPVP$$

etc, where the resolvent R is given by (28).

2.4.3. Model space with several energy levels. The formulae of the previous section are no longer valid, if there are several different energy levels in the model space. In this case we have to use the more general equation (24) as the starting point for the perturbation expansion. In analogy with (32a) we get the recursion formula

$$[\Omega^{(n)}, H_0] = QV\Omega^{(n-1)} - \sum_{m=1}^{n-1} \Omega^{(n-m)} V\Omega^{(m-1)} \tag{35}$$

ie in the lowest orders

$$[\Omega^{(1)}, H_0] = QVP$$

$$[\Omega^{(2)}, H_0] = QV\Omega^{(1)} - \Omega^{(1)}VP \qquad (36)$$

$$[\Omega^{(3)}, H_0] = QV\Omega^{(2)} - \Omega^{(1)}V\Omega^{(1)} - \Omega^{(2)}VP$$

etc.

The equations above are somewhat more complicated to solve than the corresponding equations (33) due to the commutator. Therefore, we have to analyse them further before we can formulate the equivalent of (34).

The equations (36) are of the form

$$[\Omega^{(n)}, H_0] = A. \qquad (37)$$

By taking an arbitrary matrix element of both sides we get

$$\langle r|[\Omega^{(n)}, H_0]|a\rangle = (E_0^a - E_0^r)\langle r|\Omega^{(n)}|a\rangle$$

$$= \langle r|A|a\rangle$$

or

$$\langle r|\Omega^{(n)}|a\rangle = \frac{\langle r|A|a\rangle}{E_0^a - E_0^r}. \qquad (38)$$

We assume here that $|a\rangle$ is in the model space and $|r\rangle$ in the orthogonal space, since all other elements are equal to zero (for $n > 0$) according to (8a, b). Therefore the denominator in (38) can never vanish. In order to compare this with the results for a completely degenerate model space, we form the analogous expression for the wave operator in § 2.4.2. From

$$(E_0 - H_0)\Omega^{(n)} = A \qquad (39)$$

we obtain

$$\langle r|(E_0 - H_0)\Omega^{(n)}|a\rangle = (E_0 - E_0^r)\langle r|\Omega^{(n)}|a\rangle$$

$$= \langle r|A|a\rangle$$

or

$$\langle r|\Omega^{(n)}|a\rangle = \frac{\langle r|A|a\rangle}{E_0 - E_0^r}. \qquad (40)$$

We now see that (38) and (40) differ only in their energy denominators. E_0 in (40), which is the energy of *all* states in the model space, is in the more general expression (38) replaced by the energy of the particular initial state $|a\rangle$ (which is always in the model space). The denominator is therefore always equal to the difference in zero-order energy between the initial and the final states for the operator A. This makes it convenient to describe the perturbations in terms of diagrams, as we shall demonstrate in the next section. Essentially the same result as that obtained here, can be achieved by starting from an exactly degenerate model space and summing certain diagrams to all orders

(Brandow 1967). The present approach is more direct and yields the same result without any summation of infinite series. We shall return to this question in § 4.

In the previous section we introduced the resolvent R in order to solve the equations (27). Instead of (28) we can, of course, as well use (39, 40) as the definition of this operator, ie

$$\langle r|RA|a\rangle = \frac{\langle r|A|a\rangle}{E_0 - E_0^r}. \tag{41}$$

Similarly, we can introduce a resolvent S in order to solve equations of the type (37), and define this operator by

$$\langle r|SA|a\rangle = \frac{\langle r|A|a\rangle}{E_0^a - E_0^r} \tag{42}$$

where again $|a\rangle$ is supposed to belong to the model space and $|r\rangle$ to the orthogonal space. It should be observed that the effect of this new resolvent depends not only on the final state $|r\rangle$ but also on the initial state $|a\rangle$ for the operator A that the resolvent is acting on. Therefore, S cannot be given any explicit form of the type (28), but (42) is, of course, a perfectly adequate definition of the resolvent.

The fact that the denominator in (42) depends on the initial as well as on the final state, makes it necessary to use a special notation in order to indicate exactly what operator the resolvent is acting on. For this purpose we shall use a parenthesis in the following way

$$[\Omega^{(n)}, H_0] = A \;\Rightarrow\; \Omega^{(n)} = S(A). \tag{43}$$

The right hand bracket) represents the initial state and the left hand bracket (the final state for the operator A. We shall soon see that this convention leads to a convenient and unambiguous notation when more complicated expressions are built up.

We are now in the position to write down the formal solutions of the equations (35, 36) in analogy with (32–34). Since the two sets of equations differ only in their left hand sides, we only have to replace the resolvent R in the previous equations by the new resolvent S introduced above. We then get the general formula

$$\Omega^{(n)} = S(V\Omega^{(n-1)}) - S\left(\sum_{m=1}^{n-1} \Omega^{(n-m)}V\Omega^{(m-1)}\right) \tag{44}$$

and, particularly, for the lowest orders

$$\begin{aligned}
\Omega^{(1)} &= S(VP) \\
\Omega^{(2)} &= S(V\Omega^{(1)}) - S(\Omega^{(1)}VP) \\
\Omega^{(3)} &= S(V\Omega^{(2)}) - S(\Omega^{(1)}V\Omega^{(1)}) - S(\Omega^{(2)}VP)
\end{aligned} \tag{45}$$

etc.

In order to get the equivalent of (34) we have to keep track of the different resolvents and exactly what operators they are acting on. But we can easily see that this is automatically taken care of by the parentheses. For instance, in $\Omega^{(1)}$ the resolvent acts only on VP and therefore we must keep the parenthesis around that expression, when $\Omega^{(1)}$

is substituted into $\Omega^{(2)}$ etc. We then get

$$\Omega^{(1)} = S(VP)$$

$$\Omega^{(2)} = S(VS(VP)) - S(S(VP)VP)$$

$$\Omega^{(3)} = S(VS(VS(VP))) - S(VS(S(VP)VP)) - S(S(VP)VS(VP)) - S(S(VS(VP))VP)$$

$$+ S(S(S(VP)VP)VP)$$

(46)

etc, where the resolvent S is given by (42).

Ordinary algebraic rules are used here to associate each left hand bracket with the correct right hand one. These expressions, which are valid for a model space with several energy levels, are completely analogous to the expressions (34) for an exactly degenerate model space, with the important exception that in the general case the 'range' of each resolvent has to be shown explicitly.

An even more compact notation can be used here, which we shall find useful in the following, namely

$$\Omega^{(1)} = (V)$$

$$\Omega^{(2)} = (V(V)) - ((V)V)$$

(47)

$$\Omega^{(3)} = (V(V(V))) - (V((V)V)) - ((V)V(V)) - ((V(V))V) + (((V)V)V)$$

etc.

We have here simply removed all resolvents and projection operators. It is understood that each right hand bracket is associated with a state in the model space and each left hand bracket with a state in the orthogonal space. Each complete parenthesis yields an energy denominator, equal to the corresponding unperturbed energy difference.

2.5. Non-hermiticity of the effective Hamiltonian

We have previously mentioned (§ 2.2) that the effective Hamiltonian is generally not hermitian, since it has non-orthogonal eigenvectors. In this section we shall demonstrate this non-hermiticity more explicitly.

The hermitian adjoint of the wave operator Ω (7a) is defined by

$$\langle \Psi^\alpha | = \langle \Psi_0^\alpha | \Omega^+ \qquad \alpha = 1, 2, \ldots, d$$

$$Q\Omega^+ = 0.$$

(48)

Its expansions are obtained by forming the hermitian adjoint of (36), ie

$$[H_0, \Omega^{(1)+}] = PVQ$$

$$[H_0, \Omega^{(2)+}] = \Omega^{(1)+}VQ - PV\Omega^{(1)+}$$

(49)

etc.

