JOURNAL OF COMPUTATIONAL PHYSICS 109, 108–114 (1993)

On the Easy Generation of the Diagrams Representing the Coupled Cluster Method

S. R. HUGHES AND UZI KALDOR

School of Chemistry, Tel Aviv University, 69 978 Tel Aviv, Israel

Received July 29, 1992; revised February 1, 1993

A method is presented for the easy generation for all values of n of the Goldstone diagrams which represent the n body equation of the coupled cluster method. A static, finite system of fermions describable by a single-reference state is considered in full detail and in this case a simple polynomial expression is obtained for the number of diagrams contributing to the n body equation. This expression goes like $n^3/12$ for large n. Applications to other classes of system are also considered.

1. INTRODUCTION

This paper is concerned with the coupled cluster method (CCM) of quantum many-body theory. This, together with the extended coupled cluster method (ECCM), has by now achieved a high level of theoretical development and a wide range of applications in many areas of physics and chemistry. References [1, 2] are recent reviews of the subject.

As is well known, the CCM is based on the ansatz $|\Psi\rangle = \exp(S)|\Phi\rangle$, where $|\Psi\rangle$ is the true, interacting ground state of the system and S is the sum of a series of cluster operators S_n, which represent the amplitudes for the excitation of n-particle clusters out of some initially defined model state $|\Phi\rangle$. The value of n runs from one up to the number of particles in the system. $|\Phi\rangle$ is normally taken to be some state of the system for which all or part of the interaction has been turned off. For example, the problem of liquid ^{4}He is usually studied using a zero momentum Bose condensate as the model state. The CCM achieves a microscopic decomposition of the Schrödinger equation into a set of non-linear, coupled equations for the S_n. Once these have been solved, the S_n are substituted into an expression for the energy which the formalism also produces. The method can be established in a purely algebraic fashion, but it is also possible to interpret it completely in terms of the Goldstone diagrams of many-body perturbation theory. The CCM has two important advantages over rival methods. The first is that it displays the linked cluster theorem at all levels of truncation. This means that it generates no unlinked Goldstone diagrams and so it produces energies which scale like the number of particles. The second is that it actually has a well-defined truncation scheme in the first place, which is an essential requirement for practical applications. This is called the SUBn scheme and is defined by setting $S_m = 0$ for $m > n$. It sums well-defined, infinite sets of many-body perturbation theory (MBPT) Goldstone diagrams, which increase in size for increasing n. This eliminates the problem of knowing which is the next most important set of diagrams to sum. The working equations used for computations by the present authors are written and solved directly in terms of a *finite* set of CCM Goldstone diagrams. These feature symbols representing both the S_n amplitudes and the interaction potentials. It is possible in principle to perform a recursion on this set and so obtain the *infinite* set of MBPT Goldstone diagrams, which feature symbols only for the interaction potentials that a particular equation represents.

The process of deriving the CCM equations algebraically is straightforward but tedious and requires prohibitively large amounts of work even for relatively small values of n in the SUBn scheme. In such cases it is usual to fall back upon the diagrammatic description of the CCM. Because each term in the algebraic equations is associated with a CCM Goldstone diagram, it is possible to proceed via a two-stage process, viz.,

(1) Draw all permitted diagrams for the m body equation in SUBn for all values m and n of interest.

(2) Write down the corresponding algebraic equations and/or encode the diagrams directly into a computer program.

The second step is easy to automate given the proper software and at least one prescription [3] exists in the literature for the automation of the first step. The purpose of the present paper is to present another such prescription,

0021-9991/93 $5.00

one which has the virtue of great simplicity. Several other papers have appeared which are concerned with the automatic generation of diagrams, but these are confined to MBPT [4–7].

To summarise the remainder of the present paper, Section 2 gives a brief account of the conventional derivation of the CCM equations, followed by the alternative derivation of Arponen [8] which is the jumping off point of the diagrammatic analysis. Quantum chemical calculations are the main interest of the present authors, so in Section 3 we give in full the prescription for writing down the n-body equation for a static, single-reference state, finite fermionic system. A by-product of this is the derivation of a simple polynomial expression for $N(n)$, the number of CCM Goldstone diagrams on the RHS of the n-body equation. Section 4 indicates how classes of systems other than that in Section 3 may be treated. It also features remarks on the difficulty encountered in attempting to extend the present method to the ECCM. Section 5 is the conclusion.

