Progress of Theoretical Physics, Vol. 40, No. 4, October 1968

A Systematic Approach to the Problems of Random Lattices. I

——A Self-Contained First-Order Approximation Taking into Account the Exclusion Effect——)

Fumiko YONEZAWA

Research Institute for Fundamental Physics
Kyoto University, Kyoto

(Received June 3, 1968)

A systematic method of approximation for the electronic state of a randomly doped lattice or a vibration spectrum of a disordered lattice is given in the present series of work by means of an investigation of the one-electron Green's function. In the present article, an exact form of the first-order self-energy is, with the help of a diagrammatic consideration, evaluated on rigorously including the "exclusion effect". The resulting "exact" first-order self-energy agrees with the lowest-order approximant of the "total first-order self-energy which has been previously obtained by the author and Matsubara, and satisfies the same equation derived by Taylor for the case of lattice vibrations. It is also identical with the approximation developed by Onodera and Toyozawa. Thus, one of the objects of the analysis given in the present work is to offer a mathematically correct interpretation of these methods. A systematic way to proceed to higher-order approximations is discussed.

§ 1. Introduction

The problem of evaluating the electronic structure of random alloys or liquid metals, or of calculating the frequency spectra of disordered lattices has long attracted a considerable amount of interest; and yet no systematic techniques to treat this problem have been established by now. Most of the attempts developed so far[1] are dependent on mathematical procedures characteristic of one dimension, and it is often very difficult to extend these one-dimensional models to more realistic three dimensional cases.

New approaches to the problem by means of the Green's function method[2] or the multiple-scattering theory have recently been developed by several authors. The method is advantageous in that the formulations are applicable to models of any dimension. Another merit of the Green's function method is that it offers insight into the actual dynamics of electrons or of the lattice.

In spite of the fact that the Green's function method excels in these points, techniques based upon the method are not regarded to be well-established because of some difficulties contained in the obtained results. For instance, attempted approaches fail to explain the fine structure of the spectrum which has been

*) A preliminary report of this work has been published in Prog. Theor. Phys. **39** (1968), 1076.

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

proved to appear by the machine calculation of Dean and his co-workers[3] for one and two dimensions and, more recently, by Payton and Visscher[4] for one, two and three dimensions. Alternatively, the formulation along this line has not been successful in telling something about the special frequencies which have been analytically investigated by Matsuda, Hori, etc.[5]

Another catastrophe of the Green's function method takes place when the "excluded volume" corrections are treated carelessly. On one hand, complete neglect of the exclusion effect leads to the breakdown of the host-defect dual symmetry (or host-impurity dual symmetry). On the other hand, the exclusion effect when incorrectly taken into account, even though it may fulfill the required symmetry, introduces spurious poles into the self-energy, and the convergence of the self-energy is not ensured.

These two difficulties are actually related; both originate in the wrong guiding principle in approximating the iterative type of solution.

An object of the present work is to provide a systematic way to overcome these difficulties of the Green's function method. In § 2, an explanation is given for the term "the exclusion effect". When carrying out the average over all the random configurations of impurity atoms, it is necessary to take care that no more than one impurity atom can be on a given lattice site. This means that all impurity-site summation must be partitioned or decoupled in such a manner that none of the summation indices can be the same. This partitioning procedure gives rise to the restriction over the summation indices. When the distribution of impurity atoms are completely random, averaging is effected by the replacement of the summation over all impurity-sites by the product of the impurity concentration c and the summation over all the lattice sites. The restrictions over the summation indices are preserved throughout this averaging procedure and, after averaging, must be removed by some way. In § 3, a basic procedure for taking account of the exclusion effect is presented for later use. It is emphasized in § 4 that the removement of the restrictions over the summation indices must be performed such that the self-containedness is ensured in the limits of the given stage of approximation. Diagrammatic consideration facilitates the self-contained treatment of the exclusion effect. The solution attained in this manner is verified to be free from the above mentioned difficulties of the Green's function method.

The general principle for the proceeding approximation is described in § 5. In the last section, some comment is made on an appropriate approximation for a given concentration c. The fine structure present in the energy spectra may be obtained if a valid approximation for a given c is used.

For simplicity of explanation, the formulation is carried out, throughout this paper, for electrons moving in the aperiodic potential field due to the randomly-distributed impurity atoms. Potential of the individual impurity atom is additive and comprises no internal freedom. A one-electron model is employed.

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

Although the discussion is limited to the restricted model described in the above, the method developed herein is widely applicable to general models.

§ 2. The exclusion effect

Let us consider a crystal with many substitutional impurity atoms where the concentration of the impurity atoms is c. Possibilities of multiple occupancy of a lattice point or occurence of unoccupied lattice site are avoided as a proviso.

One of the cruces in the problem of random systems is the way how the effect of "random" configurations should be treated. The situation becomes difficult, especially when the correlation between impurity atoms is important. If it is permitted to neglect the impurity-impurity correlation, so that the impurity atoms are regarded to be distributed over the lattice sites in a completely random manner, every possible configuration occurs with an equal weight. In this case, the averaging procedure over all possible configurations is effected by an easier operation as

$$\left\langle \sum_{\{l_1,\cdots,l_j\}}{}' F(\boldsymbol{R}_{l_1}, \boldsymbol{R}_{l_2},\cdots, \boldsymbol{R}_{l_j})\right\rangle = c^j \sum_{n_1,\cdots,n_j}{}' F(\boldsymbol{R}_{n_1}, \boldsymbol{R}_{n_2}, \cdots, \boldsymbol{R}_{n_j}), \tag{2·1}$$

where $F(\boldsymbol{R}_{l_1}, \boldsymbol{R}_{l_2}, \cdots, \boldsymbol{R}_{l_j})$ is an arbitrary function of j impurity atoms at $\boldsymbol{R}_{l_1}$, $\boldsymbol{R}_{l_2}$, $\cdots$, $\boldsymbol{R}_{l_j}$ respectively; the brackets $\langle\ \rangle$ indicate the configuration average; $\sum'_{\{l_1 l_2 \cdots l_j\}}$ designates the summation over all the impurity sites while $\sum'_{n_1,n_2,\cdots,n_j}$ means the summation over all the lattice points. The prime on each summation implies that none of the summation indices can be the same.

The exclusion effect originally comes from this restriction over the summation indices. The total Green's function is written in terms of the multiple scattering formulation as follows;

$$G = G_0 + \sum_{\alpha} G_0 T_\alpha G_0 + \sum_{\alpha \neq \beta} G_0 T_\alpha G_0 T_\beta G_0 + \sum_{\alpha \neq \beta \neq \gamma} G_0 T_\alpha G_0 T_\beta G_0 T_\gamma G_0 + \cdots, \tag{2·2}$$

where T_α is a scattering matrix. The restrictions over the summation indices appearing in each term of the right-hand member of Eq. (2·2) is such that every pair of succeeding summation indices are different from each other. In the fourth term for instance, the restriction indicates that $\alpha \neq \beta$ and $\beta \neq \gamma$ while $\gamma \neq \alpha$ is not implied. It is necessary to take the average of Eq. (2·2) over all the possible configurations, in order to obtain a closed form of the solution of the Dyson equation,

$$\langle G \rangle = G_0 + G_0 \sum \langle G \rangle, \tag{2·3}$$

in which notation $\sum$ designates the sum of all proper self-energy parts. Averaging is carried out on the principle stated by Eq. (2·1). For this purpose, terms from and after the fourth term of Eq. (2·2) must be so partitioned that all the summation indices are different from one another. On taking the fourth term as an example, the cases $\alpha \neq \gamma$ and $\alpha = \gamma$ are to be distinguished. After averaging,

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

the restriction over the summation indices need be removed in some manner to enable the Dyson equation to be solved. This was first pointed by Langer.[3] A general technique for complicated removement procedure has previously been given (Yonezawa and Matsubara,[6] hereafter to be referred to as YMII.) Let us first reformulate this procedure from a slightly different point of view with emphasis upon the "excluded volume" corrections. In the course of formulation, partial repetition of the theoretical routine in YMII is intended exclusively for the convenience of later development.

§ 3. A general method of taking account of the exclusion effect

3·1 *Formulation and the second-order moment*

The averaged one-electron Green's function in a random system is written in an iterative expression;

$$\langle G_{\boldsymbol{k}\boldsymbol{k}'}\rangle = G_0(\boldsymbol{k})\,\delta(\boldsymbol{k}-\boldsymbol{k}') + G_0(\boldsymbol{k})\,M_1(\boldsymbol{k}-\boldsymbol{k}')\,V_{\boldsymbol{k}\boldsymbol{k}'}G_0(\boldsymbol{k}') + G_0(\boldsymbol{k})\sum_{\boldsymbol{k}_1} M_2(\boldsymbol{k}-\boldsymbol{k}_1\,;\,\boldsymbol{k}_1-\boldsymbol{k}')\,V_{\boldsymbol{k}\boldsymbol{k}_1}V_{\boldsymbol{k}_1\boldsymbol{k}'}G_0(\boldsymbol{k}_1)G_0(\boldsymbol{k}') + \cdots, \tag{3·1}$$

where

$$M_s(\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_s) = \Big\langle \prod_{j=1}^{s} \sum_{\{l_j\}} \exp(-i\boldsymbol{p}_j\cdot\boldsymbol{R}_{l_j})\Big\rangle. \tag{3·2}$$