These equations can, of course, be solved by means of a resolvent, similar to the S operator introduced above. However for the present purpose this will not be necessary. From the first equation above we get directly

$$\langle a | \Omega^{(1)+} | r \rangle = \frac{\langle a | V | r \rangle}{E_0^a - E_0^r}$$

(50)

where, as usual, we assume that $|a\rangle$ is in the model space and $|r\rangle$ in the orthogonal space. An arbitrary matrix element of the first-order term of the effective Hamiltonian (11), $PV\Omega^{(1)}$, is according to (46)

$$\langle b|V\Omega^{(1)}|a\rangle = \sum_{r\notin D} \frac{\langle b|V|r\rangle\langle r|V|a\rangle}{E_0^a - E^r} \tag{51a}$$

while the corresponding element of the hermitian adjoint operator, $\Omega^{(1)+}\,VP$, is according to (50)

$$\langle b|\Omega^{(1)+}V|a\rangle = \sum_{r\notin D} \frac{\langle b|V|r\rangle\langle r|V|a\rangle}{E_0^b - E_0^r}. \tag{51b}$$

The two expressions above differ in their energy denominators; the former contains the energy of the initial state $|a\rangle$, while the latter contains the energy of the final state $|b\rangle$. The operator $PV\Omega^{(1)}$ is therefore not hermitian in the general case. Both the initial and the final states are in the model space, and if this space is completely degenerate, there is no difference between the two expressions, and the operator becomes hermitian. However the higher-order terms become non-hermitian also in this case, as demonstrated by des Cloizeaux (1960). Then symmetric terms, like $PVRVRVP$, are hermitian, while 'mirror' terms like, PVR^2VPVP and $PVPVR^2VP$, are the hermitian adjoints. Since the expression for the $PV\Omega$ (using 34) is not completely symmetric, it follows that the effective Hamiltonian (11) is non-hermitian. In the more general case studied here, where the model space need not be completely degenerate, it follows from the analysis above that also symmetric terms are generally non-hermitian.

des Cloizeaux (1960) has shown that it is possible to transform the effective operator (without changing its eigenvalues) to a hermitian form. In the case of an exactly degenerate model space (which is the only case he considers), this leads to a symmetrized form in terms of the P, V and R operators. A transformation of a similar kind can, in principle, be applied in the more general case. This transformation, however, will be very complicated, and we shall not consider this problem further here. There is, of course, no particular difficulty in solving the eigenvalue equation (12) as it stands, even if the effective Hamiltonian is non-hermitian. We shall return to this problem briefly in § 3.4.

3. Linked-diagram expansion

3.1. General

As mentioned previously, the linked-diagram theorem can be proved by time-dependent (Goldstone 1957, Johnson and Baranger 1971) or by time-independent (Brandow 1967, Sandars 1969) methods. In the present section we shall extend the time-independent proof of Sandars, to the case of a multi-configurational model space, using the new RS formalism discussed above. Essentially the same results as those obtained here can be derived by other (time-dependent or time-independent) methods, and we shall return to the question of comparing different methods in § 4. In that section we shall also make some comments about the convergence of the perturbation expansion, particularly in the light of recent works of Weidenmüller *et al* (Schucan and Weidenmüller 1972, 1973, Hofmann *et al* 1973, 1974).

3.2. *Diagrammatic representation of perturbations*

In order to use the diagrammatic representation of the perturbation expansion, it is necessary to have an orbital description of the wavefunction. Therefore, we assume that our model Hamiltonian (2) is composed of one-particle operators

$$H_0 = \sum_{i=1}^{N} h_0(i) \tag{52}$$

where index i runs over the N particles of the system. Our basis functions (3) can then be single determinants of one-particle orbitals

$$\phi_a = |a\rangle = \det\{\phi_1(1)\phi_2(2)\ldots\phi_N(N)\} \tag{53}$$

with

$$h_0\phi_i = \epsilon_i\phi_i.$$

Obviously, we then have

$$H_0|a\rangle = E_0^a|a\rangle \quad \text{with} \quad E_0^a = \sum_i \epsilon_i \tag{54}$$

where i runs over the orbitals of the determinant $|a\rangle$. States with the same eigenvalue of the model Hamiltonian are said to form a configuration. The model space is defined by the eigenfunctions of H_0 belonging to one or several configurations.

It is convenient to separate the orbitals into the following three categories:

(*a*) *core orbitals*, defined as orbitals occupied in all determinants of the model space,

(*b*) *valence or open-shell orbitals*, defined as orbitals occupied in some but not all determinants of the model space,

(*c*) *excited orbitals*, defined as orbitals not occupied in any determinant of the model space.

We shall assume that all orbitals with the same eigenvalue belong to the same category. As an example we can consider the Be atom with the 2s and 2p shells regarded as valence shells. The orbitals of the 1s shell would then be the only core orbitals, and the model space would consist of the configurations $1s^2 2s^2$, $1s^2 2s 2p$ and $1s^2 2p^2$.

In the diagrammatic representation we shall essentially follow the conventions adopted in the current literature. Thus, a core orbital is represented by a vertical line with an arrow pointing downwards and an excited orbital by a vertical line with an arrow pointing upwards. For valence orbitals we use Sandars' notation with a double arrow, normally pointing upwards (see figure 2). (In Sandars' diagrams time flows from the right to the left, but we shall here use the normal convention of a vertical time axis.) The perturbation V is assumed to consist of one-particle (f) and two-particle (g) operators, ie in second-quantized form (see eg Judd 1967)

$$V = \sum_{pr} a_p^+ \langle p|f|r\rangle a_r + \tfrac{1}{2} \sum_{pqrs} a_p^+ a_q^+ \langle pq|g|rs\rangle a_s a_r. \tag{55}$$

As usual, the interactions are represented in the diagrams by horizontal, broken lines and there is a matrix element associated with each such line. The creation operators are associated with outgoing orbital lines and absorption operators with incoming orbital lines with respect to the interaction vertex, as illustrated in figure 3. The wavefunction or the wave operator can also be represented by similar diagrams. Since the wave operator operates to the right only on the model space (according to equation (8a)), the corresponding diagrams can have no free core or excited lines at the bottom.
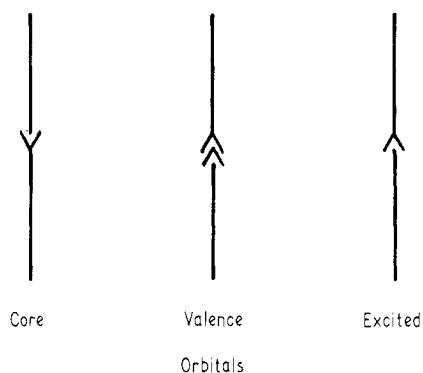
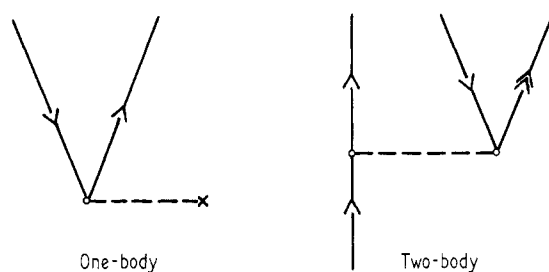**Figure 2.** Notations for orbital lines in the diagrammatic representation.

Figure 3. Examples of one- and two-body operator diagrams.

The rules for operating with an interaction diagram on a wave operator diagram can be obtained directly from the algebra of second quantization (see eg Goldstone 1957, Judd 1967, Sandars 1969). If the interaction diagram has free core or excited orbital lines at the bottom, these must be connected with corresponding outgoing lines of the wave operator diagram, since they are not allowed in the final diagram. Valence orbital lines at the bottom of the interaction diagram may or may not be connected with outgoing valence lines of the wave operator diagram. All possible ways of making these connections lead to allowed diagrams. This technique is illustrated in figure 4.

