

SUCCESSIVE APPROXIMATION METHODS IN CLASSICAL STATISTICAL MECHANICS

by G. E. UHLENBECK

University of Michigan

1. *Introduction.* In this report I will try to summarize some work on graph theoretical methods in classical statistics, which will appear in a review article by the author and G. W. Ford *). On this occasion I would also like to make a plea for a more standard terminology in the use of graphs in various approximation methods and it seems best to stick as close as possible to the accepted mathematical terms.

a. Let me begin therefore with some basic notions connected with graphs. A (*linear*) *graph* is a collection of points with lines between certain pairs of points. In general there may be more than one line connecting two points; also loops may occur. If any two points are connected by at most one line and if there are no loops present, we call the graph *simple*. Otherwise the graph is called *not simple*. (See fig. 1)

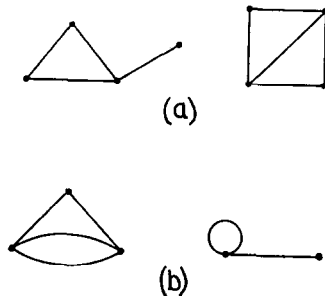


Fig. 1. a) Simple linear graphs b) Not simple graphs

A subset of points which are joined successively by lines is called a *path* connecting the initial and final points. If the final point coincides with the initial point we speak of a *cycle*. A *connected* graph is one in which there is at least one path between any two points. Otherwise the graph is *disconnected*.

*) G. W. Ford and G. E. Uhlenbeck: The theory of linear graphs with applications to the theory of the virial development of the properties of gases. To appear in: Progress of Statistical Physics, Vol. I. All references can be found there.

Mostly it is sufficient to study the connected graphs. The *degree* of a point is the number of lines incident upon it. A graph is *directed* if directions are specified along the lines; usually the direction is indicated by an arrow. *Euler graphs* are directed graphs with as many lines coming into, as out of, each point. Directed graphs are of importance in e.g. quantum field theory. An *articulation point* is a point which, if omitted, would cause the graph to fall into disconnected parts. A graph without articulation points is called a *star*. (See fig. 2)

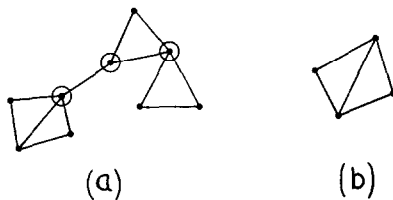


Fig. 2. a) Graph with articulation points b) Star.

A general connected graph can be considered as built up out of the constituent stars hung together in the articulation points. If the constituent stars are just single lines we call the graph a *Cayley tree*. If the stars are polygons then we will speak of a *Husimi tree*. If more general types of stars are used we will speak of star trees. A *pure* star tree is one which consists of only one type of stars; otherwise the star tree is *mixed*. Of course if no restrictions are made on the type of stars, then the notions of star tree and connected graph are synonymous.

b. Graphs always appear as a symbolization of the different terms in a successive approximation method and the number of points usually determines the order of the approximation. In each successive approximation method one can therefore distinguish two problems:

1. The combinatorial problem: how many "different" terms are there in n^{th} order, or how many "different" graphs of n points are there.
2. The integral problem: what is the contribution of each graph. Mainly the combinatorial problem will be considered here.

c. In the combinatorial problem one distinguishes *labeled graphs* and *free* (or *topological*) graphs. In a labeled graph different points are distinguished by some index. Also, since a point often represents a molecule, a point can have more than one label, as for instance in mixture problems. In a free graph the points are regarded as not distinguishable. Of fundamental importance is the number of different labeled graphs corresponding to a given free graph. To enumerate the possibilities the notion of the *group* of a graph is needed. The group of a graph is the group of automorphisms, i.e. of one to one correspondences of the points leaving the connections invariant. The group can be considered as a permutation group of the points. For

example, for an n -gon, the group is the dihedral group of order $2n$, and for a so-called *complete* graph of n points (i.e., the graph where all pairs of points are joined by lines) the group is the symmetric groups of degree n . The order of the group is the *symmetry number* s of the graph. To a free graph of p points there are $p!/s$ labeled graphs. Intermediate between the labeled and the free graphs are the *rooted* graphs, in which one point, the root, is specified. The *main leaves* of a rooted star tree are the stars which have the root point in common.