2. DERIVATIONS OF THE CCM EQUATIONS

The usual starting point for the derivation of the CCM equations is the N-particle, time-independent Schrödinger equation

$$H|\Psi\rangle = E|\Psi\rangle \tag{1}$$

and the $\exp(S)$ ansatz

$$|\Psi\rangle = \exp(S)|\Phi\rangle \tag{2}$$

which was described in the Introduction. The operator S may be expressed in the rather generalised form

$$S = \sum_i S_i C_i^\dagger, \tag{3}$$

where the configuration creation operators $C_i^\dagger$ create an i particle configuration with respect to the model state. They are each composed of a string of single particle creation operators and the suffix i is not necessarily an integer but rather a shorthand notation for the set of particles created. The C_i configuration destruction operators are defined analogously. It is now assumed that all bra and ket states of the many-body Hilbert space can be reached by an appropriate superposition of the states $C_i^\dagger|\Phi\rangle$ and $\langle\Phi| C_i$. It is further assumed that $\langle\Phi| C_i C_j^\dagger|\Phi\rangle$ may be written as $\delta(i, j)$, where the latter is a notational kronecker delta which is to be understood in a set-theoretical sense and that some form of Wick's theorem exists for use on operators arranged in normal order.

Assuming that $\langle\Phi|\Phi\rangle$ is normalised to one we use (1) and (2) to obtain trivially

$$\langle\Phi| e^{-S} H e^{S}|\Phi\rangle = E \tag{4}$$

for the energy and

$$\langle\Phi| C_i e^{-S} H e^{S}|\Phi\rangle = 0 \tag{5}$$

for the S_i operators. Making use of the well-known commutator expansion

$$e^{-S} A e^{S} = A + [A, S] + \frac{1}{2!}[[A, S], S] + \cdots \tag{6}$$

which terminates after a finite number of terms for $A = H$, we can perform the requisite expansions and obtain algebraic equations for E and the S_i.

We now describe Arponen's alternative derivation [8] of the CCM equations. In his seminal work on the ECCM, he defines the expectation value of an arbitrary operator as

$$\mathcal{J}_A = \langle\Phi| \Omega e^{-S} A e^{S}|\Phi\rangle \tag{7}$$

and hence a new energy functional,

$$\mathcal{J}_H = \mathcal{J}_T + \mathcal{J}_V = \langle\Phi| \Omega e^{-S} H e^{S}|\Phi\rangle, \tag{8}$$

where $\mathcal{J}_T$ and $\mathcal{J}_V$ are the kinetic and potential energy expectation values, respectively, and

$$\Omega = 1 + \sum_i C_i \Omega_i. \tag{9}$$

Note that Ω is composed entirely of destruction operators with respect to the model state. Assuming that the kinetic energy operator is diagonal, it can be shown that

$$\mathcal{J}_T = \sum_i t_i S_i \Omega_i + T_0, \tag{10}$$

where t_i is the kinetic energy of the state $C_i^\dagger|\Phi\rangle$ relative to $|\Phi\rangle$ and T_0 is the kinetic energy of the model state. By requiring the stability of $\mathcal{J}_H$ against all variations of Ω_i and S_i we obtain

$$S_i = -\frac{1}{t_i} \cdot \frac{\delta \mathcal{J}_V}{\delta \Omega_i} \tag{11}$$

and

$$\Omega_i = -\frac{1}{t_i} \cdot \frac{\delta \mathcal{J}_V}{\delta S_i} \tag{12}$$

which are known as the Dyson equations because of their similar appearance to the equation of the same name from quantum field theory. It is now possible to use (8), (11), and (12) to re-derive the algbraic form of the CCM equations, provided that use is made of the result that the Ω_i make a zero contribution to the expectation value (7) in the case $A = H$.

3. DIAGRAMMATIC PRESCRIPTION

We confine ourselves in this section to a static, finite system of fermions describable by a single reference state $|\Phi\rangle$. Extensions to other classes of system are considered in Section 4.

