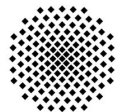

An Introduction to Graph Theory, the Mayer Cluster Expansion and M.S. Wertheim's Multiple Density Approach

Thomas Bissinger



University of Stuttgart
Germany

Forschungsarbeit supervised by

Prof. Dr.-Ing. J. Groß

M. Sc. Wasilios Zmpitas

Institut für Technische Thermodynamik, Universität Stuttgart, December 2015

An Introduction to Graph Theory, the Mayer Cluster Expansion and M.S. Wertheim's Multiple Density Approach

T. Bissinger

December 2, 2015

Abstract

In this work, we give an introduction to the fundamentals of graph theory. We explore M.S. Wertheim's multiple density formalism in analogy to the procedure of the Mayer cluster expansion. The parameter λ' in the second order TPT approximation by Wertheim is fitted to simulation results to obtain more accurate approximations than with the former fit.

1 Introduction

This introductory article is meant to give the key ingredients to understanding M.S. Wertheim's theory, presented in references^{1,2,3,4,5}. To that end, Section 2 will introduce the fundamental ideas of graph theory, which will be applied to the grand canonical partition function to derive the *Mayer Cluster Expansion* (MCE) in Section 3. Since Wertheim follows an analogous argumentation to the MCE, Section 4 will simply illustrate the major variations proposed by Wertheim and explain their advantages. Section 5 uses *Thermodynamic Perturbation Theory* (TPT) to yield approximations to compressibility values. These results are compared and fitted to simulation results. Section 6 gives concluding remarks and an outlook.

2 Graph Theory

2.1 Basic Definitions

Fundamental to everything that follows is the basic notion of a graph. This section shall give a definition of various common elements of graphs and mention three important lemmas that are often used in graph theory. To simplify the application in the following sections, we will use physically interpretable function names. However, the theory is purely mathematical and does not require any physical argumentation by itself.

The basic notion of a graph is described in a first definition.

Definition 1

Let $\mathbf{r} \in \Omega$ with a domain $\Omega \subset \mathbb{R}^d$. Let $z_i : \Omega \mapsto \mathbb{R}$ and $f_\mu : \Omega \times \Omega \mapsto \mathbb{R}$, $i, \mu \in \mathbb{N}$ be functions in one and two variables, respectively. We call z_i the properties and f_μ the bonds. With them, the following equations hold:

- i. A labelled point with a property corresponds to a function in one variable:

$$z_1 \bigcirc_{\mathbf{r}_1} = z_1(\mathbf{r}_1). \quad (1)$$

- ii. Two labelled points with a bond correspond to a function in two variables:

$$z_1 \bigcirc_{\mathbf{r}_1} \xrightarrow{f_\mu} z_2 \bigcirc_{\mathbf{r}_2} = z_1(\mathbf{r}_1) f_\mu(\mathbf{r}_1, \mathbf{r}_2) z_2(\mathbf{r}_2). \quad (2)$$

- iii. This can be generalized as

{A graph with n labelled points and K bonds with $m(\mu, ij)$ bonds of type μ connecting points i and j }

$$= \prod_{i=1}^n z_i(\mathbf{r}_i) \prod_{j>i}^n \prod_{\mu=1}^K \frac{1}{m(\mu, ij)!} f_\mu(\mathbf{r}_i, \mathbf{r}_j)^{m(\mu, ij)}. \quad (3)$$

A quick remark on the last item of the definition: if there is no bond of type μ present between points i and j , $m(\mu, ij)$ will be zero and the contribution of the bond-term in the product will be one. For all application we will discuss in this article, $m(\mu, ij)$ will be one or zero. Some authors, including for example Stell⁶, call such graphs were not more than one bond between two particles is allowed *simple graphs*. Since we will only meet such simple graphs, we keep the nomenclature short by just calling them graphs.

As a first advantage of graph theory, we can see that large products of functions can be simplified by replacing them with rather simple drawings. The true power within this ansatz is the easy manipulation of complicated (sums of) functions by set theoretic or topological considerations on graphs.

Note that the label of a labelled point is not the property z_i but the fact that a fixed variable \mathbf{r} is assigned to that point. For

brevity, the names of properties, labels or bonds can be omitted when it is clear from the context. Sometimes different graphical representations are used to distinguish between bond types or properties.

The next definition introduces integration into graph theory.