In order that the closed solution may be obtained from Eq. (3·1), the moments $M_s(\boldsymbol{p}_1, \cdots, \boldsymbol{p}_s)$ should be written in more compact forms. Before evaluating $M_s(\boldsymbol{p}_1, \cdots, \boldsymbol{p}_s)$ for an arbitrary s, it is instructive to show how the moments for small s are derived in a primitive way. The first-order moment is easily calculated;

$$\begin{aligned} M_1(\boldsymbol{p}) &= \Big\langle \sum_{\{l\}} \exp(-i\boldsymbol{p}\cdot\boldsymbol{R}_l)\Big\rangle = c\sum_n \exp(-i\boldsymbol{p}\cdot\boldsymbol{R}_n) \\ &= Nc\delta(\boldsymbol{p}) \equiv C_1(p), \end{aligned} \tag{3·3}$$

in which $C_1(\boldsymbol{p})$ designates the first-order cumulant of $\rho(\boldsymbol{p}) = \sum_{\{l\}} \exp[-i\boldsymbol{p}\cdot\boldsymbol{R}_l]$. Next, consider the second-order moment which by definition is written as

$$M_2(\boldsymbol{p}_1, \boldsymbol{p}_2) = \Big\langle \sum_{\{l_1\}} \sum_{\{l_2\}} \exp[-i\boldsymbol{p}_1\cdot\boldsymbol{R}_{l_1} - i\boldsymbol{p}_2\cdot\boldsymbol{R}_{l_2}]\Big\rangle. \tag{3·4}$$

As explained in some detail in the preceding section, two cases $l_1 = l_2$ and $l_1 \neq l_2$ must be distinguished during the averaging operation and this necessitate the partitioning the summation into two terms such as;

$$\begin{aligned} M_2(\boldsymbol{p}_1, \boldsymbol{p}_2) &= \Big\langle \sum_{\{l\}} \exp[-i(\boldsymbol{p}_1+\boldsymbol{p}_2)\cdot\boldsymbol{R}_l]\Big\rangle + \Big\langle \sum_{\{l_1\}\neq\{l_2\}} \sum \exp[-i\boldsymbol{p}_1\cdot\boldsymbol{R}_{l_1} \\ &\quad - i\boldsymbol{p}_2\cdot\boldsymbol{R}_{l_2}]\Big\rangle = c\sum_n \exp[-i(\boldsymbol{p}_1+\boldsymbol{p}_2)\cdot R_n] + c^2 \sum_{n_1\neq n_2} \exp[-i(\boldsymbol{p}_1\cdot\boldsymbol{R}_{n_1} \\ &\quad + \boldsymbol{p}_2\cdot\boldsymbol{R}_{n_2})], \end{aligned} \tag{3·5}$$

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

where the configuration average has been taken according to Eq. (2·1). In performing the second sum, the restriction over the summation indices is removed by adding and then subtracting the terms in which $n_1=n_2$, thus

$$\begin{aligned} M_2(\boldsymbol{p}_1, \boldsymbol{p}_2) &= Nc\delta(\boldsymbol{p}_1+\boldsymbol{p}_2) + c^2\{\sum_{n_1}\exp(-i\boldsymbol{p}_1\cdot\boldsymbol{R}_{n_1})\sum_{n_2}\exp(-i\boldsymbol{p}_2\cdot\boldsymbol{R}_{n_2}) \\ &\quad -\sum_{n}\exp[-i(\boldsymbol{p}_1+\boldsymbol{p}_2)\cdot\boldsymbol{R}_n]\} \\ &= Nc\delta(\boldsymbol{p}_1+\boldsymbol{p}_2) + c^2\{N\delta(\boldsymbol{p}_1)\cdot N\delta(\boldsymbol{p}_2) - N\delta(\boldsymbol{p}_1+\boldsymbol{p}_2)\}. \end{aligned} \tag{3·6}$$

Rearrangement of the terms according to the powers of N and to the patterns of delta functions yields

$$\begin{aligned} M_2(\boldsymbol{p}_1, \boldsymbol{p}_2) &= N(c-c^2)\delta(\boldsymbol{p}_1+\boldsymbol{p}_2) + Nc\delta(\boldsymbol{p}_1)\cdot Nc\delta(\boldsymbol{p}_2) \\ &\equiv C_2(\boldsymbol{p}_1, \boldsymbol{p}_2) + C_1(\boldsymbol{p}_1)C_1(\boldsymbol{p}_2), \end{aligned} \tag{3·7}$$

where $C_2(\boldsymbol{p}_1, \boldsymbol{p}_2)$ indicates the second-order cumulant of $\rho(\boldsymbol{p})$. Thus the exclusion effect in this case is interpreted to be a correction of order c^2 to the coefficient of $\delta(\boldsymbol{p}_1+\boldsymbol{p}_2)$, and $C_2(\boldsymbol{p}_1, \boldsymbol{p}_2)$ is defined by subtracting this correction from c.

3·2 *A general procedure*

Similar steps are required to get every higher-order moment $M_s(\boldsymbol{p}_1, \cdots, \boldsymbol{p}_s)$ by;

(1) first partitioning all impurity-site summations in $M_s(\boldsymbol{p}_1, \cdots, \boldsymbol{p}_s)$ according to the multiplicity of overlap among the summation indices $\boldsymbol{R}_{l_1}, \cdots, \boldsymbol{R}_{l_s}$, followed by the configuration average as shown by Eq. (2·1);

(2) removing the restrictions over the summation indices by addition and then subtraction of appropriate quantities; and

(3) lastly, reclassifying the resulting terms according to the powers of N and to the pattern of the delta functions.

3·3 *Interpretation by means of diagrams*

Diagrammatic consideration makes it easier to understand the mentioned complicated procedures; this is possible because the terms on the right-hand members of Eqs. (3·5) through (3·7) and the similar terms of higher-order moments are subject to a direct interpretation by means of diagrams.

The prescriptions for representing moments and cumulants in terms of diagrams are summarized as follows;

With reference to Figs. 1 and 2, a horizontal single line represents the unperturbed propagator or Green's function $G_0(\boldsymbol{k})$ of an electron while a double line designates the perturbed one-electron Green's function $G(\boldsymbol{k})$. Interaction or scattering by an impurity atom at $\boldsymbol{R}_n$ is shown by a dotted line connecting the propagator and a vertex n; to this interaction line, $V_{\boldsymbol{k}\boldsymbol{k}'}$ is assigned. A

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

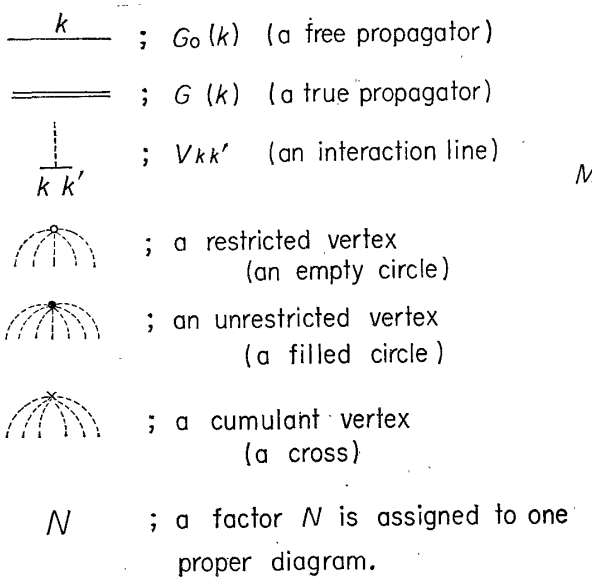

Fig. 1. Prescription for counting diagrams.

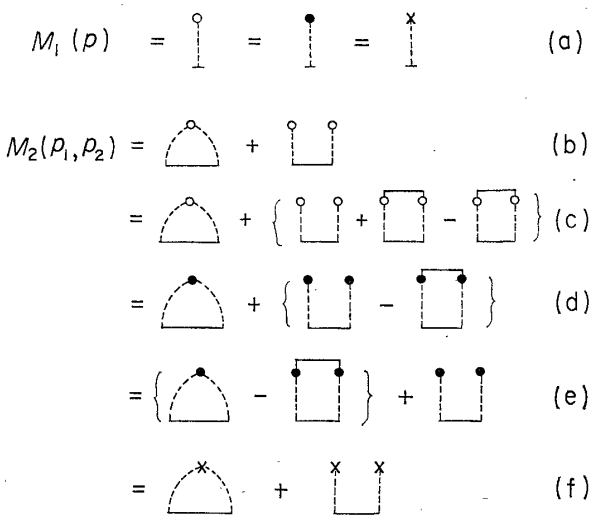

Fig. 2. Diagrammatic interpretation of the processes of evaluating $M_1(p)$ and $M_2(p_1, p_2)$.

vertex n may be an empty or filled circle, or a cross acccording to three different types of diagrams. First, a diagram which contains empty circles is called a restricted diagram; in this sort of diagram, the summation indices represented by the empty circles carry the same restriction as the topological pattern of the diagram shows. In Fig. 2(b), the first two diagrams with empty circles are the restricted ones; these diagrams in this order correspond to the first and second terms in Eq. (3·5), respectively. In the next place, a diagram with filled circles is called an unrestricted diagram in the sense that the summation indices corresponding to filled circles are not restricted. As a result, the summation over each summation index is carried out independently; a factor c is assigned to each filled circle and a factor N to each proper part. Finally, a diagram with a cross is called a renormalized diagram or cumulant diagram; an sth-order cumulant $C_s(\boldsymbol{p}_1, \cdots, \boldsymbol{p}_s)$ is assigned to a cross from which s interaction lines start.

A process of obtaining $M_1(\boldsymbol{p})$ is depicted in Fig. 2(a). It must be noted that, when considering a diagram containing one circle alone, there is no need of distinguishing between the three types of diagrams, since in this case sum is taken over 'one' summation index alone. Similarly, this is the case when a diagram contains more than one circle so connected together by a solid line that the connected diagram is topologically equivalent to the former diagram. In Figs. 2(b) through (f), the complicated procedure of evaluating $M_2(\boldsymbol{p}_1, \boldsymbol{p}_2)$ is shown. Figure 2(b) represents the partitioning and averaging. The removement of the restriction over the summation indices is expressed by Fig. 2(c), where a term $c^2 \exp[-i(\boldsymbol{p}_1+\boldsymbol{p}_2)\cdot\boldsymbol{R}_n]$ is added and then subtracted. A diagram corresponding to $c^2 \exp[-i(\boldsymbol{p}_1+\boldsymbol{p}_2)\cdot\boldsymbol{R}_n]$ in Figs. 2(c) and 2(d) comprises a solid line connecting more than one circle; this diagram is introduced for the purpose

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

of removing the restricted sum. Consequently, the summation indices represented by the the filled circles in Fig. 2(d) and figures thereafter become independent of one another. The obtained diagrams are rearranged as shown in Fig. 2(e). The definition of $C_2(\boldsymbol{p}_1, \boldsymbol{p}_2)$ is apparent from Figs. 2(e) and (f).