We begin by making use of the fact [8] that $\mathscr{J}_T$ can be expressed as the sum of the Hugenholtz "skeleton" diagrams given in Fig. 1. The empty circles are one- and two-body interaction operators and the filled circles are the S_i amplitudes themselves. The boxes are the Ω_i amplitudes. The functional differentiation required by (11) now involves no more than the removal in all possible ways of the Ω_i boxes from each diagram. Thus, the single Ω_i amplitude appearing in each of the labelled diagrams in Fig. 1 is removed and the unlabelled diagrams are deleted altogether. Figure 2 shows the resultant general equation for S_i abbreviated in an obvious manner. A cross upon a line indicates the possibility of further lines emerging from the same amplitude. In some cases, for example, L, N, Q and the term on the RHS of Fig. 2, it is actually *necessary* to add further lines in order to obtain a legitimate contribution to the equation of interest. Not all diagrams contribute to all equations. For example, A contributes only to the one-body equation and E only to the two-body equation. All terms in an equation must have an equal, even number of open lines. This may be achieved by replacing a line marked by a cross with an appropriate number of lines. If this cannot be done for a diagram, then that diagram does not contribute.

The basic idea is now straightforward. First, all of the objects in Fig. 2 are expanded in order to produce a set of "nucleus" diagrams. Then, "accreted" diagrams are generated from each nucleus diagram. This is done for a particular nucleus diagram by first defining an "available" amplitude to be one with at least one line not linked to the interaction. Accreted diagrams are then generated by adding extra vertices in all possible distinct ways to the available amplitudes, such that the accreted diagrams have the requisite number of vertices for whatever n-body equation is currently of interest. For example, diagram F in Fig. 2 is expanded to produce the two nucleus diagrams in Fig. 3a, and these contribute to the two-body equation. By the addition of $n-2$ extra vertices, each of these two nucleus diagrams will generate a contribution to all higher n-body equations of exactly one accreted diagram, as illustrated in Figs. 3b and c for the cases $n = 3$ and $n = 7$.