The standard rules for evaluating diagrams are well known, and we shall not be concerned with their details here. The reader is referred to more complete works on the subject (eg Kelly 1969). We only mention in passing that there is an energy denominator for each interaction of a wave operator diagram and a phase factor depending on the structure of the diagram. The emphasis of the present work is on the modification of the linked-diagram theorem due to the energy splitting of the model space, which will essentially affect the denominators.

3.3. Linked diagram theorem

3.3.1. Formulation of the problem. We shall first make some definitions. A part of a diagram that is not connected to the rest of the diagram by any orbital or interaction lines is said to be *disconnected*. If a disconnected part has no other free lines than valence lines (or no free lines at all), it is said to be *closed*, and the entire diagram is then defined
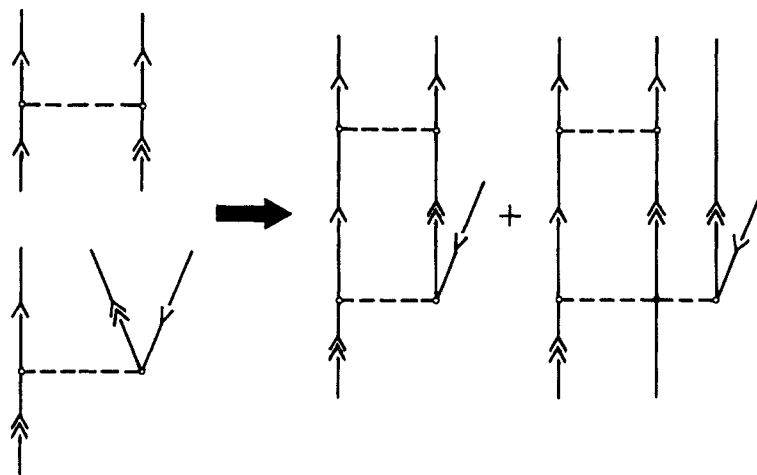
Figure 4. Illustration of the rules for operating on a wave operator diagram by the V interaction.

as *unlinked*. All other diagrams are *linked*. Examples of closed, linked and unlinked diagrams are given in figure 5. (Note that linked diagrams may consist of disconnected parts as long as no part is closed. Note also that the in- and outgoing lines of a closed part may in the general case represent different valence shells.)

The linked-diagram theorem states that the wavefunction or the wave operator is in each order represented by all possible linked diagrams obtained by operating with the perturbation on the wave operator of the next lower order according to the rules previously given. In order to prove this theorem in the general case of a model space that may contain several energy levels, we shall start from the recursion formula (35)

$$[\Omega^{(n)}, H_0] = QV\Omega^{(n-1)} - \sum_{m=1}^{n-1} \Omega^{(n-m)}V\Omega^{(m-1)}. \tag{56}$$

When V operates on $\Omega^{(n-1)}$, linked as well as unlinked diagrams may be formed. The linked-diagram theorem then implies that the unlinked diagrams obtained in this
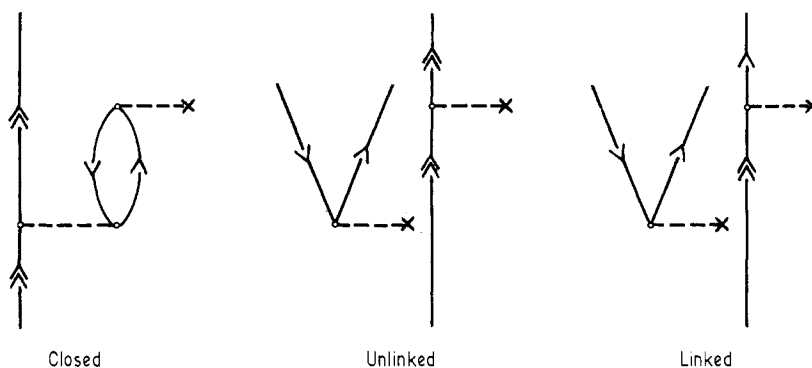

| Closed | Unlinked | Linked |

Figure 5. Examples of closed, unlinked and linked diagrams.

operation are essentially cancelled by the remaining terms of (56) so that only linked diagrams remain. We shall now prove this theorem by the method of induction.

We assume that the linked-diagram theorem holds to order $(n-1)$, and we shall show that it then holds in order n. $\Omega^{(n-1)}$ is now assumed to be linked, which means that it does not have any closed parts. When V operates on $\Omega^{(n-1)}$, unlinked diagrams may be formed in different ways. If $\Omega^{(n-1)}$ contains disconnected parts, ie parts that are not connected by any orbitals or interaction lines, one or several of these can be closed by the nth interaction (see figure 6). Besides, of course, the last interaction can form a closed
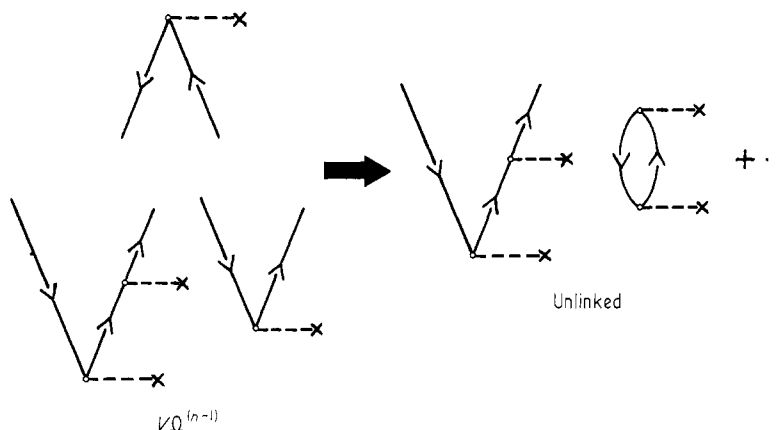
Figure 6. Example of formation of unlinked diagrams by operating on a wave operator diagram.

part by itself, or it can close the entire diagram. In the last case the final state will be in the model space, and the diagram will be cancelled by the Q operator in (56). Since all diagrams of $\Omega^{(n-1)}$ are assumed to be linked, there cannot be more than one closed part in a diagram of $QV\Omega^{(n-1)}$, and the number of interactions in the closed part can be $m = 1, 2 \ldots (n-1)$. We shall now show that all unlinked diagrams of $QV\Omega^{(n-1)}$, with m interactions in the closed part, are cancelled by the term $\Omega^{(n-m)}V\Omega^{(m-1)}$ in (56) (after introducing so-called backwards or folded diagrams). Then the linked-diagram theorem follows directly.

3.3.2. Cancellation of unlinked diagrams: an illustrative example. As a simple illustration of the cancellation of unlinked diagrams we consider the disconnected wave operator diagram in figure 7. It follows from the rules for constructing the diagrams that all possible permutations of the relative orderings of the interactions of the two parts will appear as allowed diagrams (see figure 8). It is also well known that the denominators can in such a case be factorized into a product of the denominators of the two parts, considered as separate diagrams. This is the 'factorization theorem' (Frantz and Mills 1960, see also the Appendix).

We now operate on the diagrams in figure 8 so that unlinked diagrams are produced (see figure 9). The factorization theorem can be applied to these unlinked diagrams as well, and the result is then a product of a linked and a closed diagram (apart from the order of the creation and absorption operators).
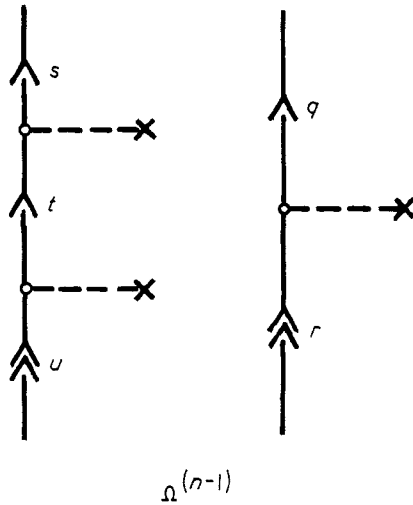
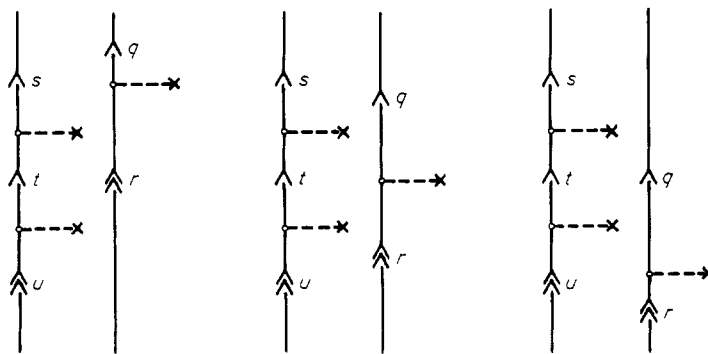**Figure 7.** Example of disconnected wave operator diagram.