When the exclusion effect is completely neglected, the contribution from the second term in Fig. 2(e) is not taken into account, so that only a factor c is assigned to a crossed vertex instead of a cumulant $C_2(\boldsymbol{p}_1, \boldsymbol{p}_2)$.

In a similar manner, the three steps for $M_3(\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3)$ are shown in Figs. 3(a), (b) and (c), and $M_3(\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3)$ is obtained as shown in Fig. 3(d) in terms of a sum of cumulant diagrams. The forth-order cumulant $C_4(\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3, \boldsymbol{p}_4)$ is determined through the evaluation procedure of $M_4(\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3, \boldsymbol{p}_4)$ as given by Fig. 4.

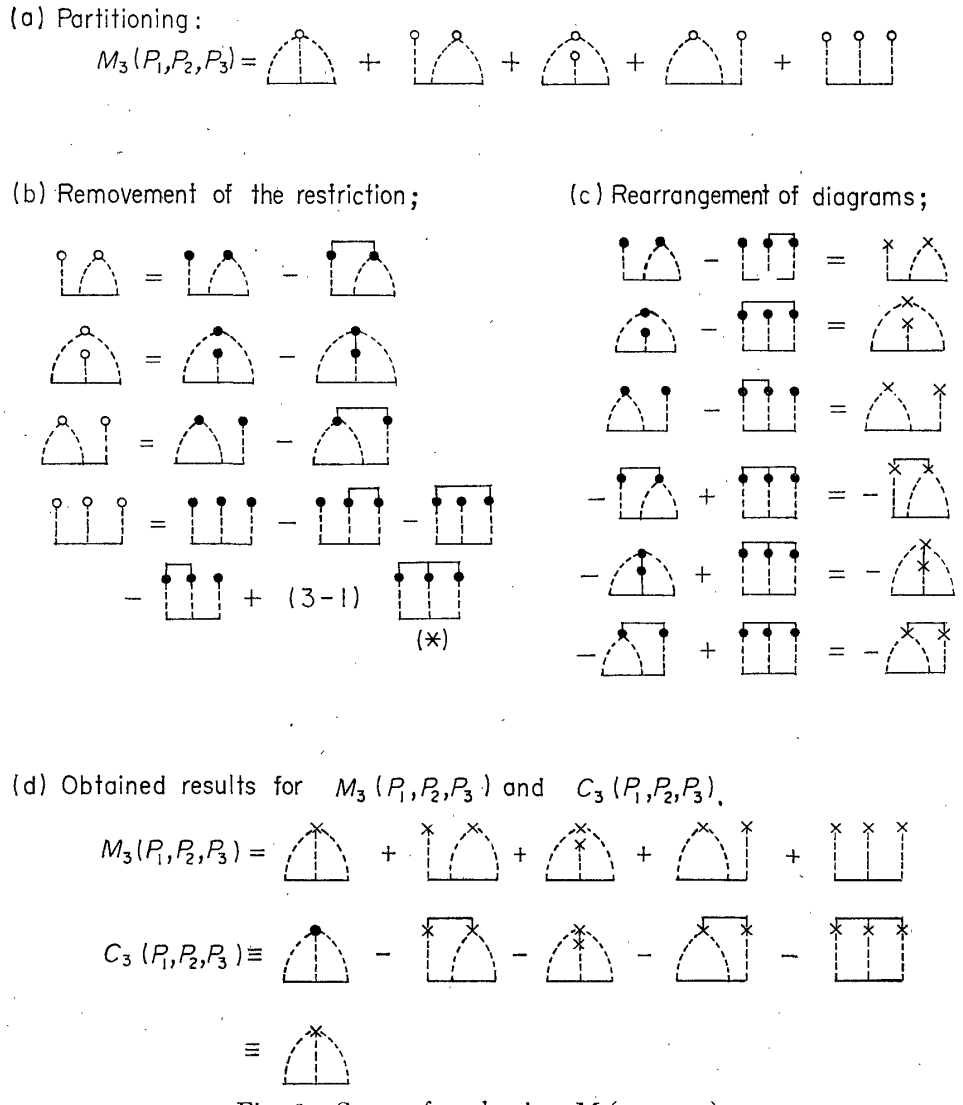

Fig. 3. Steps of evaluating $M_3(p_1, p_2, p_3)$.

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

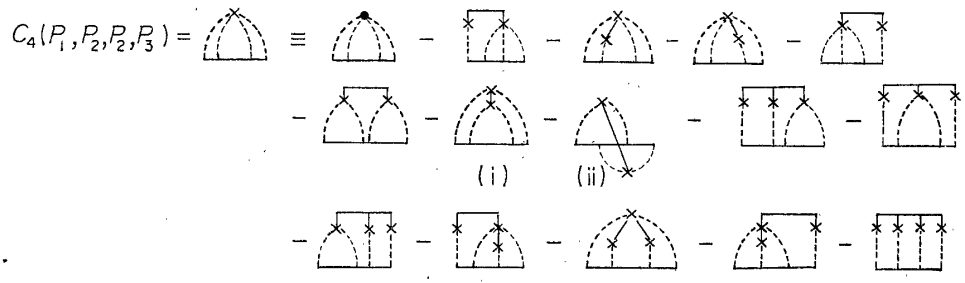

Fig. 4. Forth-order cumulant.

3·4 *Cumulant averages of general order*

Let us next consider how the cumulant average of order s is generally derived. As is easily predicted, an sth-order cumulant is of the form;

$$C_s(\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_s) = NP_s(c)\,\delta(\boldsymbol{p}_1+\boldsymbol{p}_2+\cdots+\boldsymbol{p}_s). \tag{3·8}$$

The coefficients $P_s(c)$ for $s\leq 4$ are obtained by counting diagrams in Figs. 2 and 3 according to the above prescription;

$$P_1(c) = c, \tag{3·9a}$$

$$P_2(c) = c - P_1(c)\,P_1(c) = c - c^2, \tag{3·9b}$$

$$P_3(c) = c - 3P_1(c)\,P_2(c) - (P_1(c))^3 = c - 3c^2 + 2c^3, \tag{3·9c}$$

$$\begin{aligned} P_4(c) &= c - 4P_1(c)\,P_3(c) - 3P_2(c)\,P_2(c) - 6P_1(c)\,P_1(c)\,P_2(c) - (P_1(c))^4 \\ &= c - 7c^2 + 12c^3 - 6c^4. \end{aligned} \tag{3·9d}$$

The consideration for $P_s(c)$ with small s suggests that $P_s(c)$ for an arbitrary s is determined as

$$\begin{aligned} P_s(c) &= c\,(\equiv\text{a contribution from a diagram with a filled circle from which } s \text{ interaction lines start.}) - (\text{correction factors coming from those diagrams which are topologically equal to the diagram of the first term when connected by a solid line.}) \\ &= c - \sum_{(2)} P_{m_1}(c)\,P_{s-m_1}(c) - \sum_{(3)} P_{m_1}(c)\,P_{m_2}(c) \\ &\quad \times P_{s-m_1-m_2}(c) - \cdots \sum_{(s-1)} P_2(c)\,(P_1(c))^{s-2} - (P_1(c))^s, \end{aligned} \tag{3·10}$$

where $\sum_{(l)}$ indicates all partitions of $p_1, p_2, \cdots, p_s$ into l groups. It is concluded from Eq. (3·10) that $P_s(c)$ is a polynomial of order n in c.

It has been shown in YMII that $P_s(c)$ is systematically obtained from a generating function;

$$g(x:c) = \log(1-c+ce^x) = \sum_{s=1}^{\infty} P_s(c)\,x^s/s! \tag{3·11}$$

3·5 *First-Order self-energy*

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

On summarizing, an exact self-energy $\boldsymbol{\Sigma}$ satisfying the Dyson equation (2·3) is derived by:

(1) first drawing all possible proper cumulant diagrams; and

(2) counting the contributions from these diagrams by means of the prescriptions as shown in Fig. 1.

Practically, however, summing all possible diagrams is impossible, so that the most important diagrams are selected by some criteria. For instance, if the repeated scattering with one impurity atom is most important and it is permitted to regard only those diagrams shown in Fig. 5 are summed as a first-order self-energy $\boldsymbol{\Sigma}(1)$. The result is

$$\boldsymbol{\Sigma}(1)=N\boldsymbol{V}\sum_{s=1}^{\infty}P_s(c)\,(\boldsymbol{G}\boldsymbol{V})^{s-1}, \tag{3·12}$$

where $\boldsymbol{\Sigma}$, $\boldsymbol{G}$ and $\boldsymbol{V}$ designates matrices whose $\boldsymbol{kk'}$-elements are $\boldsymbol{\Sigma}_{\boldsymbol{kk'}}$, $\langle \boldsymbol{G}_{\boldsymbol{kk'}}\rangle$ and $\boldsymbol{V}_{\boldsymbol{kk'}}$, respectively. When the potential is of delta-function type, $\boldsymbol{V}_{\boldsymbol{kk'}}$ is independent of $\boldsymbol{k}$ and $\boldsymbol{k'}$, and thus can be expressed just by $\boldsymbol{V}$. In this case, $\boldsymbol{\Sigma}_{\boldsymbol{kk'}}$ also becomes independent of $\boldsymbol{k}$ and $\boldsymbol{k'}$ inasmuch as $\langle \boldsymbol{G}_{\boldsymbol{kk'}}\rangle = G_{\boldsymbol{k}}(E)\,\delta_{\boldsymbol{kk'}}$ Thus upon putting $Z=\sum_{\boldsymbol{k}} G_{\boldsymbol{k}}(E)$, $\boldsymbol{\Sigma}$ is written as

$$\Sigma=NV\sum_{s=1}^{\infty}P_s(c)\,(ZV)^{s-1}. \tag{3·12a}$$

It must be noted hereupon that $\boldsymbol{\Sigma}(1)$ has been derived through a mathematically formal steps. In other words, analytic properties of Eq. (3·12) are not evident. Actually, as will be shown later, $\boldsymbol{\Sigma}(1)$ in Eq. (3·12) is not a correct first-order approximation since self-containedness is not retained, and accordingly $\boldsymbol{\Sigma}(1)$ in Eq. (3·12) suffers several unreasonable properties.