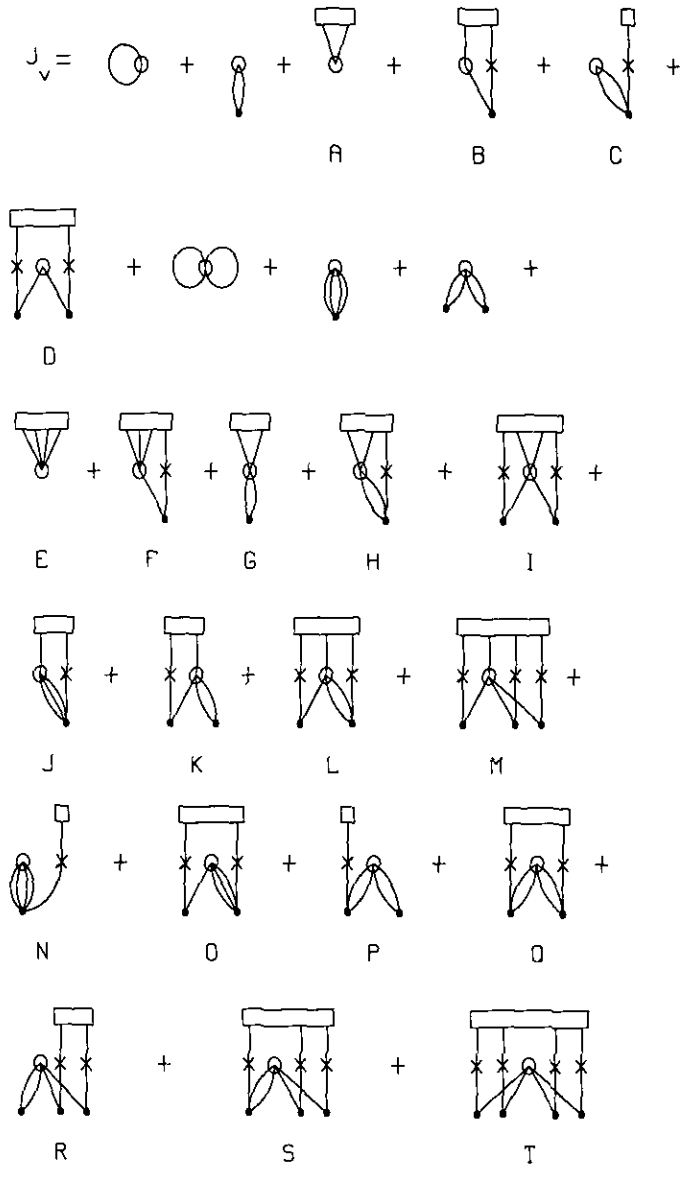

FIG. 1. The equation for $\mathscr{J}_v$ expressed in terms of Hugenholtz skeletons. A cross upon a line indicates the possibility and/or necessity of further lines emerging from the same amplitude. For further explanation of symbols, see text.

There is a further illustration of this process in Fig. 4. In Figs. 3–6, the open circle and the horizontal dashed line are one- and two-body interaction operators, respectively; the filled-in circle is now an S_1 amplitude; and a horizontal solid line from which n particle hole excitations emerge is an S_n

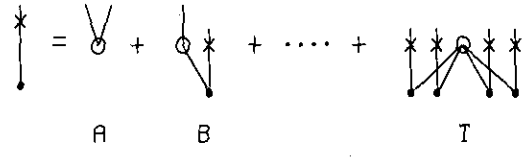

FIG. 2. The general equation for S_n obtained from Fig. 1 by functional differentiation with respect to Ω and written in an abbreviated form.

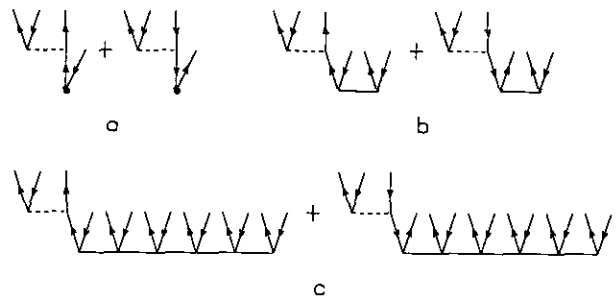

FIG. 3. Contribution of diagram F (Figs. 1 and 2) to: (a) the two-body equation; (b) the three-body equation; and (c) the seven-body equation.

amplitude. As higher and higher values of n are considered, the accreted diagrams will come to have the appearance of long homogeneous strings of vertices interrupted by occasional bursts of inhomogeneity resulting from the nucleus diagrams. There is no danger of overcounting, because the original Hugenholtz skeletons are all distinct and this cannot be altered by the subsequent addition of extra vertices. The full set of nucleus diagrams is given in Fig. 5. The nucleus and accreted diagrams form a pair of complementary subsets of the totality of CCM diagrams. The former is finite, having 74 members, and the latter is infinite.

The various n-body equations and their respective values of the function $N(n)$ are now considered in turn. $N(n)$ is defined as the number of diagrams on the RHS of the n-body equation. The one-body equation is a special case, being the only one to consist *entirely* of nucleus diagrams. There are 30 of these which can contribute, so

$$N(1) = 30. \tag{13}$$

The two-body equation is also a special case, being the only one to consist of *both* nucleus *and* accreted diagrams. There are 44 and 34 of these, respectively, and so

$$N(2) = 78. \tag{14}$$

Hereafter, *all* equations consist *entirely* of accreted

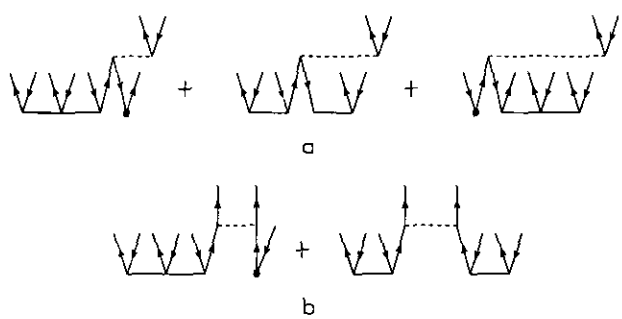

FIG. 4. Contribution of (a) diagram $I1$ and (b) diagram $I3$ (Fig. 5) to the four-body equation.