Figure 8. Possible permutations of the interactions in the diagram of figure 7.

We then consider the term $\Omega^{(n-m)}V\Omega^{(m-1)}$ in (56). Since Ω operates to the right only on the model space (8a), we can write this term as $\Omega^{(n-m)}PV\Omega^{(m-1)}P$. Therefore the V interaction must here close the diagrams of $\Omega^{(m-1)}$, and the result is a product of wave operator diagrams, $\Omega^{(n-m)}$, and closed diagrams, $PV\Omega^{(m-1)}P$. Therefore this term will contain the product of the linked and the closed parts of the unlinked diagram in figure 9, which will then be essentially cancelled.

The cancellation is exact, possibly apart from the order of the creation and absorption operators, which we have not yet considered. Generally, the order of these operators is obtained by reading the outgoing, free lines of the diagram in one direction (normally from the left to the right) and the incoming, free lines in the opposite direction. Thus, for the unlinked diagram in figure 9 (which is a *single* diagram) this order is $a_s^+ a_p^+ a_r a_u$. For the corresponding product in $\Omega^{(n-m)}PV\Omega^{(m-1)}P$, on the other hand, this order is $a_s^+ a_u a_p^+ a_r$. Using the commutation rules we have

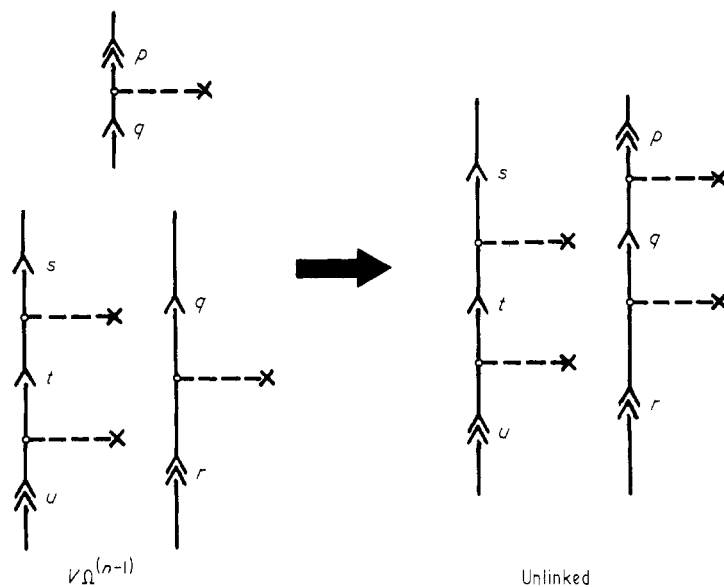$$a_s^+ a_p^+ a_r a_u = -\delta(p, u)a_s^+ a_r + a_s^+ a_u a_p^+ a_r. \tag{57}$$

Figure 9. Formation of an unlinked diagram by operating on the diagram of figure 7.

This means that we have to add a diagram corresponding to $-\delta(p, u)a_s^+ a_r$ in order to get an exact cancellation in this case. This diagram should have the same matrix elements and energy denominators as the original diagram in figure 9, and it can be obtained by joining the p and u lines (see figure 10). Such diagrams are called 'backward' by Sandars (1969) since a valence line is directed downwards, and they are essentially equivalent to the 'folded' diagrams discussed by Brandow (1967).
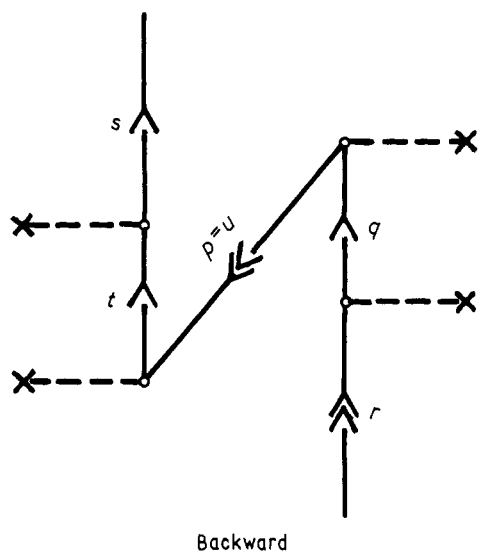

Backward

Figure 10. 'Backward' diagram associated to the diagram in figure 9.

In the time-dependent approach similar diagrams appear in order to compensate for the incorrect time ordering, caused by the factorization (Johnson and Baranger 1971, Kuo *et al* 1971). In the present work we shall not make any more detailed comparison of the backward diagrams used here and the different types of folded diagrams found in the literature.

3.3.3. Cancellation of unlinked diagrams: general. In order to show that there is a general cancellation of unlinked diagrams, we have primarily to establish a one-to-one correspondence between the factorized, unlinked diagrams of $QV\Omega^{(n-1)}$, with m interactions in the closed part, and the terms of $\Omega^{(n-m)}V\Omega^{(m-1)}$. Secondly, the creation and absorption operators have to be rearranged, which gives rise to the backward diagrams. We shall first establish this one-to-one correspondence and then return to the question of constructing the backward diagrams in § 3.3.4.

It follows from the rules of constructing the diagrams and the linked-diagram theorem that each separated part of a disconnected wave operator diagram is itself an allowed wave operator diagram (see figure 7) and that all combination of allowed wave operator diagrams lead to allowed diagrams. The linked-diagram theorem is assumed to be valid up to order $(n-1)$ and, hence, all disconnected diagrams of $\Omega^{(1)}$ $(1 \leqslant n-1)$ can be analysed in terms of wave operator diagrams of lower order. Obviously, the disconnected diagrams of $\Omega^{(1)}$ consists of all 'distinct' combinations of $\Omega^{(k_1)}, \Omega^{(k_2)}, \ldots$ with $\Sigma k_i = 1$. ('Distinct' is here used to exclude horizontal permutations of diagrammatic parts.) By means of the factorization theorem $\Omega^{(1)}$ can then be expressed by means of products of the corresponding wave operator diagrams (disregarding the order of the creation and absorption operators).

Let us by $\Omega_c^{(k)}$ denote all connected diagrams of $\Omega^{(k)}$. It then follows from the arguments above that we can formally express the wave operators of the lowest orders in the following way

$$\Omega^{(1)} : \Omega_c^{(1)}$$

$$\Omega^{(2)} : \Omega_c^{(2)} + \frac{1}{2!}\Omega_c^{(1)}\Omega_c^{(1)}$$

$$\Omega^{(3)} : \Omega_c^{(3)} + \Omega_c^{(2)}\Omega_c^{(1)} + \frac{1}{3!}\Omega_c^{(1)}\Omega_c^{(1)}\Omega_c^{(1)} \tag{58}$$

$$\Omega^{(4)} : \Omega_c^{(4)} + \Omega_c^{(3)}\Omega_c^{(1)} + \frac{1}{2!}\Omega_c^{(2)}\Omega_c^{(2)} + \frac{1}{2!}\Omega^{(2)}\Omega^{(1)}\Omega^{(1)} + \frac{1}{4!}\Omega^{(1)}\Omega^{(1)}\Omega^{(1)}\Omega^{(1)}$$

etc. We have here avoided the equality sign, since the order of the creation and absorption operators is not the same on both sides. In the general case we have

$$\Omega^{(1)} : \sum \prod_{k=1}^{1} \frac{1}{n_k!}[\Omega_c^{(k)}]^{n_k} \qquad 1 \leqslant n-1. \tag{59}$$

n_k is here the number of separated parts of the same order k, and the summation is performed over all combinations with

$$\sum n_k k = 1. \tag{60}$$

The factor $n_k!$ in the denominator of (59) compensates for the number of permutations, which do not lead to distinct diagrams. A formula analogous to (59) has been given by Brandow (1967) and others, for the effective interaction.