2. *Application to the theory of non-ideal gases.* Although the methods can also be applied to the calculation of distribution functions, we will restrict ourselves to the calculation of the classical partition function for an N -particle system with additive two-body forces

$$Z(V, T, N) = \frac{1}{N! \lambda^{3N}} \int_{\mathcal{V}} \dots \int d\mathbf{r}_1 \dots d\mathbf{r}_N e^{-\beta \sum_{i < j} \Phi(r_{ij})};$$

$$\lambda = h(2\pi m k T)^{-1/2}, \quad \beta = 1/kT. \quad (2.1)$$

The usual trick for the low density case, due to Mayer, is to express the integrand in product form by

$$e^{-\beta \sum_{i < j} \Phi(r_{ij})} = \prod_{i < j}^N (1 + f_{ij}), \quad (2.2)$$

$$f_{ij} = e^{-\beta \Phi(r_{ij})} - 1. \quad (2.3)$$

If one represents each factor f_{ij} in the terms of the expansion of the product form by a line one gets all N -point labeled graphs, connected and disconnected, from N separate points up to the complete graph with $\frac{1}{2}N(N-1)$ lines. The two basic theorems of Mayer's theory can be regarded as special cases of two general theorems in the theory of graphs. We will sketch the proof of these general theorems, and indicate the relation to the Mayer theory.

Theorem I. Consider the quantity F_N defined by

$$F_N = \sum_{(G_N)} W(G_N), \quad (2.4)$$

where $W(G_N)$ is a weight assigned to the graph G_N of N points and where the sum runs over all N -point labeled graphs. The weight $W(G_N)$ is supposed to have the properties:

a. $W(G_N)$ is independent of the labeling of the points

b. $W(G_N) = \prod_{(\text{all } C_l)} W(C_l)$

where C_l is one of the disjoint connected parts of G_N . We introduce also quantities f_l by a definition analogous to (2.4):

$$f_l = \sum_{(C_l)} W(C_l) \quad (2.5)$$

where the sum runs over all connected graphs of l points.

Theorem I then states:

$$1 + F(x) = \exp f(x) \quad (2.6)$$

where $F(x)$, $f(x)$ are the generating functions (counting series):

$$F(x) = \sum_{N=1}^{\infty} F_N \frac{x^N}{N!}; \quad f(x) = \sum_{l=1}^{\infty} f_l \frac{x^l}{l!}$$

The proof is based on the following lemma:

Product theorem. The generating function for the collection of labeled graphs, consisting of any graph of one collection plus any graph of a second collection, is the product of the generating functions of the two collections taken separately (always assuming that the weights have the properties *a*) and *b*). This may seem almost obvious, but the proof (which we omit) really needs some thought. For the validity of the theorem it is essential that the generating functions are defined with the factorials in the denominators. If $F_m(x)$ is the counting series for graphs of m disconnected parts, then

$$F(x) = \sum_{m=1}^{\infty} F_m(x),$$

and

$$F_m(x) = \frac{1}{m!} [f(x)]^m.$$

$m!$ is required because the m parts are chosen from the *same* collection of connected graphs, and any permutation of the m parts leads to the same disjoint graph. From this eq. (2.6) follows immediately. (Eq. (2.6) has been rediscovered many times!) $W(G)$ is arbitrary and could even be an operator. We now give two specializations of this theorem.

a) Take

$$W(G) = y^k$$

where k = number of lines of graph G and y is an arbitrary variable. Theorem I then yields the *first Riddell formula*:

$$1 + N(x, y) = \exp C(x, y), \quad (2.7a)$$

where

$$N(x, y) = \sum_{p=1}^{\infty} \frac{x^p}{p!} \sum_{k=0}^{\frac{1}{2}p(p-1)} N_{pk} y^k, \quad (2.7b)$$

$$C(x, y) = \sum_{p=1}^{\infty} \frac{x^p}{p!} \sum_{k=p-1}^{\frac{1}{2}p(p-1)} C_{pk} y^k. \quad (2.7c)$$

N_{pk} = number of labeled (p, k) graphs (graphs of p points, k lines),

$$N_{pk} = \binom{\frac{1}{2}p(p-1)}{k}.$$

C_{pk} = number of labeled, connected (p, k) graphs.