TABLE I

Numbers of Accreted Diagrams Contributing to the n-body Equation, $n \geqslant 3$

Number	Corresponding nucleus diagrams
0	A, E, $G1$, $G2$
1	$B1$, $B2$, $C1$, $C2$, $F1$, $F2$, $H1$–$H6$, $J1$–$J6$, $K1$–$K4$, $N1$–$N5$, $P1$–$P4$
n	D, $O1$, $O2$, $O4$, $O5$, $R1$, $R2$
$n-1$	$I1$, $I2$, $L1$–$L10$, $O3$, $O6$, $Q3$, $Q5$, $Q7$
$n/2$, n even; $(n-1)/2$, n odd	$I3$, $I4$, $Q1$, $Q2$, $Q4$, $Q6$
$n(n-1)/2$	$M1$, $M2$, $S1$–$S4$
$n^2/4$, n even; $(n^2-1)/4$, n odd	$S5$, $S6$
$n(n+2)(n-2)/12 + n/2$, n even; $n(n+1)(n-1)/12$, n odd	T

diagrams. In Table I are listed the numbers of these generated by each nucleus diagram for $n \geqslant 3$. The summation of these contributions is a trivial matter and the answer is

$$N(n) = (7n^2 + 48n + 28)/2 + [n/2 + n(n+2)(n-2)/12] \quad (n \geqslant 3 \text{ and even}), \tag{15a}$$

$$= (7n^2 + 48n + 21)/2 + [n(n+1)(n-1)/12] \quad (n \geqslant 3 \text{ and odd}). \tag{15b}$$

It now remains to prove that the various functions of n given in Table I are correct. The zero written against A, E, $G1$, and $G2$ indicates that they generate no contribution for $n \geqslant 3$. This is because they have no available amplitudes. As for the remainder, it is first necessary to consider the problem of the number of distinct arrangements of p additional vertices among q *distinguishable* amplitudes. Table II lists these results for $1 \leqslant q \leqslant 3$. The cases $q = 1$ and $q = 2$ are straightforward, but $q = 3$ is slightly more difficult. In this case, m of the p additional vertices are allocated to one particular amplitude and the other $p - m$ are shared between the other two amplitudes for $0 \leqslant m \leqslant p$. Therefore, employing the result for $q = 2$, the result for $q = 3$ is given by