We now operate on $\Omega^{(1)}$ with the interaction V so that unlinked diagrams are obtained. As mentioned previously, the last interaction can form a closed part by itself, or, if the wavefunction is disconnected, it can close one or several parts. (If V is an N-body operator, it is obviously possible to connect and close up to N separated parts of the wave operator diagram by a single interaction.) As an example we consider the expression for $\Omega^{(4)}$ above. By closing $\Omega_c^{(1)}$, for instance, we get a closed part with *two* interactions, for which we use the symbol $\overline{\Omega_c^{(1)}}$. This operation on $\Omega_c^{(3)}\Omega_c^{(1)}$ then gives $\Omega_c^{(3)}\overline{\Omega_c^{(1)}}$. The fourth term of $\Omega^{(4)}$ has two factors of $\Omega_c^{(1)}$, and closing one or the other leads to equivalent diagrams. We can therefore close one of them and remove the factor of $1/2$, ie $\Omega_c^{(2)}\Omega_c^{(1)}\overline{\Omega_c^{(1)}}$. For the last term in $\Omega^{(4)}$ we can close $\Omega_c^{(1)}$ in four equivalent ways, and we should therefore remove a factor of four from the denominator. All unlinked diagrams with two interactions in the closed part obtained by operating with V on $\Omega^{(4)}$ can then be expressed as

$$\Omega_c^{(3)}\overline{\Omega_c^{(1)}} + \Omega_c^{(2)}\Omega_c^{(1)}\overline{\Omega_c^{(1)}} + \frac{1}{3!}\Omega_c^{(1)}\Omega_c^{(1)}\Omega_c^{(1)}\overline{\Omega_c^{(1)}} = \Omega^{(3)}\overline{\Omega_c^{(1)}}.$$

Unlinked diagrams with *three* interactions in the closed part can be obtained by closing $\Omega_c^{(2)}$, which in the case of $\Omega^{(4)}$ leads to $\Omega^{(2)}\overline{\Omega_c^{(2)}}$ in the same way as above. But we can also in some cases close two diagrams of $\Omega_c^{(1)}$ simultaneously. From the term

$$\frac{1}{2!}\Omega_c^{(2)}\Omega_c^{(1)}\Omega_c^{(1)}$$

we then get

$$\frac{1}{2!}\Omega_c^{(2)}\overline{\Omega_c^{(1)}\Omega_c^{(1)}}.$$

In the last term of $\Omega^{(4)}$ we can close two of the $\Omega_c^{(1)}$ parts in

$$\binom{4}{2} = \frac{4!}{2!2!}$$

equivalent ways, and the result is

$$\frac{1}{2!2!}\Omega_c^{(1)}\Omega_c^{(1)}\overline{\Omega_c^{(1)}\Omega_c^{(1)}}.$$

Altogether we then get by closing two $\Omega^{(1)}$ parts of $\Omega^{(4)}$

$$\frac{1}{2!}\Omega_c^{(2)}\overline{\Omega_c^{(1)}\Omega_c^{(1)}} + \frac{1}{2!2!}\Omega_c^{(1)}\Omega_c^{(1)}\overline{\Omega_c^{(1)}\Omega_c^{(1)}} = \frac{1}{2!}\Omega^{(2)}\overline{\Omega_c^{(1)}\Omega_c^{(1)}}.$$

Combining this with the result above, $\Omega^{(2)}\overline{\Omega_c^2}$, we see that all unlinked diagrams with three interactions in the closed part, obtained by operating with V on $\Omega^{(4)}$, are given by

$$\Omega^{(2)}\overline{\Omega_c^{(2)}} + \frac{1}{2!}\Omega^{(2)}\overline{\Omega_c^{(1)}\Omega_c^{(1)}} = \Omega^{(2)}\overline{\Omega^{(2)}}.$$

By means of the expressions (59–60) it is easy to extend the arguments above to the general case. The result can be expressed so that there is a one-to-one correspondence between the factorized, unlinked diagrams of $QV\Omega^{(n-1)}$, with m interactions in the closed

part, and the terms of $\Omega^{(n-m)}\overline{\Omega^{(m-1)}}$. Instead of $\overline{\Omega^{(m-1)}}$ we can use the previous notation, $PV\Omega^{(m-1)}$, and we have the important result

$$\{QV\Omega^{(n-1)}\}_{\text{UL}} = \sum_{m=1}^{n-1} \Omega^{(n-m)}PV\Omega^{(m-1)} + \frac{\text{backward}}{\text{diagrams}} \tag{61}$$

where UL indicates unlinked diagrams. It then follows from (56) that all unlinked diagrams are cancelled, and the final result becomes

$$[\Omega^{(n)}, H_0] = \{V\Omega^{(n-1)}\}_{\text{L}} \tag{62}$$

where L indicates linked *and* backward diagrams. This is the *linked-diagram theorem*, which states that the wave operator diagrams of a certain order are obtained by operating with the perturbation V on the wave operator diagrams of the next lower order, keeping only linked (and backward) diagrams. The rules for operating with the perturbation on a wave operator diagram are briefly described in § 3.2, and the rules for obtaining backward diagrams will be discussed in thé following section. The commutator in (62) yields an energy denominator, which is equal to the difference in the unperturbed energy of the initial state for the first interaction and the final state for the nth interaction (see equation (38)), ie between the bottom and the top of the diagram. This is in accordance with the normal rules for evaluating diagrams, but the important point is here that this has been shown without any assumption about the energies of the model space. This will be discussed further and compared with other ways of handling this problem in § 4.

3.3.4. Backward diagrams. In the previous section we have proved that there is a complete cancellation of unlinked diagrams in the wave operator if certain 'backward' diagrams are included. This cancellation was shown by establishing the one-to-one correspondence between the factorized, unlinked diagrams of $QV\Omega^{(n-1)}$ and the diagrams of $\Omega^{(n-m)}V\Omega^{(m-1)}$ in (56). However the order of the creation and absorption operators is not the same in the two terms. Therefore, these operators have to be rearranged, which leads to the backward diagrams. In this section we shall analyse these diagrams in some detail.

The order of the creation and absorption operators are in the unlinked diagrams of $QV\Omega^{(n-1)}$ given by the ordinary diagrammatic rule, for instance, by reading outgoing lines (creation operators) from the left to the right and incoming lines (absorption operators) from the right to the left. In $\Omega^{(n-m)}V\Omega^{(m-1)}$, on the other hand, these operators are associated with the two parts, $\Omega^{(n-m)}$ and $(P)V\Omega^{(m-1)}$, separately.