Equations (2.7) constitute a set of functional relations, which can be solved to find C_{pk} in terms of the known N_{pk} . It can be shown that, if $k > p \log p$, C_{pk} and N_{pk} are asymptotically equal for large p .

b) The connection of the general result with the Mayer theory is established by taking:

$$W(G_N) = \lambda^{-3N} \int_V \dots \int dr_1 \dots dr_N \prod_{G_N} f_{ij}, \tag{2.8}$$

where the product is taken over all pairs of points (i, j) connected by a line in G_N . Writing (for historical reasons)

$$b_l(V, T) = \frac{1}{Vl!} \int_V \dots \int dr_1 \dots dr_l \sum_{(C_l)} \prod_{(C_l)} f_{ij} \tag{2.9}$$

and introducing the cluster integral function

$$\chi(V, T, z) = \sum_{l=1}^{\infty} b_l(V, T) z^l,$$

one then gets for the grand-partition function:

$$\mathcal{Z}(V, T, z) = \sum_{N=0}^{\infty} Z(V, T, N) (\lambda^3 z)^N$$

with $Z(V, T, N)$ as defined in eq. (2.1) and, conventionally, $Z(V, T, 0) \equiv 1$, from theorem I the 1st Mayer theorem:

$$\mathcal{Z}(V, T, z) = \exp\{V\chi(V, T, z)\} \tag{2.10}$$

From eq. (2.10) one gets in the usual way in the gas phase (for $N, V \rightarrow \infty$, $v = V/N$ finite):

$$\left. \begin{aligned} p/kT &= \bar{\chi}(T, z_0) \\ \frac{1}{v} &= z_0 \frac{\partial}{\partial z_0} \bar{\chi}(T, z_0) \end{aligned} \right\} \tag{2.11}$$

$$\bar{\chi} = \sum_{l=1}^{\infty} \bar{b}_l(T) z^l; \quad \bar{b}_l(T) = \lim_{N \rightarrow \infty} b(Nv, T)$$

It should be noted that this reduction of the graphs occurring in the Mayer theory is independent of the assumption of additive forces; since for non-additive forces properties a) and b) for $W(G)$ still hold, Eqs. (2.10) and (2.11) remain valid in this case. Note also that for (2.10) there is no need to pass to the limit of an infinite system.

Theorem II. We now study the reduction of connected graphs. Suppose the weight $W(C_p)$ has the following properties:

- a. $W(C_p)$ is independent of labeling
- b. $W(C_p) = \prod_{(S_m)} W(S_m)$

where S_m is a labeled star of m points occurring in the connected graphs C_p . Let

$$f_p = \sum_{(C_p)} W(C_p), \quad r_m = \sum_{(S_m)} W(S_m). \tag{2.12}$$

Then theorem II states:

$$T(z) = z \exp \frac{dS(T)}{dT}$$

where

$$T(z) = z \frac{df}{dz} = \sum_{p=1}^{\infty} pf_p \frac{z^p}{p!} \quad (2.13)$$

$$S(y) = \sum_{m=2}^{\infty} r_m \frac{y^m}{m!}$$

are the counting series for the *rooted* connected graphs and for the stars.

Proof: Consider first for simplicity pure star trees, where the constituent star has g points, weight W , and symmetry number s , and write:

$$T(z) = \sum_{n=0}^{\infty} T_n(z) \quad (2.14)$$

where $T_n(z)$ is the counting series for rooted graphs with n main leaves; $T_0(z) \equiv z$.

Now we can construct rooted star trees with n main leaves by selecting n rooted star trees with *one* main leaf and hanging them all on the root. From the product theorem follows:

$$T_n(z) = \frac{z}{n!} \left(\frac{T_1(z)}{z} \right)^n. \quad (2.15)$$

Since $T_1(z)/z$ is the counting series for rooted star trees with one main leaf in which the root point is erased; the factor z accounts for the root and the $n!$ is needed since the n root points are in fact identical. To find $T_1(z)$ we start with the main leaf and hang rooted star trees on the $(g-1)$ points other than the root. Again from the product theorem,

$$T_1(z) = \frac{zqW}{s} T^{g-1}(z) \quad (2.16)$$

The factor z accounts for the root, and W accounts for the weight of the main leaf; the factor q is the number of ways in which we can select the root point and the factor $1/s$ is needed since there are s equivalent arrangements of root and rooted star trees on the main leaf because of its symmetry. Hence

$$T(z) = z \exp \left\{ \frac{qW}{s} T^{g-1} \right\} \quad (2.17)$$

Furthermore, in this case $S(y)$ consists of only *one* term $r_q y^q / q!$ with $r_q = (q!/s)W$ since there are $q!/s$ different labeled stars of the given type. Hence $S(y) = Wz^q/s$, so that (2.17) can be written in the form (2.13). The generalisation to mixed star trees is straightforward. Finally that

$T(z) = zdf/dz$ is clear, because each labeled graph of p points corresponds to p rooted labeled graphs since each point can in turn be preferred as the root.