$$\sum_{m=0}^{p} (p - m + 1) = (p+1)(p+2)/2. \tag{16}$$

TABLE II

Number of Distinct Ways of Adding p Vertices to q Distinguishable Amplitudes

q	Number
1	1
2	$p+1$
3	$(p+1)(p+2)/2$

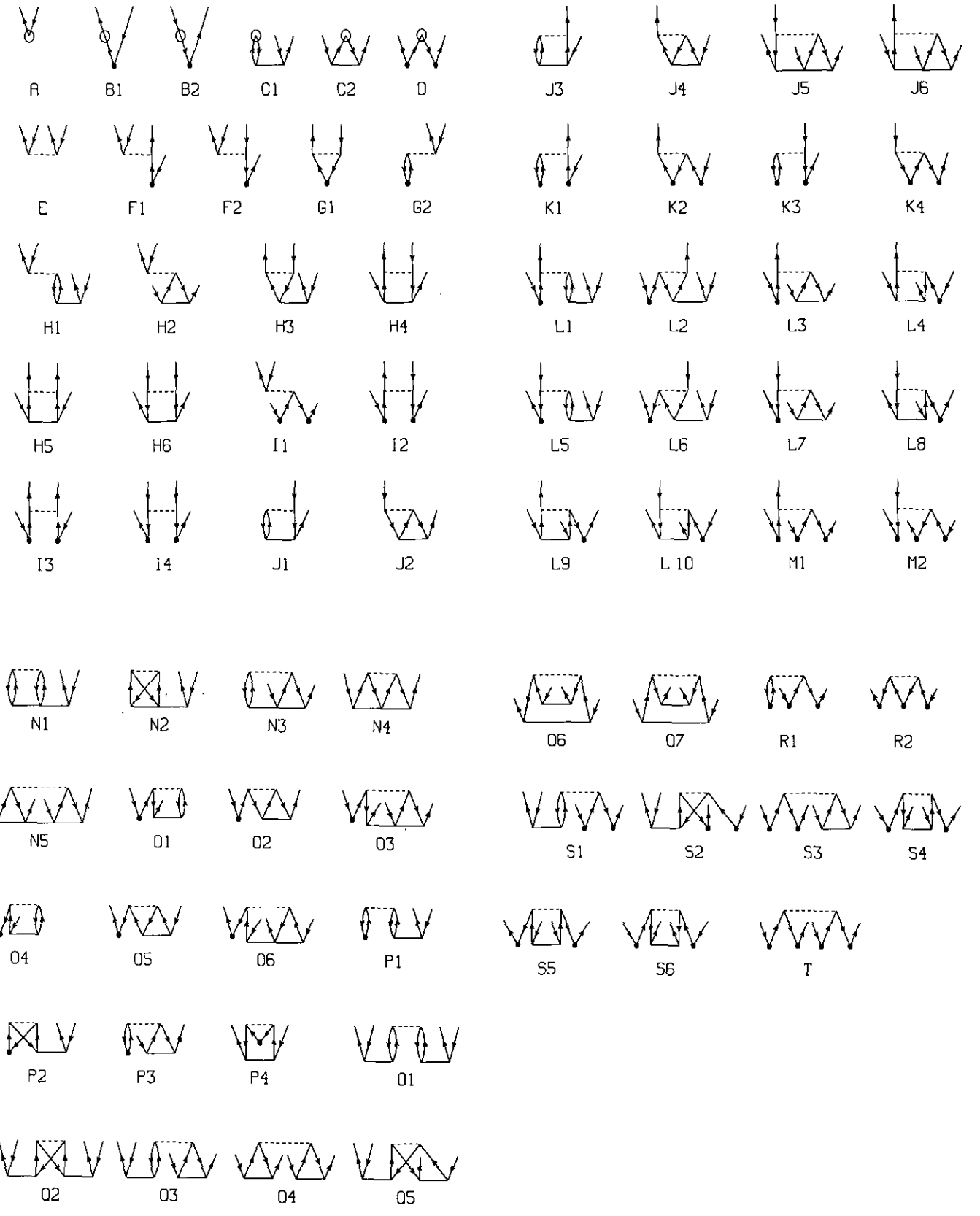

FIG. 5. Complete set of nucleus Goldstone diagrams for the CCM.

If the p additional vertices are added to two *indistinguishable* amplitudes then the number of distinct arrangements is given by $f_2(p)$, where

$$f_2(p) = \mathrm{Int}(p/2+1). \tag{17}$$

The only functions of n in Table I that are unaccounted for by Table II and Eq. (17) are those belonging to $S5$, $S6$, and T. The first two each have three available amplitudes, two of which are indistinguishable. Using (17), the number of distinct ways of adding the p additional vertices is therefore given by

$$\sum_{m=0}^{p} f_2(m) = (p+1)(p+3)/4, \qquad p \text{ even} \tag{18a}$$

$$= (p+2)^2/4, \qquad p \text{ odd}. \tag{18b}$$

Finally, there is nucleus diagram T, which is slightly tricky. Its contribution, $f_T(p)$, will first be stated and then explained,

$$f_T(p) = \sum_{m=0}^{p} f_2(m)(p-m+1) + \sum_{\substack{m=0 \\ m \text{ even}}}^{p} [-(p-m+1) + f_2(p-m)]. \tag{19}$$

Diagram T has four S_1 amplitudes, of which two are connected to the potential by particle lines and two by hole lines. These are dubbed the particle and the hole amplitudes, respectively. Of the p additional vertices, m are allocated to the particle amplitudes and $p-m$ to the hole amplitudes for all $0 \leqslant m \leqslant p$. The particle amplitudes are initially indistinguishable and so there are $f_2(m)$ distinct arrangements of their m allocated vertices. However, the addition of these vertices causes the hole amplitudes to become distinguishable in general. As a result, this $f_2(m)$ must be multiplied by a factor of $(p-m+1)$ for the distinct arrangements of the $p-m$ vertices allocated to the hole amplitudes, see Table I. This explains the first term in (19). The tricky part is caused by the cases when m is even, in which there is one arrangement of the m vertices allocated to the particle amplitude such that the hole amplitudes become *indistinguishable*. It is then necessary to subtract off one of the terms $(p-m+1)$, which is appropriate to distinguishable amplitudes, and replace it by a term $f_2(p-m)$, which is appropriate to indistinguishable amplitudes. This is illustrated by Fig. 6, which is the contribution of T to the five-body equation. The summations on the RHS of (19) can easily be executed.