Let us consider a simple but non-trivial example shown in figure 11. A^+, B^+, C and D represent here groups of creation and absorption operators as indicated in the figure. If this diagram represents $QV\Omega^{(n-1)}$, the order of these operators should be, for instance, A^+B^+CD, while if it represents $\Omega^{(n-m)}V\Omega^{(m-1)}$ the order should be A^+DB^+C. Therefore, in order to get agreement we have to move the operators, for instance, those of the closed part in the unlinked diagram (B^+C) to the right of the operators of the linked part (A^+D). This leads to contractions between the valence shell operators of B^+ and D. Similar permutations are made, when ordinary diagrams are constructed by means of the rules of second-quantization (see § 3.2). It then follows that the backward diagrams are obtained by joining the valence lines of B^+ and D in all possible ways (see figure 12). Of course, only lines belonging to the same shell shall be joined. From the commutation rules for the creation and absorption operators one easily finds that the diagrams (a), (d) and (e) in figure 12 will have a phase factor -1 in addition to the usual phase rule
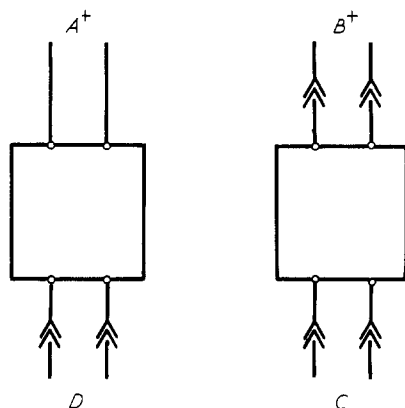
2462   *I Lindgren*

Figure 11. General unlinked diagram.

mentioned in § 3.2. The general phase rule can be formulated so that there is a change in sign, if there is an even number of crossings, when the backward lines are drawn. This is in accordance with the rule found by Sandars (1969) in the single-configurational case.

Generally, backward diagrams appear as soon as the normal rules for constructing the diagrams would lead to an unlinked diagram. Before the latter is discarded, all possible backward diagrams must be constructed by joining the outgoing valence lines of the closed part with the incoming valence lines of the linked part.

3.4. Effective-operator diagrams

In the previous sections we have been mainly concerned with the linked-diagram theorem for the wave operator (7a). In practical applications, however, it is usually more convenient to work with effective operators. By means of (11) we can easily construct the diagrams for the effective Hamiltonian. In the nth order ($n > 0$) we have

$$H_{\text{eff}}^{(n)} = PV\Omega^{(n-1)} \tag{63}$$

which means that the nth order diagrams of the effective Hamiltonian are obtained by closing the wave operator diagrams of order $(n-1)$ by the perturbation V in all possible ways. Since the diagrams of Ω are linked, ie do not have any closed part, there cannot be any disconnected closed parts in the diagrams of H_{eff}. If the wave operator diagram is disconnected, the separated parts must be connected by the closing interaction. The exact rules for forming the diagrams of the effective Hamiltonian follows from (62) and (63). It should be noted that there is no energy denominator associated with the last interaction in (63).

When the diagrams of H_{eff} have been constructed and evaluated with the desired accuracy, the eigenvalue equations (12) can be solved. This leads to a matrix equation of the same order as the dimensionality of the model space. Usually, however, this can be split up into a number of equations of lower order, using the known constants of the motion. Considering again as an example the Be atom with the configurations $1s^2 2s^2(^1S)$, $1s^2 2s2p(^1P, ^3P)$ and $1s^2 2p^2(^1S, ^1D, ^3P)$ in the model space, we have a total of 28 states in that space. With the electrostatic interaction as the only perturbation S, L, M_S, M_L and parity are good quantum numbers, and states which differ in any of these do not mix. Therefore, the 28-dimensional problem can be separated into 26
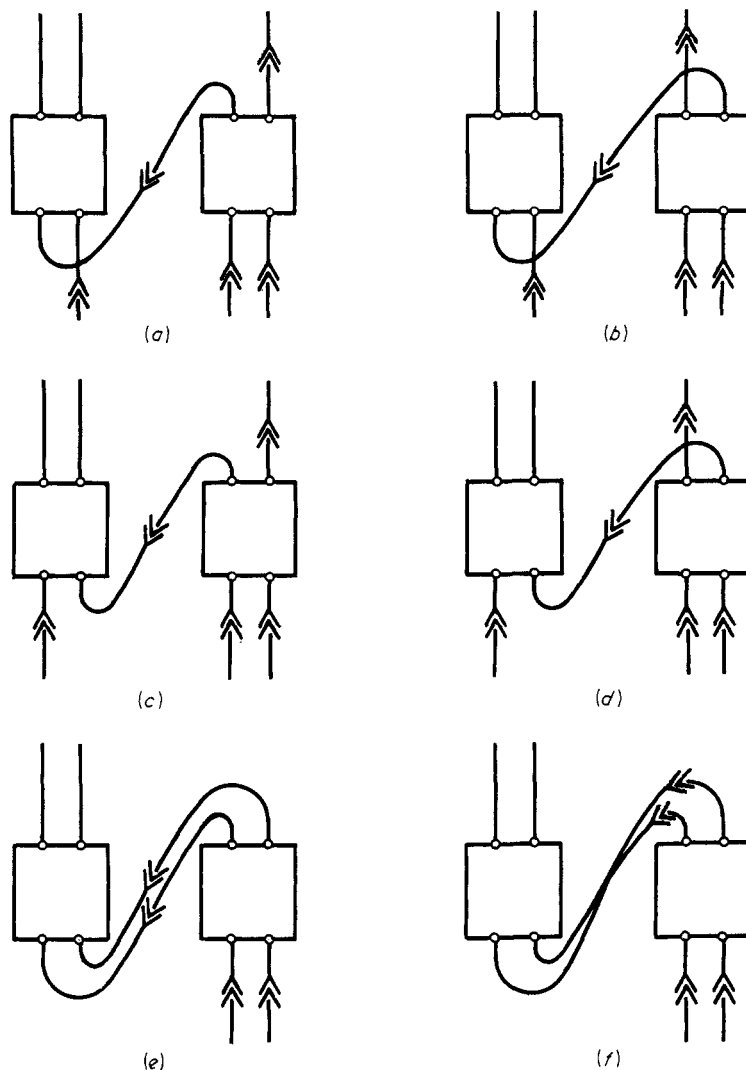
Figure 12. Backward diagrams associated with the unlinked diagram in figure 11.

one-dimensional ones (for the P and D states) and a two-dimensional one (for the two ^{1}S states). In the latter case there are four groups of diagrams (see figure 13), depending on the initial and the final states, and each group contributes to an element of the 2×2 matrix of the effective Hamiltonian. Of course, it is necessary here to include H_0 in the effective Hamiltonian, since its (diagonal) elements are not the same in the two states. Solving this two-dimensional eigenvalue equation leads to the energies of the two ^{1}S states and the mixing of the two configurations in these cases.

Some precaution is required here, since the effective Hamiltonian is not generally hermitian, as demonstrated in § 2.5. This means that in our example diagrams starting from the $2s^2$ level and ending at the $2p^2$ level are not identical to diagrams going in the opposite direction. The reason for this is, of course, that with our conventions the initial state enters in the denominators while the final state does not. It is possible to define
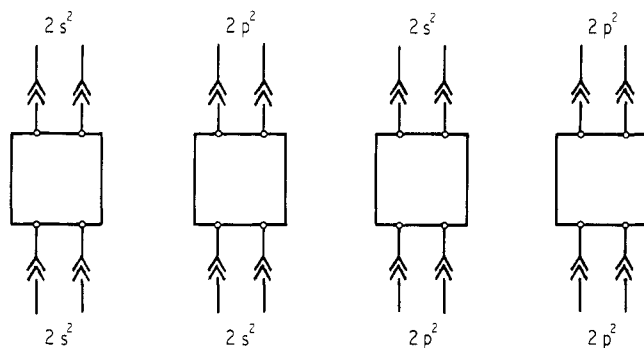
Figure 13. The four groups of effective two-body operator diagrams, appearing in the Be problem mentioned in the text. Each group of diagrams contributes to one matrix element of the 2×2 matrix for the ^1S state. (In addition one-body operator diagrams as well as H_0 contribute to the diagonal elements.)

the diagrams in such a way that the initial and final states appear in a symmetric way, which obviously leads to a hermitian operator (Johnson and Baranger 1971). Such a convention, however, is less convenient in a formalism like ours, which is based on the wave operator concept. In the wave operator the final state is not yet known, and a symmetric denominator cannot be used. Therefore we prefer to use the unsymmetric form of the effective operators.