Applications of theorem II.

a) By taking again

$$W(C) = y^k$$

($k =$ number of lines) one obtains from (2.13) the 2nd Riddell formula:

$$\begin{aligned} z(x, y) &= x \frac{\partial}{\partial x} C(x, y) \\ \frac{\partial}{\partial z} S(z, y) &= \ln \frac{z}{x} \end{aligned} \tag{2.18}$$

with

$$S(x, y) = \sum_{p=1}^{\infty} \frac{x^p}{p!} \sum_{k=p}^{\frac{1}{2}p(p-1)} S_{pk} y^k.$$

S_{pk} is the number of labeled stars of p points, k lines. From (2.18) one can find S_{pk} in terms of C_{pk} and one can show that, if $k > p \log p$, S_{pk} and C_{pk} are asymptotically equal for large p . Hence for large p , $k > p \log p$, the overwhelming majority of graphs are stars.

b) One can use (2.13) by taking $W = 1$, to count all kinds of labeled star trees. For example from (2.17) follows for the counting series of rooted labeled Cayley trees the functional equation $T(z) = z \exp T(z)$, from which one derives that there are p^{p-2} different labeled Cayley trees of p points. Similar results one can find for Husimi trees.

c) For additive forces the weight

$$W(C_l) = \lim_{V \rightarrow \infty} \frac{1}{V} \int_V \dots \int_V dr_1 \dots dr_l \prod_{C_l} f_{ij}$$

fulfills the two requirements and therefore theorem II is applicable. $T(z)$ becomes

$$z \frac{d\bar{\chi}}{dz} = \frac{1}{v}$$

and calling

$$\bar{\beta}_{m-1} = \lim_{V \rightarrow \infty} \frac{1}{V(m-1)!} \int_V \dots \int_V dr_1 \dots dr_m \sum_{(S_m)} \prod_{S_m} f_{ij} \tag{2.19}$$

one gets

$$S'(y) = \frac{d}{dy} \sum_{m=2}^{\infty} \frac{y^m}{m} \bar{\beta}_{m-1} = \sum_{v=1}^{\infty} \bar{\beta}_v y^v \equiv \varphi(y) \tag{2.20}$$

From (2.13) one now obtains the 2nd Mayer theorem:

$$\frac{1}{v} = z e^{\varphi(1/v)} \tag{2.21}$$

$$z = x e^{-\varphi(x)}, \quad x \equiv 1/v$$

which expresses formally z in x , or inverts the second Mayer eq. $1/v = z d\bar{\chi}/dz$. From the first Mayer equation one then gets

$$\begin{aligned} \frac{p}{kT} &= \chi(z) = \int_0^z \frac{x}{z} dz = \int_0^x e^{\varphi(x)} d(xe^{-\varphi(x)}) \\ &= x - \int_0^x x\varphi'(x) dx \\ &= \frac{1}{v} - \sum_{\nu=1}^{\infty} \frac{\nu}{\nu+1} \beta_{\nu} \left(\frac{1}{v}\right)^{\nu+1}; \end{aligned} \quad (2.22)$$

which is the virial expansion.

3. *Further developments.* The two Mayer theorems reduce the problem of calculating the partition function to the calculation of the β_{ν} , i.e., to the star integrals. Any further reduction must make distinctions between different types of stars. It seems reasonable to classify the stars somehow according to their complexity or connectivity, because the more complex or connected the stars are, the smaller the corresponding star integral will be. Dr. Meeron has gone farthest in this direction and since we will hear about these developments from Prof. J. de Boer I will only give here some of the definitions and simple theorems.

a) The *connectivity* $\kappa(G)$ of a graph G is the minimum number of points, whose omission together with the lines incident upon them, would separate the graph. For a complete graph of N points, the connectivity, by definition, is $N - 1$. For disconnected graphs $\kappa = 0$. For graphs with an articulation point $\kappa = 1$; $\kappa \geq 2$ for stars.

From the (few) known theorems we only mention:

1) $\kappa(G) \leq$ minimum degree of the points of G . The more evenly the degrees of the points are distributed, the higher the connectivity.