It is important to realise that the expressions in Table II and Eqs. (17)–(19) are given in terms of p, the number of vertices to be *added* to a given nucleus diagram. However,

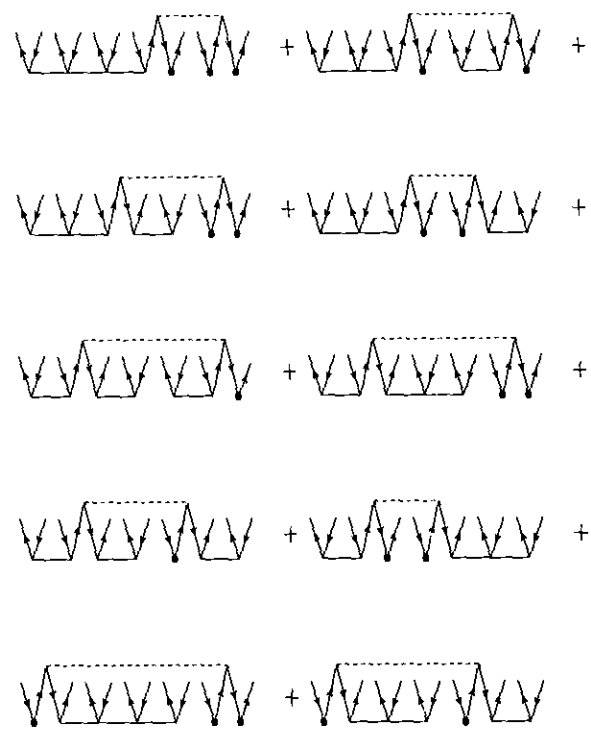

FIG. 6. Contribution of diagram T (Fig. 5) to the five-body equation.

the nucleus diagrams all have either one or two pairs of uncontracted lines to begin with. Therefore, when evaluating expressions for the n-body equation, p must be substituted in all cases with $n-1$ or $n-2$ as appropriate.

It should be noted that the expression $N(n)$ for the number of diagrams contributing to the n-body equation goes like $n^3/12$ for large n, and that the $n^3/12$ contribution is generated solely by nucleus diagram T.

4. FURTHER EXTENSIONS

In view of Eqs. (7) and (12) and Fig. 1, it might be thought that the derivation of a set of diagrammatic equations for the Ω_i should be attempted, thus enabling the calculation of the expectation values of an arbitrary operator. However, it is well known that the equations for the Ω_i are unlinked in general and that this problem is avoided only when the energy expectation value is calculated, in which case it can be shown that, as already stated, the Ω_i make a zero contribution. In particular, the removal of the S amplitude of Fig. 1N at once yields an unlinked diagram for Ω. This problem is solved by the ECCM [8] in which the diagrammatic expression for $\mathscr{J}_\nu$ is obtained by resolving the Ω_i boxes into a set of upper, de-excitation amplitudes and preserving only those diagrams which satisfy the so-called double linking requirements, viz.,

(1) From each lower, excitation amplitude (now called σ_i) at least one line is connected to the interaction.

(2) Every upper, de-excitation amplitude (now called $\tilde{\sigma}_i$) must be connected to the interaction and/or have connections to at least two σ_i amplitudes.