It is also easy to construct the effective-operator diagrams corresponding to some other quantity (h) than the energy. Usually, one is interested in the effect of some small perturbation, like the hyperfine interaction. In order to get the 'effective interaction' h_{eff} one can then replace the perturbation V by $V + h$ and consider terms linear in h. It is easily found that the corresponding diagrams are obtained by replacing one of the V interactions in the diagrams of H_{eff} by h in all possible ways. The energy shift due to this interaction is according to (14) given by

$$\Delta E^\alpha = \langle \Psi_0^\alpha | h_{\text{eff}} | \Psi_0^\alpha \rangle. \tag{64}$$

Of course, the model functions Ψ_0^α have still to be determined by means of the eigenvalue equation (12). In many cases, however, these functions are uniquely determined for symmetry reasons, as illustrated for the Be atom above. Then the energy shift in (64) can be determined entirely without considering the main perturbation.

We shall not consider the effective-operator diagrams further in the present work, but we shall return to them in more detail in the connection with applications to the hyperfine interaction and other problems (Garpman *et al* 1974).

4. Discussion and conclusions

In the previous sections we have proved the linked-diagram theorem for a model space of the multi-configurational type, starting from an extended form of the Rayleigh–Schrödinger perturbation expansion. Similar results can be obtained by other methods, and it is therefore of interest to compare the various approaches in order to find out to what extent they are identical and to what extent there are significant differences. An instructive comparison between time-dependent and time-independent methods has recently been made in the review article by Barrett and Kirson (1973), and, therefore,

we shall not be further concerned with that aspect of the problem here. Instead, we shall make some comparisons between the present approach and that of Brandow (1967), which are both time-independent. The main difference between the two approaches is that Brandow starts from the Brillouin–Wigner type of perturbation expansion, while in the present work the Rayleigh–Schrödinger type is used.

In order to deal with a model space that is not exactly degenerate, Brandow introduces an additional perturbation

$$V_1 = \sum a_i^+ (\epsilon_i - \epsilon_0) a_i \tag{65}$$

summed over all *valence* shells, and the same term is removed from the model Hamiltonian H_0

$$H_0' = H_0 - V_1. \tag{66}$$

This has no effect on the orbitals, but it will shift the orbital energy for *all* valence shells to an arbitrary value ϵ_0. This means that all configurations of the model space (which differ only in the occupation numbers of the valence shells) will be exactly degenerate. It is then possible to use the technique for a single-configurational model space, provided the effect of the perturbation (65) is considered.

The extra perturbation (65) has only diagonal elements, and, therefore, it is easy to sum its contributions to all orders. Let us consider the situations in figure 14, where V_1
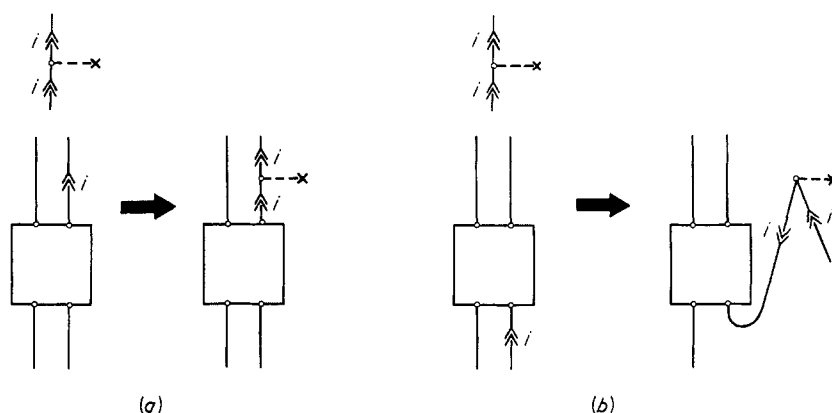

Figure 14. Modification of valence-orbital lines by the diagonal perturbation (65). Modification of an incoming line leads to a backward diagram (*b*).

is operating on diagrams with outgoing or incoming valence lines. The former case leads to a normal and the latter case to a backward diagram (see § 3.3.4). The resulting diagram is equal to $\pm (\epsilon_i - \epsilon_0)/\Delta E$ times the original diagram, where ΔE is the last energy denominator of that diagram. The minus sign appears in the backward case (*b*), since there is no crossing. This process can be repeated, and it leads to a geometric series, which is easily summed,

$$1 \pm \frac{\epsilon_i - \epsilon_0}{\Delta E} + \left(\frac{\epsilon_i - \epsilon_0}{\Delta E} \right)^2 + \ldots = \frac{\Delta E}{\Delta E \mp (\epsilon_i - \epsilon_0)}. \tag{67}$$

ΔE is the difference between the energy of the initial and the final states with the *modified* model Hamiltonian (66). Therefore, the effect of multiple interactions by the artificial

perturbation V_1 can be included to all orders simply by replacing the modified orbital energy ϵ_0 by the original energy ϵ_i for the valence shells in the denominators. This means that the energy denominator is now given by the energy difference between the initial and final states corresponding to the *original* model Hamiltonian H_0. It can be shown that this result holds also in more complicated cases than that shown in figure 14. Exactly the same result is obtained with the technique introduced in the present work, using the RS formalism for a multi-configurational model space, without any summation of infinite series (see equations 38 and 62).

There is one important formal difference between the result obtained in this section and that obtained previously. The series (67) is certainly convergent, if

$$|\epsilon_i - \epsilon_0| < |\Delta E|$$

or, more generally, if

$$|\Delta_1 E| < |\Delta E| \tag{68}$$

where $\Delta_1 E$ is the energy shift caused by the perturbation V_1 in (65), and ΔE is the distance to an unperturbed level outside the model space. Therefore, (68) is always fulfilled, if there is no overlap between the zero-order energies of the model space and those of the orthogonal space (for a reasonable choice of ϵ_0). This is certainly a *sufficient* condition for the convergence of the series (67). This result is in agreement with the more general conclusion of Schucan and Weidenmüller (1972), namely that the entire perturbation expansion can be expected to be convergent, if there is no overlap between the *true* energies of the states corresponding to the model space and those corresponding to the orthogonal space. In the present approach, however, there is no perturbation of the type (65), and, hence, there is no restriction, similar to (68), on the zero-order splitting of the model space. But the two approaches lead to identical results, which means that the summation of the diagrams due to V_1 in (67) can be made, even if (68) is not fulfilled. Therefore the zero-order splitting of the model space has no effect on the convergence, which instead should be determined by the shifts of the model states due to *real* perturbations relative to the levels originating from the orthogonal space. If crossings occur (when the perturbation is turned on smoothly), convergence problems are likely to appear, and non-perturbative techniques, such as Padé approximants (Hofmann *et al* 1973, 1974) may be used. On the other hand, it is not necessary that the model space yields *all* states within a certain energy range for the perturbation expansion to converge. Instead, one can include the strongest interacting levels in the model space (regardless of their zero-order energy separation), and leave out weakly interacting ones, even if they are close in energy. This gives great flexibility in choosing the model space, which can be a useful tool, in addition to the non-perturbative methods mentioned above, in reducing convergence difficulties.

Acknowledgments

Appendix. Factorization theorem

The factorization theorem, which is a crucial part of the time-independent proof of the linked-diagram theorem, can be formulated in the following way. Consider a diagram, which has no intermediate state within the model space and which is partly disconnected at one end (see figure A1). If a summation is performed over all possible relative orderings
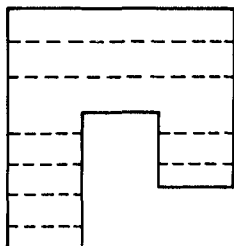

Figure A1. Partially disconnected diagram.

between the interactions of the disconnected parts (keeping the relative ordering within each part intact), the results can be expressed by means of a single diagram with the denominators determined independently for each part of the diagrams (see eg Frantz and Mills 1960).