Definition 2

A field point is a filled point (usually without a label) and means integration over the coordinate:

$$i. \quad \text{●}^{z_1} = \int_{\Omega} d\mathbf{r}_1 z_1(\mathbf{r}_1). \quad (4)$$

$$ii. \quad \text{●}^{z_1} \text{---}_{f_{\mu}} \text{●}^{z_2} = \int_{\Omega} d\mathbf{r}_2 \int_{\Omega} d\mathbf{r}_1 z_1(\mathbf{r}_1) f_{\mu}(\mathbf{r}_1, \mathbf{r}_2) z_2(\mathbf{r}_2). \quad (5)$$

iii. And the general version:

$$\begin{aligned} & \{ \text{A graph with } n \text{ labelled points and } m \text{ field points} \\ & \text{and arbitrary bonds} \} \\ & = \int_{\Omega} d\mathbf{r}_{n+m} \cdots \int_{\Omega} d\mathbf{r}_{n+1} \times \{ \text{the same graph with} \\ & \text{all field points carrying label } \mathbf{r}_{n+1} \text{ to } \mathbf{r}_{n+m} \}. \end{aligned} \quad (6)$$

With graphs, the interdependences of integration variables are clearly visible. This helps to easily detect *steric incompatibilities*, e.g. graphs vanishing because their geometry always results in a value of zero. In the hard spheres case we will study, this is always due to overlapping spheres.

The variables $\mathbf{r}_1, \dots, \mathbf{r}_N$ are often abbreviated by $1, \dots, N$. This is also done if the orientations $\Omega_1, \dots, \Omega_N$ of the particles play an important role (i.e. if they are non-spherical), where the variable i then contains both \mathbf{r}_i and Ω_i .

2.2 Graph Symmetry

With the above definition, labels of field points can be set arbitrarily or omitted entirely, all resulting in the same graph. This is in agreement with the mathematical intuition that integration variables can be named arbitrarily as long as they stay well-defined and distinguishable. However, we will see in Section 3 that when using graphs to analyze the grand canonical ensemble, the prefactors of the arising graphs play an important role. Therefore, some authors (e.g. Morita & Hiroike in their definitions 3 and 3'⁷) use a definition that distinguishes between graphs with labelled field points and graphs with unlabelled field points. We will also give this definition for completeness, but note that we will not stick to it. For a deeper understanding, we will consider the graph

$$\begin{array}{c} \text{●} \\ | \\ \text{●} \text{---} \text{●} \end{array} := \begin{array}{c} z \text{●} \\ | \\ z \text{●} \text{---}_{f} z \text{●} \end{array}. \quad (7)$$

We will use the shorthand form on the left hand side of (7) to keep the overview. Obviously, as long as we label the field points in a distinguishable manner, their labels are arbitrary and only serve as dummy names for the integration process. It is, however, convenient to introduce an equivalency relation between two graphs with labelled points and field points carrying labels: two graphs are *topologically indistinguishable* if points with the same label i carry the same properties and pairs of points with labels i and j are linked by the same bonds in both graphs. For example,

$$\begin{array}{c} \text{●}_3 \\ | \\ \text{●}_1 \text{---} \text{●}_2 \end{array} \cong \begin{array}{c} \text{●}_3 \\ | \\ \text{●}_2 \text{---} \text{●}_1 \end{array}. \quad (8)$$

where the \cong sign means topological indistinguishability. Conversely, two graphs are topologically distinguishable – indicated here by the $\not\cong$ sign – if the bonds or properties assigned to their points differ, as in

$$\begin{array}{c} \text{●}_3 \\ | \\ \text{●}_1 \text{---} \text{●}_2 \end{array} \not\cong \begin{array}{c} \text{●}_2 \\ | \\ \text{●}_1 \text{---} \text{●}_3 \end{array}, \quad (9)$$

The numerical value represented by all graphs in (8) and (9) is the same, and we did already note above that from a mathematical point of view the labelling of the integration variables is of no importance. So why is this distinction necessary at all?

There are multiple answers to that question. First of all, there is no unique way to transform some expression containing properties, bonds and integrals into a graph. Mathematically, the expressions

$$\begin{array}{c} \text{●}_3 \\ | \\ \text{●}_1 \text{---} \text{●}_2 \end{array} \cong \begin{array}{c} \text{●}_1 \text{---} \text{●}_3 \text{---} \text{●}_2 \end{array} \quad (10)$$

represent the exact same object. This is a first reason to know which graphs are actually the same even if they may appear to be different at first glance.