This prescription ensures a properly linked description of an arbitrary operator. The ECCM Dyson equations, cf. (11) and (12),

$$\sigma_i = -\frac{1}{t_i} \cdot \frac{\delta \mathscr{J}_V}{\delta \tilde{\sigma}_i} \tag{20}$$

and

$$\tilde{\sigma}_i = -\frac{1}{t_i} \cdot \frac{\delta \mathscr{J}_V}{\delta \sigma_i} \tag{21}$$

can now be used in principle to obtain equations for σ_i and $\tilde{\sigma}_i$ written in terms of ECCM Goldstone diagrams. In practice, however, there are considerable difficulties impeding an analysis of the ECCM, analogous to that of Section 3. If there is only one σ or one $\tilde{\sigma}$ amplitude in a given ECCM nucleus diagram, then the functional differentiation reduces the diagram to one which is of similar difficulty to the "normal" CCM (NCCM). However, if this is not the case then the combinatorics of the lines linking the multiple σ_i and $\tilde{\sigma}_i$ amplitudes appear to preclude such an easy scheme as has been obtained for the NCCM. Furthermore, the ECCM analogue of Fig. 2T causes violent proliferation of diagrams even for so crude an approximation as CCD, that is, in which $S_n = 0$ for all $n \neq 2$.

The prescription given in Section 3 is designed for a static, finite, single-reference state fermionic system. However, it is very easy to use these diagrams to obtain those for an open shell system, which will in general require multiple reference states. If there are m particles added to the closed shell, then bend down m hole lines in all possible ways for each diagram in order to form valence particle lines. A similar rule holds for particles deleted from a closed shell state. It then remains only to draw the so-called folded diagrams which are few in number and easy to write down by inspection. For further details, see Ref. [9]. The dynamic case may also be treated and Arponen gives rules for doing this.

Application to a system of bosons in the thermodynamic limit is also possible. Essentially, this involves the deletion of all diagrams featuring an S_1 amplitude (because $S_1 \neq 0$ violates momentum conservation in this case) and also all diagrams which can be obtained from other diagrams by an exchange of hole lines (because all holes are equivalent and so the same physical process is described). Although for small n it is well known that $N_{\text{fermi}}(n) \gg N_{\text{bose}}(n)$, for large n the same $n^3/12$ limit is recovered. The culprit is once again diagram T of Fig. 6. Although in the bosonic case it makes its first appearance only in the six-body equation, thereafter the pattern of additional vertices is the same as for the fermionic equations, except that it starts four equations later in the hierarchy.

5. CONCLUSION

A method has been presented for the easy automatic generation of the diagrams representing the equations for *all* the CCM cluster amplitudes S_n in the case of static, single-reference state fermionic systems. This has been achieved by defining a set of "nucleus" diagrams from which all other diagrams can be obtained in a straightforward manner by means of the addition of extra vertices to the cluster amplitudes. It will readily be seen that a program for the automatic generation of diagrams will need only to store the nucleus diagrams and thereafter consist of at most a few simple nested DO loops for each diagram. A simple polynomial expression for $N(n)$, the number of diagrams on the RHS of the n-body equation, has also been obtained and this goes like $n^3/12$ for large N, this contribution coming entirely from just one of the nucleus diagrams. Application of the same method to other classes of systems has been sketched. In particular, we have obtained the perhaps rather surprising result that for large n, $N(n)$ has the same limit for a thermodynamic system of bosons as for a finite system of fermions.

ACKNOWLEDGMENTS

The research reported here was supported by the U.S.–Israel Binational Science Foundation. S.R.H. thanks the Committee for Planning and Budgeting of the Israeli Council on Higher Education for a fellowship.

REFERENCES

1. R. F. Bishop, *Theor. Chim. Acta* **80**, 95 (1991).
2. R. J. Bartlett, *J. Phys. Chem.* **93**, 1697 (1989).
3. S. A. Kucharski and R. J. Bartlett, *Adv. Quantum. Chem.* **18**, 281 (1986).
4. J. Paldus and H. C. Wong, *Comput. Phys. Commun.* **6**, 1 (1973); H. C. Wong and J. Paldus, *Comput. Phys. Commun.* **6**, 9 (1973).
5. U. Kaldor, *J. Comput. Phys.* **20**, 432 (1976).
6. V. Kvasnicka, V. Laurinc, and S. Biskupic, *Mol. Phys.* **39**, 143 (1980).
7. Z. Csépes and J. Pipek, *J. Comput. Phys.* **77**, 1 (1988).
8. J. S. Arponen, *Ann. Phys. (N.Y.)* **151**, 311 (1983).
9. I. Lindgren and J. Morrison, *Atomic Many-Body Theory* (Springer, Berlin, 1982).