A detailed discussion concerning this point will be given in a succeeding section.

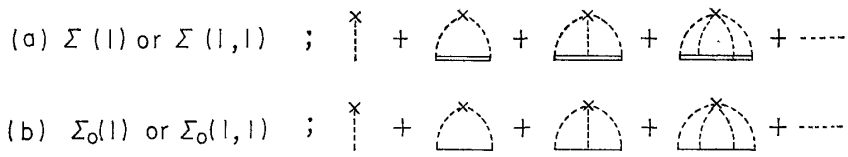

Fig. 5. First-order self-energy.

§ 4. The self-contained first-order approximation

4·1 *Difficulties of* $\boldsymbol{\Sigma}(1)$

The first-order self-energy $\boldsymbol{\Sigma}(1)$, which is derived by the principle of the preceding section and is expressed in terms of $P_s(c)$, has the following two deficiencies;

(1) There appear spurious poles in $\boldsymbol{\Sigma}(1)$.

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

(2) The sum of the series $P_s(c)\,(\boldsymbol{GV})^{s-1}$, or the right-hand members of Eq. (3·12), is not convergent, so that $\boldsymbol{\Sigma}(1)$ is not defined in an analytic sense.

In order to see how the first difficulty arises, let us expand $\boldsymbol{\Sigma}(1)$ in powers of c. For this purpose, an explicit form of $P_s(c)$ for an arbitrary s is evaluated in Appendix A; $P_s(c)$ becomes

$$P_s(c)=\sum_{m=1}^{s}P_{s,m}c^m\equiv\sum_{m=1}^{s}\left[\sum_{k=1}^{m}(-1)^{k-1}\binom{m-1}{k-1}k^{s-1}\right]c^m. \tag{4·1}$$

On inserting Eq. (4·1) into $P_s(c)$ in Eq. (3·12), $\boldsymbol{\Sigma}(1)$ is written as

$$\begin{aligned}\boldsymbol{\Sigma}(1)&=N\boldsymbol{V}\sum_{s=1}^{\infty}P_s(c)\,(\boldsymbol{GV})^{s-1}\\&=N\boldsymbol{V}\sum_{m=1}^{\infty}\left[\sum_{s=m}^{\infty}\sum_{k=1}^{m}(-1)^{k-1}\binom{m-1}{k-1}(k\boldsymbol{GV})^{s-1}\right]c^m\\&=N\boldsymbol{V}\sum_{m=1}^{\infty}\left[\sum_{k=1}^{m}(-1)^{k-1}\binom{m-1}{k-1}\frac{(k\boldsymbol{GV})^{m-1}}{\mathbf{1}-k\boldsymbol{GV}}\right]c^m.\end{aligned} \tag{4·2}$$

The coefficient of c has a pole characterized by the relation $\mathbf{1}-\boldsymbol{GV}=0$; this is the equation which determines the energy of the localized mode or resonance mode for one impurity problem, and thus this is a pole with a physical meaning. On the other hand, the coefficient of c^n for $n\geq 2$ has poles defined by the equation $\mathbf{1}-k\boldsymbol{GV}=0$ $(k=1, 2, \cdots, n-1, n)$. Among n poles, one with $k=1$ is physically allowed as explained in the above, while the other $(n-1)$ poles are spurious and proved to disappear when the self-energy diagrams up to nth order are calculated. These poles are interpreted as follows: $\mathbf{1}-k\boldsymbol{GV}=0$ $(k\geq 2)$ is the equation to give the localized mode or resonance mode for the system with only one impurity of potential $k\boldsymbol{V}$. This situation enables us to consider alternatively that the pole given by $\mathbf{1}-k\boldsymbol{GV}=0$ corresponds to a localized mode or resonance mode due to k impurity atoms existing on the same lattice site at once. Now that the probability of the multiple occupancy of a lattice site by more than one impurity atom is avoided in the present formulation, those $(n-1)$ poles are physically meaningless.

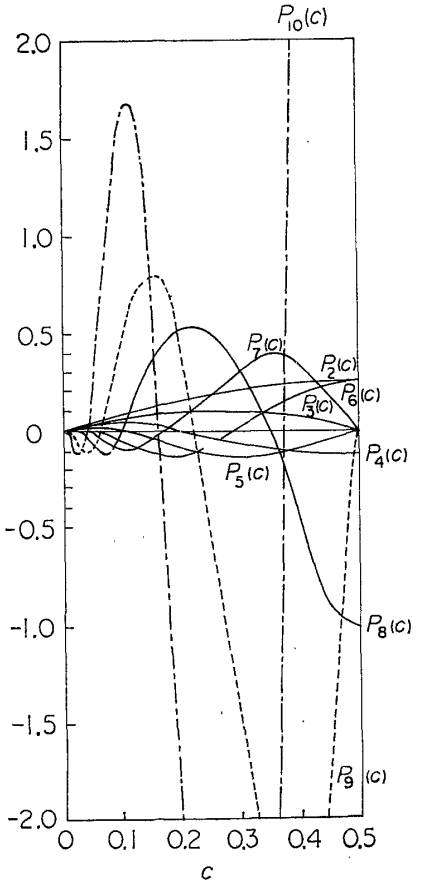

Fig. 6. $P_s(c)$ vs. c.

The second difficulty that the convergence of $\boldsymbol{\Sigma}(1)$ is not assured is recognized on noting

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

that $P_s(c)$ diverges as s increases. This divergent tendency of $P_s(c)$ is obviously seen from Fig. 6 in which $P_s(c)$ vs. c curves are shown for $2 \leq s \leq 10$.

4·2 *Self-containedness*

These difficulties of the first-order self-energy $\Sigma(1)$ result from the fact that the applied approximation is not self-contained. In other words, the relation between the degree of approximation and the correction factors to renormalize the comulant is not properly appreciated. For instance, let us recall how the fourth-order cumulant $C_4(\boldsymbol{p}_1, \cdots, \boldsymbol{p}_4)$ has been defined. As shown in Fig. 4, $C_4(\boldsymbol{p}_1, \cdots, \boldsymbol{p}_4)$ is determined on subtracting correction or renormalization factors from c; these correction factors originally come from all those diagrams which are of higher-order in concentration c than, and of the same order in interaction V with, the first diagram in Fig. 4. The correction from the crossed diagram (ii) in Fig. 4 is also included. (We call crossed diagrams such as diagram (ii) "irreducible" cluster diagrams in the sense that these diagrams give the essential cluster effect depending on the distance between two impurities, while these diagrams such as diagram (i) in Fig. 4 are not essentially cluster-diagrams since they can be reduced to a renormalization of $\boldsymbol{G}(\boldsymbol{E})$.)

Similarly, the coefficient $P_s(c)$ of s-th-order cumulant for $s \geq 5$ contains correction factors due to several crossed diagrams.

It must be noted that, when the first-order self-energy as pictured in Fig. 5 is to be calculated, the effect of all the crossed diagrams are neglected. Thus, it happens that, in spite of the fact that the diagram itself is not included in the given stage of approximation, the correction from the diagram is taken into ac-

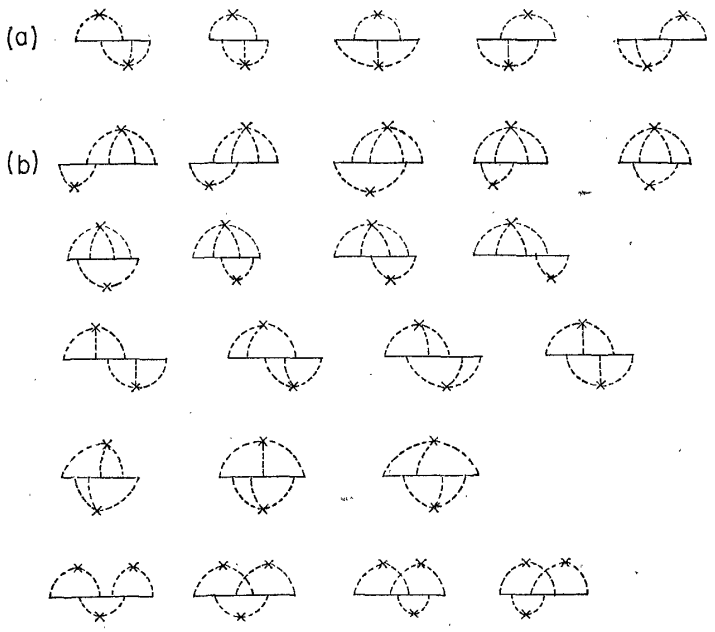

Fig. 7. Some of the correction factors to be included in 5th and 6th order cumulant.

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

count when $P_s(c)$ are determined. This is where the self-containedness is broken down.

This breakdown of the self-containedness can be understood easily by seeing the examples shown in Figs. 7(a) and (b). Five diagrams in Fig. 7(a) are not included in the first approximation, but the correction factors due to them are taken into consideration when $P_5(c)$ is evaluated; similarly, the corrections of 19 diagrams in Fig. 7(b) is included into $P_6(c)$ as renormalization factors though the diagrams themselves are eliminated from Fig. 5.