We shall begin the proof of this general theorem by considering a simple example, shown in figure A2. If we let a and r denote the matrix elements associated with the
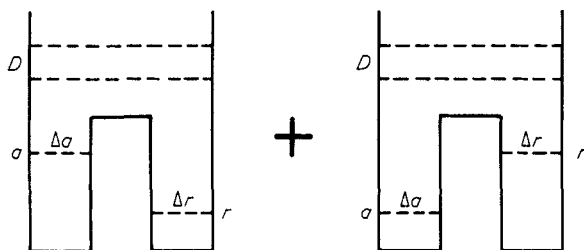

Figure A2. Illustration of the simple factorization in equation (A1).

interactions indicated and Δ_a and Δ_r the corresponding excitation energies, we can express the sum of two diagrams as

$$\frac{ar}{(\Delta_a + \Delta_r)\Delta_r}D + \frac{ar}{\Delta_a(\Delta_a + \Delta_r)}D = \frac{ar}{\Delta_a\Delta_r}D \qquad (A1)$$

where D stands for the remaining part of the diagram. This is the simplest possible illustration of the factorization theorem, and it is the only case we have to calculate explicitly. The general theorem can be proved by means of repeated use of this relation.

Next we consider a more general diagram shown in figure A3, and introduce the following notations:

$$(A) = (a(b(c(\,\ldots\,)))) = \frac{abc\,\ldots}{(\Delta_a + \Delta_b + \Delta_c + \,\ldots\,)(\Delta_b + \Delta_c + \,\ldots\,)(\Delta_c + \,\ldots\,)}$$

$$A = a(b(c(\,\ldots\,))) = \frac{abc\,\ldots}{(\Delta_b + \Delta_c + \,\ldots\,)(\Delta_c + \,\ldots\,)} \tag{A2a}$$

$$(B) = (b(c(\,\ldots\,))) = \frac{bc\,\ldots}{(\Delta_b + \Delta_c + \,\ldots\,)(\Delta_c + \,\ldots\,)}$$
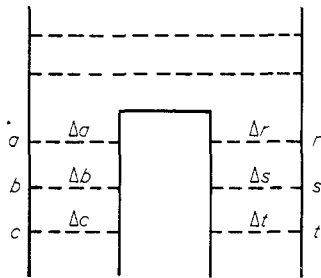
etc.

Figure A3. Illustration of the notations used in the mathematical expressions.

Again, $a, b, c \ldots$ represent matrix elements, $\Delta_a, \Delta_b, \Delta_c, \ldots$ the corresponding excitation energies, and the parentheses indicate the denominators in the same manner as in § 2 (equation (47)). We then have the following relations

$$(A) = (a(B)) = (a(b(C))) = \,\ldots$$
$$A = a(B) = a(b(C)) = \,\ldots\,. \tag{A3}$$

Similarly, for the other part of the diagram

$$(R) = (r(s(t(\,\ldots\,))))$$
$$R = r(s(t(\,\ldots\,))) \tag{A2b}$$
$$(S) = (s(t(\,\ldots\,)))$$

which gives

$$(R) = (r(S)) = (r(s(T))) = \,\ldots$$
$$R = r(S) = r(s(T)) = \,\ldots\,. \tag{A3a}$$

With these notations the factorization shown in (A1) can be written

$$(a(r)) + ((a)r) = (a)(r). \tag{A4a}$$

It is easy to show that this relation holds also if a and r represent several interactions. With the notations introduced above we then have, for instance,

$$(A(R)) + ((A)R) = (A)(R). \tag{A4b}$$

It follows directly from the definitions (A2) that we can permute the terms in such expressions, provided we move also the associated parentheses. For instance,

$$(A)(R) = (R)(A) \quad \text{and} \quad ((A)R) = (R(A)). \tag{A5}$$

By means of the relations (A3–5) we then get

$$(A)(R) = (A(R)) + (R(A)) = (a(B)(R)) + (r(S)(A)) = (a(B)(R)) + (r(A)(S)). \tag{A6a}$$

In the same way

$$\begin{aligned}(B)(R) &= (b(C)(R)) + (r(B)(S)) \\ (A)(S) &= (a(B)(S)) + (s(A)(T)).\end{aligned} \tag{A6b}$$

Combining (A6a) and (A6b) we get

$$(A)(R) = (a(b(C)(R))) + (a(r(B)(S))) + (r(a(B)(S))) + (r(s(A)(T))). \tag{A7}$$

When this process is continued, it leads to

$$(A)(R) = P(a(b(c(\ldots (r(s(t(\ldots)))) \ldots)))) \tag{A8}$$

where P represents a summation over all permutations of $a, b, c, \ldots, r, s, t, \ldots$, such that the relative ordering within the two groups $(a, b, c, \ldots$ and $r, s, t, \ldots)$ is unchanged. In (A8) all right hand parentheses are at the far right, which means that the denominators are given in accordance with the rules for evaluating diagrams. The left hand side of (A8) represents a factorized product of the two parts of the diagram taken separately. Therefore, (A8) is a formulation of the factorization theorem, which is thereby proved.

References

Balian R and DeDominicis C 1971 *Ann. Phys., N.Y.* **62** 229
Barrett B R and Kirson M W 1973 *Adv. Nucl. Phys.* **6** ed M Baranger and Vogt (New York: Plenum) 219–85
Bloch C 1958 *Nucl. Phys.* **6** 329–47
Bloch C and Horowitz J 1958 *Nucl. Phys.* **8** 91–105
Brandow, B H 1967 *Rev. Mod. Phys.* **39** 771–828
Brueckner K A 1955 *Phys. Rev.* **100** 36–45
des Cloizeaux J 1960 *Nucl. Phys.* **20** 321–46
Eden R J and Francis N C 1955 *Phys. Rev.* **97** 1366–79
Feshbach H 1958 *Ann. Phys., N.Y.* **5** 357–90
—— 1962 *Ann. Phys., N.Y.* **19** 287–313
Feynman R P 1949 *Phys. Rev.* **76** 749–59
Frantz L M and Mills R L 1960 *Nucl. Phys.* **15** 16–33
Garpman S, Lindgren I, Lindgren J and Morrison J 1974 *Phys. Rev.* to be published
Goldstone J 1957 *Proc. R. Soc.* A **239** 267–79
Johnson M B and Baranger M 1971 *Ann. Phys., N.Y.* **62** 172–226
Judd B R 1967 *Second Quantization in Atomic Spectroscopy* (Baltimore: John Hopkins)
Kato T 1949 *Prog. Theor. Phys.* **4** 514–23
—— 1950 *Prog. Theor. Phys.* **5** 95–101
Kelly H P 1963 *Phys. Rev.* **131** 684–99
—— 1969 *Adv. Chem. Phys.* **14** 129–90
Kuo T T S, Lee S Y and Ratcliff K F 1971 *Nucl. Phys.* A **176** 65–88
Lindgren I *Comments on Atomic and Molecular Physics* to appear
Löwdin P O 1966 *Perturbation Theory and its Applications in Quantum Mechanics* ed C H Wilcox (New York: Wiley)
—— 1968 *Int. J. Quant. Chem.* **2** 867–931

Messiah A 1965 *Quantum Mechanics* (Amsterdam: North Holland) chap 16

Schucan T H and Weidenmüller H A 1972 *Ann. Phys., N.Y.* **73** 108–35

—— 1973 *Ann. Phys., N.Y.* **76** 483–509

Sandars P G H 1969 *Adv. Chem. Phys.* **14** 365–419

Møller C 1945 *K. Dan. Vid. Selsk.* **23** No 1

—— 1946 *K. Dan. Vid. Selsk.* **22** No 19

Hofmann H M, Lee S Y, Richert J, Weidenmüller H A and Schucan T H 1973 *Phys. Lett,* **45** B 421–4

—— 1974 *Ann. Phys., N.Y.* to appear