2) For a (p, k) graph,

$$\max \kappa = \begin{cases} 0 & \text{if } k < p - 1; \text{ disconnected graph,} \\ 1 & \text{if } k = p - 1; \text{ the graph is a Cayley tree,} \\ \left[\frac{2k}{p} \right] & \text{if } k > p - 1. \end{cases}$$

$$\min \kappa = \begin{cases} 0 & \text{if } k \leq \binom{p-1}{2}, \\ k - \binom{p-1}{2} & \text{if } k > \binom{p-1}{2}. \end{cases}$$

This theorem, which is due to Harary, gives information on the range of values of κ ; but unfortunately we know nothing about the distribution of κ , or even its average value.

b) The *complexity* $d(C_p)$ of a graph C_p is defined in terms of the graph matrix $d_{ij}(C_p)$. The graph matrix $d_{ij}(C_p)$ is a $p \times p$ matrix with elements

defined by

$$d_{ij}(C_p) = \begin{cases} -1 & \text{if the line } (i, j) \text{ occurs in } C_p, \\ 0 & \text{if the line } (i, j) \text{ does not occur in } C_p. \end{cases}$$

$$d_{ii}(C_p) = \text{degree of point } i \text{ in } C_p.$$

It is easy to see that the determinant $\|d_{ij}\| = 0$ and that all minors of order $p - 1$ are equal. Their common value $d(C_p)$ is called the graph complexity.

1) The relation of the complexity $d(C_p)$ to the integral problem is obvious if one takes in the Mayer weight function (2.8) the Gaussian function

$$f_{ij} = - e^{-\alpha|r_i-r_j|^2} \tag{3.1}$$

which does not correspond to a real physical potential, but which should represent qualitatively a soft repulsive potential. One now finds

$$W(C_p) = (-1)^k \left(\frac{\pi}{\alpha}\right)^{\frac{1}{2}(p-1)} [d(C_p)]^{-\frac{1}{2}}. \tag{3.2}$$

This result shows directly, that the larger the complexity of the graph, the smaller the corresponding integral.

2) $d(C_p)$ = number of labeled Cayley trees which are subgraphs of C_p (Kirchhoff). Taking for C_p the complete graph gives therefore for the total number of labeled Cayley trees of p points

$$\begin{vmatrix} p-1 & -1 & & -1 \\ -1 & p-1 & & -1 \\ & & \ddots & \\ -1 & -1 & & p-1 \end{vmatrix} = p^{p-2}.$$

3) The sum of the complexities for all (p, k) graphs is given by

$$\sum_{(C_{pk})} d(C_{pk}) = p^{p-2} \binom{(p-1)(p-2)}{k-p+1}. \tag{3.3}$$

This result is due to G. W. Ford. Since one also knows the number C_{pk} of labeled (p, k) graphs, eq. (3.3) gives the average value of the complexity for these graphs. By direct computation of the complexity for all graphs up to seven points, Ford has found empirically that the distribution of the complexity (for k in the middle of the range between $p - 1$ and $p(p - 1)/2$) is gaussian around the average value given by (3.3). It would be very valuable if this could be proved and if one could calculate the variance of the distribution.

4. *Application to plasma's (fully ionized gases).* Of course with long range forces the virial expansion has no sense and should be replaced by the so-

called Debye expansion, in which the small parameter is $\kappa v^{\frac{1}{2}}$; $1/\kappa$ is the Debye shielding radius: $\kappa^2 = (4\pi e^2/kT) \sum_{\sigma} Z_{\sigma}^2 n_{\sigma}$; $n_{\sigma} = N_{\sigma}/V =$ concentration of ions of type σ ; v is again $v = V/N = V/\sum_{\sigma} N_{\sigma}$.

Electrical neutrality $\sum_{\sigma} Z_{\sigma} n_{\sigma} = 0$ is always assumed.