4·3 *Self-contained first-order approximation*

From the above discussion, the point to be made becomes clear; the cumulant or the factor to be assigned to a vertex must vary from approximation to approximation. This is because the cumulant is determined by the renormalization factors due to higher-order diagrams which are contained in the approximation under consideration.

Now, for example let us try to develop a method to remove from $P_s(c)$ those correction factors which should not be counted, when the first-order diagrams alone are considered. On the basis of Eq. (2·14), the correctly calculated cumulant of order s is written as

$$C_s^{(1)}(\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_s) = NQ_s(c)\,\delta(\boldsymbol{p}_1+\cdots+\boldsymbol{p}_s), \tag{4·3}$$

where the superscript (1) on C_s indicates that the cumulant is self-contained within the first-order approximation and coefficient $Q_s(c)$ instead of $P_s(c)$ is used to define the cumulant. In analogy with Eq. (3·8), $Q_s(c)$ is determined by means of $Q_m(c)$ with $m<s$ in the form;

$$\begin{aligned} Q_s(c) = c &- \sum_{(2)}{}' Q_{m_1}(c)\,Q_{s-m_1}(c) - \sum_{(3)}{}' Q_{m_1}(c)\,Q_{m_2}(c)\,Q_{s-m_1-m_2}(c) \\ &\cdots - \sum_{(s-1)}{}' Q_2(c)\,(Q_1(c))^{s-2} - (Q_1(c))^s, \end{aligned} \tag{4·4}$$

in which $\sum'_{(l)}$ implies all partitions of $p_1, \cdots, p_s$ into l groups such that the partitioned diagrams are the first-order diagrams alone.

Since it is difficult to derive directly an explicit expression of $Q_s(c)$ for an arbitrary s, let us try and look at the problem from a slightly different point of view. By a careful analysis of self-energy diagrams, it is recognized that the exclusion effect is also included as the renormalization factor to interaction; in this case, the contributions from vertices should remain unrenormalized. In order to understand this alternative interpretation, let us first consider how correction due to the diagram 8(b-i) is interpreted as a renormalization factor of an interaction line: The correction factor itself in this case appears in the form as shown in Fig. 8(b-ii) with a minus sign, which is topologically identical with the first-order diagram 8(b-iii); thus it becomes possible to renormalize the middle interaction line of (b-iii) and express the renormalized interaction by a wavy line as

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

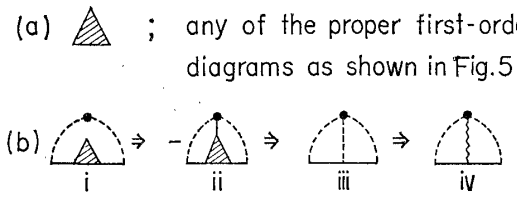

Fig. 8. Interpretation of the effects of the "exclusion effect" as a renormalization factor of V.

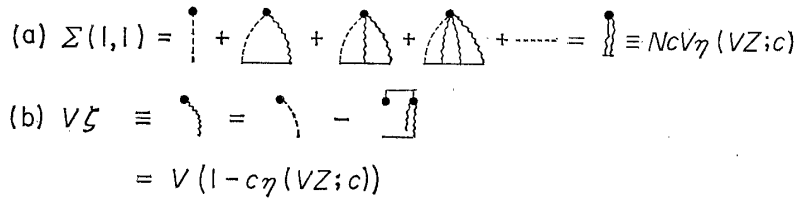

Fig. 9. Definition of ζ and η $(VZ; c)$.

in Fig. 8(b-iv). The essential points of the renormalization processes are stated in the above concerning Fig. 8 although, for more complicated diagrams, more detailed explanation will be necessary; this is given in Appendix B.

From the above explanation and with reference to Appendix B, it is seen that the self-contained first-order approximation $\boldsymbol{\Sigma}(1, 1)$ is pictured by the sum of diagrams in Fig. 9(a) by the use of a wavy line. A factor $\boldsymbol{V\zeta}$ is assigned to a wavy line where $\boldsymbol{\zeta}$ is a renormalization factor. A double wavy line is introduced to express the sum $\boldsymbol{\Sigma}(1, 1)$ and to this a factor $\boldsymbol{V\eta(GV}; c)$ is assigned. It is also apparent from the above discussion and from the explanation with regard to Fig. 8 that the single wavy line $\boldsymbol{V\zeta}$ is now in turn defined in terms of the double interaction line as depicted in Fig. 9(b).

From all the described consideration, it follows that the sum of the contributions from diagrams in Fig. 9(a) is written as

$$\begin{aligned}\boldsymbol{\Sigma}(1, 1) &= NcV + NcV\cdot G\cdot V\zeta + NcV\cdot G\cdot V\zeta\cdot G\cdot V\zeta + \cdots \\ &= \frac{NcV}{1-GV\zeta} \equiv NcV\eta(GV : c),\end{aligned} \tag{4·5}$$

while it is easily seen from Fig. 9(b) that

$$V\zeta = V - cV\eta(GV : c) = V\{1 - c\eta(GV : c)\}. \tag{4·6}$$

Therefore, ζ and $\eta(GV : c)$ are determined in a self-consistent manner, and a closed solution is yielded;

$$\eta(GV : c) = \frac{1}{1 - VG\{1 - c\eta(GV : c)\}}, \tag{4·7}$$

which fulfills the requirement that $Q_s(c)$ should satisfy Eq. (4·4). It is also easily proved that the first-order self-energy thus derived satisfies the host-defect

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

dual symmetry. An explicit form of $Q_s(c)$ is obtained on equating the coefficient of $(\boldsymbol{GV})^{s-1}$ in the expanded terms of $\boldsymbol{\Sigma}(1, 1)$

$$\boldsymbol{\Sigma}(1, 1) = N\boldsymbol{V}\sum_{s=1}^{\infty} Q_s(c)\ (\boldsymbol{GV})^{s-1} \tag{4·8}$$

and that of

$$\boldsymbol{\Sigma}(1, 1) = N\boldsymbol{V}c\boldsymbol{\eta}(\boldsymbol{GV} : c). \tag{4·9}$$

We have as a result;

$$\begin{aligned} Q_s(c) &= \sum_{m=1}^{s} Q_{s,m} c^m \\ &= \sum_{m=1}^{s} \left[(-1)^{m-1} \frac{(s+m-2)!}{m!\,(s-m)!\,(m-1)!} \right] c^m. \end{aligned} \tag{4·10}$$

A detailed derivation of $Q_s(c)$ is given in Appendix C.

It is worth mentioning hereupon that the self-contained first-order result $\Sigma(1, 1)$ satisfies the host-defect dual symmetry. Writing $\Sigma(1, 1)$ as

$$\Sigma(1, 1) \equiv NVI(x : c), \tag{4·9a}$$

where we consider delta-function type potential and $x = ZV$, we have the relation for $I(x : c)$

$$I(x : c) = \frac{c}{1 - x\{1 - I(x : c)\}} \tag{4·7a}$$

or

$$x\{I(x : c)\}^2 + (1-x)I(x : c) - c = 0. \tag{4·7b}$$

The above equation is solved for $I(x : c)$ in the form;

$$I(x : c) = \frac{1}{2x}\{-(1-x) + \sqrt{(1-x)^2 + 4cx}\},$$

by which it is easily verified that the dual symmetry

$$I(x : c) + I(-x : 1-c) = 1$$

is fulfilled. Although the above proof is carried out for the case of delta-function potential, the self-energy has the host-defect dual symmetry for the other type of potential as well.

Now, let us prove that the two difficulties of $\Sigma(1)$ discussed in subsection (4·1) are removed in the present formulation. The point that the first-order self-energy $\boldsymbol{\Sigma}(1, 1)$ evaluated in a self-contained manner is free from any spurious poles is shown by expanding $\boldsymbol{\Sigma}(1, 1)$ in powers of c; i.e.

$$\boldsymbol{\Sigma}(1, 1) = N\boldsymbol{V}\sum_{s=1}^{\infty}\sum_{m=1}^{s}(-1)^{m-1}\frac{(s+m-2)!}{m!\,(s-m)!\,(m-1)!}\, c^m (\boldsymbol{GV})^{s-1}$$

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

$$=NV\sum_{m=1}^{\infty}\left[\frac{(-1)^{m-1}(2m-2)!}{m!(m-1)!}\frac{(\boldsymbol{GV})^{m-1}}{(\mathbf{1}-\boldsymbol{GV})^{2m-1}}\right]c^m, \qquad (4\cdot11)$$

where use is made of Eq. (4·10). Thus, it is apparent that the coefficient of c^m has no other poles than $\mathbf{1}-\boldsymbol{GV}=0$. As for the second difficulty of divergence, there is no fear of occurence because $Q_s(c)$ converges as shown in Fig. 10 in which $Q_s(c)$ vs c curves for $2\leq s\leq 10$ are depicted. We have

$$|Q_s(c)|\lesssim Q_2(c)\leq Q_1(c)$$
$$\text{for all } c \ (s\geq 3). \qquad (4\cdot12)$$

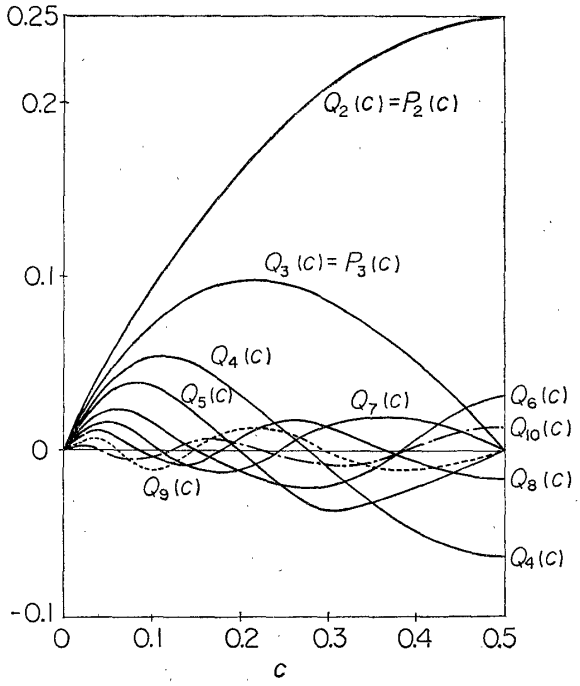

Fig. 10. $Q_s(c)$ vs. c.