It would take too long to discuss this expansion in detail, and we only want to draw attention to one aspect of the theory (due to Mayer, 1950), since it is probably the first example of the partial summing technique. Consider a plasma of different types σ of ions. In the partition function there occurs the potential energy

$$\varphi_{ij}^{\sigma\tau} = 4\pi e^2 Z_{\sigma} Z_{\tau} \Phi(R) + \Psi_{\sigma\tau}(R), \quad R = |r_i - r_j| \tag{4.1}$$

where $\Psi_{\sigma\tau}(R)$ is the short-range potential (range $a_{\sigma\tau}$), and where Φ is a (formally) damped Coulomb potential,

$$\Phi(R) = \frac{e^{-\alpha R}}{4\pi R}, \quad \alpha > 0 \tag{4.2}$$

Because of this damping one has formally made the Coulomb potential also short range and hence one can start in the same way as in the virial development. However one now expands in e^2 , so that

$$f_{ij}^{\sigma\tau} = g_{\sigma\tau}(R) + [1 + g_{\sigma\tau}(R)] \sum_{n=1}^{\infty} \frac{(-Z_{\sigma} Z_{\tau} \lambda \Phi)^n}{n!}$$

with

$$\lambda = \frac{4\pi e^2}{kT} \tag{4.3}$$

$$g_{\sigma\tau}(R) = e^{-\Psi_{\sigma\tau}(R)/kT} - 1$$

Each star is now *doubly-labeled* (according to type σ , and according to the number in each type), and if one substitutes (4.3) for $f_{ij}^{\sigma\tau}$ each star is blown up in an infinite set of (not simple) stars by interpreting: a) $g_{ij}^{\sigma\tau}$ by a dotted line say; and b) $-Z_{\sigma} Z_{\tau} \lambda \Phi$ by a solid line.

For example (fig. 3):

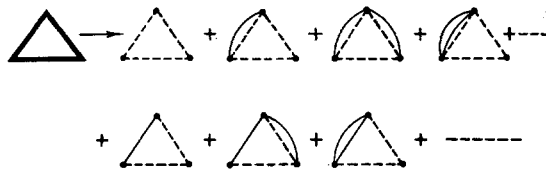


Fig. 3.

Dotted lines (g -bonds) can occur only zero or one times, whereas the solid lines (Φ -bonds) may occur any number of times. The $n!$ is the symmetry number of the n -tuple Φ -bond. Or course, if $\alpha \rightarrow 0$ all weights diverge. But now Mayer considers together all graphs which are *homeomorphic* with respect to the Φ -bonds. In general two graphs are homeomorphic if they differ

only with respect to the number of points of degree two. Thus (fig. 4):

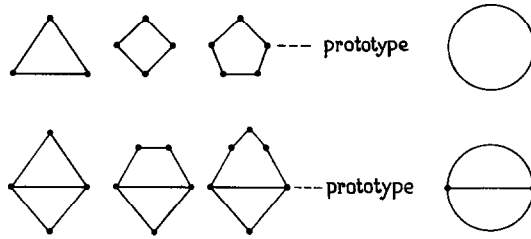


Fig. 4.

The reason why this is a good idea lies in the fact that for a *chain* the integral is of the folding type, so that the integration over the intermediate points can easily be performed. Especially one easily proves the following theorem. Let

$$\Phi_n(r_{ij}) = \int_V \dots \int dr_1 \dots dr_n \Phi(r_{i1}) \Phi(r_{12}) \dots \Phi(r_{nj}); \tag{4.4}$$

then

$$\lim_{\alpha \rightarrow 0} \sum_{n=0}^{\infty} (-\kappa^2)^n \Phi_n(r_{ij}) = \frac{e^{-\kappa|r_i - r_j|}}{4\pi|r_i - r_j|} \equiv \Phi_D(R). \tag{4.5}$$

In this way the Debye potential enters the development. In addition it is easy to see that:

- a) The contribution of the single Φ -bond or the graph \cdots gives zero because of electric neutrality;
- b) The contribution of the cycle of Φ -bonds gives the first Debye-correction to the free energy or equation of state.

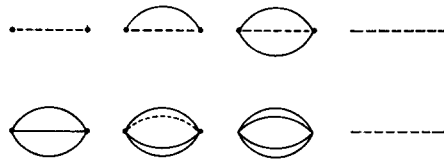


Fig. 5.

There remain the blown-up stars, as for instance (fig. 5) where now each solid line represents a Debye bond: $-Z_\sigma Z_\tau \lambda \Phi_D(R)$. We will not discuss these terms further and mention only that in this way one obtains in a systematic way corrections to the well-known Debye limit laws. The short range potential is essential, since otherwise one gets the also well-known divergences for $R \rightarrow 0$. There are now of course several parameters in the expansion, and one must assume that:

$$a_{\sigma\tau} \approx \lambda \ll v^{-\frac{1}{2}} \ll 1/\kappa \tag{4.6}$$

to assure the practicality of the combined virial and Debye expansion.