4·4 *Remarks on the renormalization factors*

It follows from the above discussion that, in order to give an approximate self-energy a physical meaning, it is required the correction factors to renormalize cumulants must be chosen such that the self-containedness in the given stage of approximation is ensured. Namely, $Q_s(c)$ which determines the cumulant of order s differs for a different type of approximation. Obviously, $P_s(c)$ is correct when all proper self-energy diagrams can be summed up and an exact self-energy is attained some way.

In order to see that the coefficient of the cumulant varies according to approximation, we attempt to calculate cumulants for another approximation. We adopt, for this purpose, an approximation in which those first-order diagrams containing the unperturbed propagator instead of the true propagator are summed (cf. Fig. 5(b)). Through completely analogous steps to the case of evaluating $\boldsymbol{\Sigma}(1,1)$, the self-energy $\boldsymbol{\Sigma}_0(1,1)$ of Fig. 5(b) is defined by Figs. 11(a) and (b) which lead to the equations;

$$\boldsymbol{\Sigma}_0(1,1)=Nc\boldsymbol{V}+Nc\boldsymbol{V}\cdot\boldsymbol{G}\cdot\boldsymbol{V}\zeta_0+Nc\boldsymbol{V}\cdot\boldsymbol{G}\cdot\boldsymbol{V}\zeta_0\cdot\boldsymbol{G}_0\cdot\boldsymbol{V}\zeta_0+\cdots$$
$$=\frac{Nc\boldsymbol{V}}{\mathbf{1}-\boldsymbol{G}_0\boldsymbol{V}\zeta_0} \qquad (4\cdot13)$$

and

$$\boldsymbol{V}\zeta_0=\boldsymbol{V}(1-c), \qquad (4\cdot14)$$

ζ_0 is the renormalization factor for the present model; ζ_0 is provided by Fig. 11(b) and written as Eq. (4·14) where the requirement for ζ_0 is such that only

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

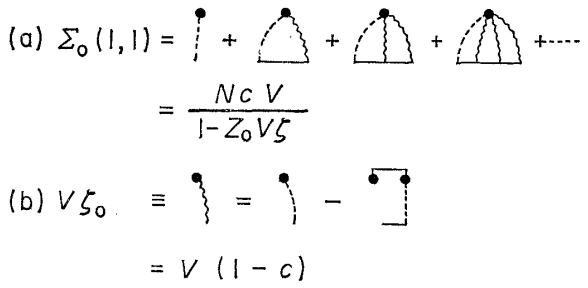

Fig. 11. Definition of ζ_0 and Σ_0 (1·1).

those corrections due to simple diagrams are taken into account while the correction factors corresponding to reducible cluster diagrams as well as irreducible cluster diagrams should be neglected.

The self-energy determined by the relations (4·13) and (4·14) is self-contained in the sense that the corrections for renormalization are counted from only those diagrams as shown in Fig. 5(b) and the sequence of said sorts of diagrams. The self-energy $\boldsymbol{\Sigma}_0(1, 1)$ is expanded in powers of $\boldsymbol{G}_0\boldsymbol{V}$ with $Q_s^0(c)$ as coefficients;

$$\boldsymbol{\Sigma}_0(1, 1) = N\boldsymbol{V}\sum_{s=1} Q_s^0(c)\,(\boldsymbol{G}_0\boldsymbol{V})^{s-1}. \tag{4·15}$$

Comparison of Eqs. (4·13) and (4·15) yields

$$Q_s^0(c) = c(1-c)^{s-1} \quad (s>1). \tag{4·16}$$

This result $\boldsymbol{\Sigma}_0(1, 1)$ agrees with the approximation proposed by Elliot and Taylor[9] or other authors as a good approximation for small c.

4·5 *Relation between* $\boldsymbol{\Sigma}(1)$ *and* $\boldsymbol{\Sigma}(1, 1)$

Before concluding this section, it is interesting to investigate the relation between $\boldsymbol{\Sigma}(1)$ given by Eq. (3·12) and $\boldsymbol{\Sigma}(1, 1)$ of Eq. (4·8) or (4·9) which is an exact first-order self-energy in the rigorous sense of the term. For this purpose, we note that, when the potential is of the delta-function type, $\Sigma(1)$ is described by the use of a continued fraction as follows:[7]

$$\Sigma(1)_{\boldsymbol{kk'}} = NV\delta_{\boldsymbol{kk'}}[c + VZf_1(VZ\,;\,c)], \tag{4·17}$$

in which

$$f_1(x : c) = \frac{c(1-c)}{1-(1-2c)x + x^2 f_2(x : c)}, \tag{4·18a}$$

$$f_n(x : c) = \frac{n(n-1)c(1-c)}{1-n(1-2c)x + x^2 f_{n+1}(x : c)}. \quad (n\geq 2) \tag{4·18b}$$

The infinite continued fraction $f_n(x : c)$ can be put in a closed form by approximating $f_2(x : c)$ by $f_1(x : c)$. This is called a first approximant. Thus $f_1(x : c)$ can be calculated and accordingly $\boldsymbol{\Sigma}(1)$ is evaluated. It is readily shown that the self-energy $\boldsymbol{\Sigma}(1)$ thus obtained is identical with the exact first-order self-energy $\boldsymbol{\Sigma}(1, 1)$ of Eq. (4·9). Namely, by putting $f_2(x : c) = f_1(x : c)$ in Eq. (4·18a), we have

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

$$f_1(x:c)=\frac{c(1-c)}{1-(1-2c)x+x^2f_1(x:c)}=\frac{c(1-c)}{1-(1-c)x+x\{c+xf_1(x:c)\}}$$

or

$$c+xf_1(x:c)=\frac{c+cx\{c+xf_1(x:c)\}}{1-(1-c)x+x\{c+xf_1(x:c)\}}, \tag{4·18c}$$

which reduces to the relation

$$c+xf_1(x:c)=\frac{c}{1-x+x\{c+xf_1(x:c)\}}. \tag{4·7a}$$

On noting that $c\eta(x:c)=c+xf_1(x,c)$, we see that the proof is fulfilled. It is also noteworthy that $\boldsymbol{\Sigma}(1,1)$ agrees with the result obtained by Taylor[8] for large c and with that by Onodera and Toyozawa.[10]

§ 5. Higher approximations

In this section, a systematic method for treating higher-order approximations is to be given. With this goal in mind, we first note that $\boldsymbol{\Sigma}(1)$ obtained in § 3 is the first-order self-energy as shown in Fig. 5(a) where $P_s(c)$ is assigned with a vertex connected to the propagator by an interaction lines. More precisely, the correction factors to renormalize a vertex originate in all possible higher-order diagrams, irreducible as well as reducible clusters. On the other hand, $\boldsymbol{\Sigma}(1,1)$ is the first-order self-energy in which only those corrections due to reducible cluster diagrams are counted. Thus, in general, $\boldsymbol{\Sigma}(1)$ including all the correction factors is expressed in the form;

$$\boldsymbol{\Sigma}(1)=\boldsymbol{\Sigma}(1,1)+\boldsymbol{\Sigma}(1,2)+\boldsymbol{\Sigma}(1,3)+\cdots, \tag{5·1}$$

in which $\boldsymbol{\Sigma}(1,2)$ is the sum of the correction factors to the first-order self-energy caused by irreducible clusters consisting of two impurity atoms; to these irreducible two impurity clusters or second-order clusters, the five diagrams of Fig. 7(a) and the first 15 diagrams of Fig. 7(b) are included. The correction factors $\boldsymbol{\Sigma}(1,2)$ are obtained by first connecting two independent verteces of two types of first-order proper diagrams and multiplying a factor (-1) to the contributions of the thus-obtained diagrams. Namely, when the sum of the contributions from the mentioned second-order clusters is written as $\sum_{\boldsymbol{R}}\mathcal{I}_2(\boldsymbol{R})$, $\boldsymbol{R}$ being the distance between two impurity atoms in a cluster, $\boldsymbol{\Sigma}(1,2)$ is expressed by $-\mathcal{I}_2(\boldsymbol{R}=0)$. The third term $\boldsymbol{\Sigma}(1,3)$ indicates the correction factors in $\boldsymbol{\Sigma}(1)$ due to irreducible clusters made of three impurity atoms, such as the last four diagrams in Fig. 7(b).

Similar equations hold for higher-order cases. That is to say, if the nth-order self-energy $\boldsymbol{\Sigma}(n)$ is evaluated in such a manner that the contribution from a vertex with s interaction lines is regarded to give a factor $P_s(c)$, then $\boldsymbol{\Sigma}(n)$ becomes

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

$$\boldsymbol{\Sigma}(n) = \boldsymbol{\Sigma}(n, n) + \boldsymbol{\Sigma}(n, n+1) + \boldsymbol{\Sigma}(n, n+2) + \cdots, \tag{5·2}$$

The notation $\boldsymbol{\Sigma}(n, n)$ designates the nth-order self-energy including the corrections from nth-order irreducible diagrams alone, whence $\boldsymbol{\Sigma}(n, m)$ is the correction factors to the self-energy of order n originating in the mth-order irreducible clusters.

It has been shown in the preceding section that the first-order self-energy calculated in a self-contained manner is given by $\boldsymbol{\Sigma}(1, 1)$. In the same way, self-contained higher-order approximations $\boldsymbol{\Sigma}_n(\boldsymbol{k})$ are generally defined in terms of the newly introduced notations $\boldsymbol{\Sigma}(n, m)$ as will be stated below. Let us first consider the exact second-order self-energy $\boldsymbol{\Sigma}_2(\boldsymbol{k})$. In order to attain the self-containedness within the second-order, it is necessary to first derive $\boldsymbol{\Sigma}(2, 2)$, the sum of the second-order self-energy diagrams with corrections from irreducible two-impurity clusters, and next to take account of the correction factors from the second-order irreducible diagrams to the first-order self-energy. Thus we have

$$\boldsymbol{\Sigma}_2(k) = \boldsymbol{\Sigma}(2, 2) + \boldsymbol{\Sigma}(1, 2). \tag{5·3}$$

In complete analogy with the second-order case, the self-contained third-order approximation is given by

$$\boldsymbol{\Sigma}_3(k) = \boldsymbol{\Sigma}(3, 3) + \boldsymbol{\Sigma}(2, 3) + \boldsymbol{\Sigma}(1, 3). \tag{5·4}$$

In summary, a systematic way to obtain higher-order approximations is explained with reference to Table 1. The sum of mth row is $\boldsymbol{\Sigma}(m)$ while the sum of nth column is $\boldsymbol{\Sigma}_n(k)$. So far, it has been widely accepted that $\boldsymbol{\Sigma}(1)$ is the first approximation and that $\boldsymbol{\Sigma}(2)$, $\boldsymbol{\Sigma}(3)$, etc. should be included succeedingly, when 2nd, 3rd, ···approximations are required. However, these conventional conclusions are actually not correct since; (a) spurious poles are present in $\boldsymbol{\Sigma}(1)$, $\boldsymbol{\Sigma}(2)$, etc.; and (b) the series expansion determining $\boldsymbol{\Sigma}(n)$ in general is not convergent, now that $P_s(c)$ is divergent.

Table I. Higher approximations for self-energy.

						total
	$\Sigma(1, 1)$	$\Sigma(1, 2)$	$\Sigma(1, 3)$	$\Sigma(1, 4)$	.....	$\Sigma(1)$
		$\Sigma(2, 2)$	$\Sigma(2, 3)$	$\Sigma(2, 4)$	.....	$\Sigma(2)$
			$\Sigma(3, 3)$	$\Sigma(3, 4)$	.....	$\Sigma(3)$
				$\Sigma(4, 4)$	.....	$\Sigma(4)$
total	$\Sigma_1(k)$	$\Sigma_2(k)$	$\Sigma_3(k)$	$\Sigma_4(k)$		Σexact

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

These difficulties are overcome by taking $\boldsymbol{\Sigma}_1(k)$, $\boldsymbol{\Sigma}_2(k)$, etc. as first, second and higher approximations. As shown in Table I, $\boldsymbol{\Sigma}_n(k)$ is determined by

$$\boldsymbol{\Sigma}_n(k)=\boldsymbol{\Sigma}(1,\ n)+\boldsymbol{\Sigma}(2,\ n)+\cdots+\boldsymbol{\Sigma}(n,\ n). \tag{5·5}$$

With the help of physical consideration of $\boldsymbol{\Sigma}(n,\ m)$, it is made clear how spurious poles are removed. For the sake of explanation, the second-order approximation is discussed as an example; from the above discussion, $\boldsymbol{\Sigma}(2,\ 2)$ is by definition

$$\boldsymbol{\Sigma}(2,\ 2)=\sum_{\boldsymbol{R}}\mathcal{J}_2(\boldsymbol{R}). \tag{5·6}$$

Note that the term, in which $\boldsymbol{R}=0$ is included, is physically meaningless owing to the fact that the multiple occupancy of a lattice site with more than one impurity atom is excluded. This term is one of the causes of unphysical poles in $\Sigma(2)$. The second term of Eq. (5·3) by definition is given by

$$\boldsymbol{\Sigma}(1,\ 2)=-\mathcal{J}_2(\boldsymbol{0}) \tag{5·7}$$

As a result, we have

$$\boldsymbol{\Sigma}_2(k)=\sum_{\boldsymbol{R}(\neq 0)}\mathcal{J}_2(\boldsymbol{R}). \tag{5·8}$$

So that $\boldsymbol{\Sigma}(1,\ 2)$ has the effect of cancelling a term of $\boldsymbol{\Sigma}(2,\ 2)$ with $\boldsymbol{R}=0$, and thus $\boldsymbol{\Sigma}_2(k)$ is free from spurious poles. It is instructive to see that the same situation takes place for the third-order approximation; i.e. $\boldsymbol{\Sigma}(3,\ 3)$ is given by a function $\mathcal{J}_3(\boldsymbol{R}_{12},\ \boldsymbol{R}_{23},\ \boldsymbol{R}_{31})$ where $\boldsymbol{R}_{12}$, $\boldsymbol{R}_{23}$ and $\boldsymbol{R}_{31}(=-\boldsymbol{R}_{12}-\boldsymbol{R}_{23})$ are respectively the distances between three pairs of impurity atoms out of atoms 1, 2 and 3;

$$\boldsymbol{\Sigma}(3,\ 3)=\sum_{\boldsymbol{R}_{12},\boldsymbol{R}_{23},\boldsymbol{R}_{31}}\mathcal{J}_3(\boldsymbol{R}_{12},\ \boldsymbol{R}_{23},\ \boldsymbol{R}_{31})\,\delta(\boldsymbol{R}_{12}+\boldsymbol{R}_{23}+\boldsymbol{R}_{31}). \tag{5·9}$$

In the sum of the right-hand member of Eq. (5·9), the following two terms, which are not actually allowed, are included; (a) Any one of $\boldsymbol{R}_{12}$, $\boldsymbol{R}_{23}$ and $\boldsymbol{R}_{31}$ vanishes while the other two variables are not zero; (this corresponds to the case in which two out of three impurity atoms exist on the same lattice site while the third atom is on some other lattice point. (b) The case in which $\boldsymbol{R}_{12}=\boldsymbol{R}_{23}=\boldsymbol{R}_{31}=0$; this means that all three impurity atoms come on the same lattice site at once. The contributions from such terms as described in (a) are cancelled by $\boldsymbol{\Sigma}(2,\ 3)$ while the effect of $\boldsymbol{\Sigma}(1,\ 3)$ is to cancel the term in which all $\boldsymbol{R}_{12}$, $\boldsymbol{R}_{23}$ and $\boldsymbol{R}_{31}$ vanish. Thus n spurious poles appear.

As for the strictly calculated nth-order self-energy $\boldsymbol{\Sigma}_n(\boldsymbol{k})$, $\boldsymbol{\Sigma}(1,\ n)$, $\boldsymbol{\Sigma}(2,\ n)$, $\cdots$, and $\boldsymbol{\Sigma}(n-1,\ n)$ (n; arbitrary) have their respective corresponding terms in $\boldsymbol{\Sigma}(n,\ n)$ with the same magnitude and a different sign, so that all spurious poles present in $\boldsymbol{\Sigma}(n,\ n)$ are completely subdued.

So far as the convergence of $\boldsymbol{\Sigma}_n(\boldsymbol{k})$ is concerned, it is shown that $\boldsymbol{\Sigma}_n(\boldsymbol{k})$ are constructed on the basis of $Q_s(c)$ obtained in § 4 and thus, in view of the convergence of $Q_s(c)$, it is concluded that $\boldsymbol{\Sigma}_n(\boldsymbol{k})$ are well-defined.

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

§ 6. Discussion

It has been presented in §§ 2 to 4 that the self-contained treatment of approximations for the self-energy is needed in order to obtain physically reasonable results, and the technique of deriving the exact form of the first-order self-energy $\Sigma(1, 1)$ is developed. It is also verified that $\Sigma(1, 1)$, which is self-contained within the scope of the first-order approximation, agrees with the first approximant of $\Sigma(1)$, the latter being attained from purely mathematical point of view. This first-order self-energy $\Sigma(1, 1)$ leads to the same equation as that of Taylor; $\Sigma(1, 1)$ is also identical with the result of Onodera and Toyozawa.

It has been shown by Taylor that $\Sigma(1, 1)$ serves as a good approximation for such an impurity concentration c as is larger than the critical percolation concentration c_p. The obtained density of states along the same approximation accounts for the fact that $\Sigma(1, 1)$ is far from a good approximation in the concentration range $c<c_p$, since the approximate self-energy $\Sigma(1, 1)$ fails to explain the fine structure of the state density in the forbidden gap, the fine structure being confirmed to exist by computer work. A similar discussion on the method of approximation and the concentration range where the considered approximation is adequate is found in the recent work by Matsuda and Okaka.[11]

These results suggest that some other approximation must be introduced in order to treat the systems with small concentration of impurity atoms. It reasonably follows without a logical gap that, for the purpose of explaining the fine structure, the higher-order self-energy which represents the effect of cluster is to be included. When c is small enough not to produce a lot of spikes in the impurity band region except the one around the local mode due to one impurity atom, $\Sigma(1, 1)$ will act as a good approximation for the true self-energy. According as c increases, the second-, third-, fourth- and higher-order self-energy should be taken into consideration step by step. The criterion up to what order the approximation should be picked up is determined by the following way; if the $(n+1)$th-order approximation does not cause a drastic change in the resulting physical quantities compared with those of the nth-order approximation, then it is sufficient to calculate up to the nth-order approximation. However, with the increase of c, it will not continue without end. At an appropriate concentration, it stops and thereafter any higher-order approximation yields no remarkable change to the results of the first-order approximation. The concentration at which this takes place gives the percolation concentration. For $c>c_p$, the result with $\Sigma_1(k)+\Sigma_2(k)$ may be almost the same as that obtained in the first approximation alone. The detailed evaluation of the second-order self-energy ($\Sigma(1, 2)$ and $\Sigma(2, 2)$) will appear in a forthcoming paper.

Acknowledgement

The author is indebted to Prof. T. Matsubara for a valuable discussion and

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

a critical reading of the manuscript. She also acknowledges helpful conversation with Prof. H. Matsuda and with Dr. S. Takeno.

Appendix A

Evaluation of an explicit form for $P_s(c)$

The coefficient $P_s(c)$ of the sth cumulant is a polynomial in c of order s so that it can be written as

$$P_s(c)=\sum_{m=1}^{s} P_{s,l}c^l. \tag{A·1}$$

Inserting Eq. (A·1) into $P_s(c)$ of Eq. (3·11), we have

$$\begin{aligned} g(x:c)\equiv\log(1-c+ce^x)&=\sum_{s=1}^{\infty}\sum_{l=1} P_{s,l}c^l x^s/s! \\ &=\sum_{l=1}^{\infty}\Big(\sum_{s=l}^{\infty} P_{s,l}x^s/s!\Big)c^l. \end{aligned} \tag{A·2}$$

The nth derivative of the left-hand member in Eq. (A·2) by c yields

$$\begin{aligned} \frac{\partial^n g(x;c)}{\partial c^n}\bigg|_{c=0}&=\frac{(-1)^{n-1}(n-1)!(e^x-1)^n}{(1-c+ce^x)^n}\bigg|_{c=0}=(-1)^{n-1}(n-1)!(e^x-1)^n \\ &=(-1)^{n-1}(n-1)!\sum_{k=0}^{n}\binom{n}{k}(-1)^{n-k}e^{kx}, \end{aligned} \tag{A·3}$$

where c is put equal to zero, while that of the right-hand member becomes

$$\frac{\partial^n g(x:c)}{\partial c^n}\bigg|_{c=0}=n!\sum_{s=n}^{\infty} P_{s,n}x^s/s!\,. \tag{A·4}$$

On taking the sth derivative by x of Eqs. (A·3) and (A·4) under the condition that $x=0$ and equating the obtained results to each other, it is concluded that

$$P_{s,n}=\sum_{k=1}^{n}(-1)^{k-1}\binom{n-1}{k-1}k^{s-1}. \tag{A·5}$$

The analytic properties of $P_s(c)$ has also been discussed by Leath and Goodmann[12] in terms of the starling formula.

Appendix B

In subsection 4·3, it is stated that the exclusion effect is interpreted either as the renormalization of vertices or of interaction lines. In order to understand this alternative interpretation of the exclusion effect, it is convenient to study the diagrammatic expansion corresponding to the iterative terms of the Green's function. Recalling the steps through which the unrestricted diagrams in Figs. 2 and 3 are introduced and further investigating the similar steps for higher-order moments, we notice how to determine a factor and a sign to be assigned

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

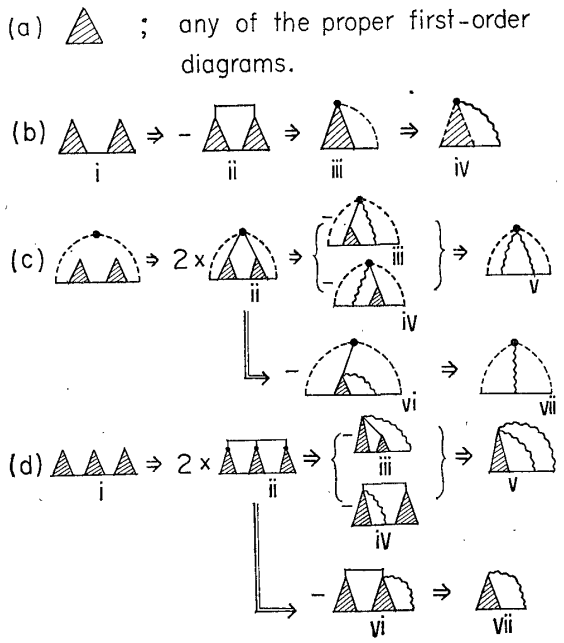

Fig. 12. Interpretation of the effects of the "exclusion effect" as a renormalization factor of V. —Complicated cases.—

to a given type of diagram. Thus, with reference to Figs. 8(a) and (b), and to Figs. 12(a) through 12(d), diagrams 8(b-ii) and 12(b-ii) carry a factor -1 while a factor $+2$ is attached to diagrams 12(c-ii) and 12(d-ii). These four diagrams appear as the corrections to the four diagrams 8(b-i), 12 (b-i), (c-i) and (d-i), respectively. Note that a hatched triangle represents any of the first-order proper self-energy diagram found in Fig. 5(b). Here, we limit our consideration to only those diagrams as included in the present approximation; in other words, all the crossed diagrams are discarded.

We have seen, in subsection 4·2, how the correction due to the diagram 8(b-i) is interpreted as a renormalization factor of an interaction line. Logical steps completely analogous with the example in § 4·2 are followed in order to ascribe the correction coming from the diagram 12(b-i) to the renormalized diagram (b-iv).

In the next place, the renormalization brought about by the diagram (c-i) is expressed by the diagram (c-ii); this correction is reduced to two renormalized diagrams v and viii through either one of two processes iii and iv and through step vi, respectively. Similar explanation by means of diagrams is embodied in Figs. 12(d) for another example. Remember that a factor -1 is introduced at every stage; thus two stages result in the plus sign. On the other hand, the factor 2 is explained by the two different steps through which to attain the final diagram.

In the same way, it is straightforward a task, laborious though it may be, to prove that correction of each higher-order diagram can be included in an appropriate first-order diagrams with renormalized interaction lines.

The important point is that, when all the crossed diagrams are omitted, the factor and the sign of a given diagram are completely included in the process of reducing the correction diagrams to renormalized interaction lines.

Appendix C

Evaluatin of an explicit form for $Q_s(c)$

Equations (4·8) and (4·9) are rewritten as

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

$$\sigma(x:c)\equiv c\eta(x:c)=\sum_{s=1}^{\infty}Q_s(c)\,x^{s-1}, \tag{C·1}$$

where $Q_s(c)$ is a polynomial in c of order s given in the form as

$$Q_s(c)=\sum_{m=1}^{s}Q_{s,m}c^m. \tag{C·2}$$

Substituting $Q_s(c)$ in Eq. (C·1) by that of Eq. (C·2), we obtain

$$\begin{aligned}\sigma(x:c)&=\sum_{s=1}^{\infty}\Big[\sum_{m=1}^{s}Q_{s,m}c^m\Big]x^{s-1}\\&=\sum_{m=1}^{\infty}\Big[\sum_{s=m}^{\infty}Q_{s,m}x^{s-1}\Big]c^m.\end{aligned} \tag{C·3}$$

The first derivative of $\sigma(x;c)$ with reference to c is obtained by the use of Eq. (4·7) as

$$\frac{\partial\sigma(x:c)}{\partial c}=\frac{1}{1-x(1-2\sigma(x:c))}\equiv F(x:c). \tag{C·4}$$

It is easily seen that the first derivative of $F(x:c)$ is determined by $F(x:c)$ itself as

$$\frac{\partial F(x:c)}{\partial c}=-2xF^3(x:c). \tag{C·5}$$

Equations (C·4) and (C·5) yield the nth derivative by c of $\sigma(x:c)$ in the form as

$$\frac{\partial^n\sigma(x:c)}{\partial c^n}=(-2x)^{n-1}3\cdot5\cdots(2n-3)\,[F(x:c)]^{2n-1}. \quad (n>2) \tag{C·6}$$

On noting that $F(x;0)=1/(1-x)$, we get

$$\begin{aligned}\left.\frac{\partial^n\sigma(x:c)}{\partial c^n}\right|_{c=0}&=\frac{(-1)^{n-1}(2n-3)!}{(n-1)!}\,\frac{x^{n-1}}{(1-x)^{2n-1}}\\&=\frac{(-1)^{n-1}(2n-2)!}{(n-1)!}\sum_{m=2n-2}^{\infty}{}_mC_{2n-2}x^{m-n+1}\\&\equiv n!\sum_{s=n}^{\infty}Q_{s,n}x^{s-1}.\end{aligned} \tag{C·7}$$

Then if we differentiate Eq. (C·7) by x and subsequently equate x to zero, the final conclusion is achieved in the form;

$$Q_{s,n}=\frac{(-1)^{n-1}(n+s-2)!}{(n-1)!\,n!\,(s-n)!}. \tag{C·8}$$

References

1) For instance, review article by A. A. Maradudin, P. Mazur, E. W. Montroll and G. H. Weiss, Rev. Mod. Phys. **30** (1958), 175.

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022

2) S. F. Edwards, Phil. Mag. **3** (1958), 1020.
R. Klauder, Ann. of Phys. **14** (1961), 43.
J. S. Langer, J. Math. Phys. **2** (1961), 584.
F. Yonezawa, Prog. Theor. Phys. **31** (1964), 357.
3) P. Dean, Proc. Roy. Soc. **A254** (1960), 507.
P. Dean and H. D. Bacon, Proc. Roy. Soc. **283** (1965), 64.
4) D. N. Payton and W. M. Visscher, Phys. Rev. **154** (1967), 802.
5) H. Matsuda, Prog. Theor. Phys. **27** (1962), 811.
J. Hori, Prog. Theor. Phys. **32** (1964), 471.
J. Hori and M. Fukushima, J. Phys. Soc. Japan **19** (1964), 296.
6) F. Yonezawa and T. Matsubara, Prog. Theor. Phys. **35** (1966), 357, 759.
7) T. Matsubara and F. Yonezawa, Prog. Theor. Phys. **37** (1967), 1346.
8) D. W. Taylor, Phys. Rev. **156** (1967), 1017.
9) R. J. Elliot and D. W. Taylor, Proc. Roy. Soc. **296** (1967), 161.
10) Y. Onodera and Y. Toyozawa, J. Phys. Soc. Japan **24** (1968), 341.
11) H. Matsuda and K. Okada, to be published in Prog. Theor. Phys.
12) P. L. Leath and B. Goodmann, Phys. Rev. **148** (1966), 968.

Downloaded from https://academic.oup.com/ptp/article/40/4/734/1926294 by guest on 21 August 2022