Another answer to this question lies in the application of graph theory to problems of physics. When applying graph theory to the grand canonical partition function in Section 3, we will encounter sums over graphs. These will only contain topologically distinguishable graphs, and a graph of N points will be weighted by a factor of $(N!)^{-1}$. When only considering the graphs in the example above, the sum will contain the expres-

$$\begin{aligned} & \frac{1}{6} \left\{ \begin{array}{c} \bullet_3 \\ \bullet_1 \text{---} \bullet_2 \end{array} + \begin{array}{c} \bullet_2 \\ \bullet_3 \text{---} \bullet_1 \end{array} + \begin{array}{c} \bullet_1 \\ \bullet_2 \text{---} \bullet_3 \end{array} \right\} \\ &= \frac{1}{2} \left\{ \begin{array}{c} \bullet_3 \\ \bullet_1 \text{---} \bullet_2 \end{array} \right\}. \end{aligned} \quad (11)$$

The symmetry number is an important property of a graph. Some authors include it directly in the basic definition of a graph, thereby distinguishing between graphs with labelled field points and graphs with unlabelled field points. We will not do so, which leads to different equations.

While integration by definition turns a labelled point into a field point, turning a field point back into a labelled point requires a slight bit of analysis. For that purpose, we introduce the variational calculus, and thereby differentiation, into the graph theoretic framework.

Let z be defined as in Definition 1 and let $\Lambda : z \mapsto \mathbb{R}$ be a functional. Then the functional derivative $\delta\Lambda/\delta z$ is defined via the variation $\delta\Lambda$ of Λ and δz of z by

Integration over the label of a labelled point is of course a functional of the property assigned to that labelled point (and the bonds adjacent to it, but we will not use that here). It is therefore an excellent candidate to use the above definition. We will

$$\frac{\delta z(\mathbf{r})}{\delta z'(\mathbf{r}')} = \begin{cases} 0 & \text{for } z \text{ independent of } z', \\ \delta(\mathbf{r} - \mathbf{r}') & \text{for } z \equiv z'. \end{cases} \quad (13)$$
$$\begin{aligned}
& \frac{\delta}{\delta z(\mathbf{r})} \frac{z}{f} \\
&= \int d\mathbf{r}' \int d\mathbf{r}'' \frac{\delta}{\delta z(\mathbf{r})} [z(\mathbf{r}') f(\mathbf{r}', \mathbf{r}'') z(\mathbf{r}'')] \\
&= \int d\mathbf{r}' \int d\mathbf{r}'' [\delta(\mathbf{r} - \mathbf{r}') f(\mathbf{r}', \mathbf{r}'') z(\mathbf{r}'')] \\
&\quad + \int d\mathbf{r}' \int d\mathbf{r}'' \delta(\mathbf{r} - \mathbf{r}'') f(\mathbf{r}', \mathbf{r}'') z(\mathbf{r}') \\
&= 2 \int d\mathbf{r}' f(\mathbf{r}, \mathbf{r}') z(\mathbf{r}') \\
&= 2 \frac{z}{f_{\mu}}.
\end{aligned} \tag{14}$$

But as is the case in equation (14), this would result in summing many graphs with equal value. This redundancy stems from the aforementioned fact that our Definition 2 has no special treatment of graph symmetry. The derivative changes the symmetry number of the graph from 2 to 1. Devision by symmetry numbers yields a useful lemma, which we will give without a prove.

Let z, z' be properties and let Γ be a graph and let the set $\{\Gamma_{\alpha}(\mathbf{r})\}_{1 \leq \alpha \leq m}$ be the set of all topologically distinguishable graphs obtainable from Γ by turning one field point with property z into a labelled point depending on \mathbf{r} with property z' . Let s be the symmetry number of Γ and let the set $\{s_{\alpha}\}_{1 \leq \alpha \leq m}$ be the set of symmetry numbers corresponding to $\{\Gamma_{\alpha}(\mathbf{r})\}$. Then

Of course, if Γ has no z field points, the derivative amounts to zero as does the sum over the empty set on the right hand side. Note again that Lemma 4 would look differently if all graphs already contained division by symmetry number.

2.4 Articulation Points

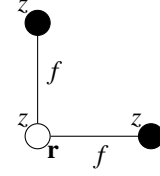
Most considerations on graphs take only the topology of a graph into account. From the topology, it is possible to derive reductions which can ultimately result in huge simplifications of originally very complicated expressions.

For example, finding independent integration variables can be easily translated to a topological consideration. That is this subsection's purpose, and we will only need a short list of definitions to get there. The following properties are relevant in graph theory:

- *connected graph*: a graph Γ with two or more points is connected if for any points $A, B \in \Gamma$ there is a path of points and bonds connecting A to B .
- *doubly, triply, ... n -connected graph*: a graph Γ with n or more points is n -connected if for any points $A, B \in \Gamma$ there are n independent paths of points and bonds connecting A to B .
- *articulation point*: a point A in a connected graph Γ . If one removes A and its adjacent bonds, the graph Γ will be divided into two or more separate graphs.
- *reducible graph*: a graph that contains articulation points or bonds that may be expressed in a more convenient way.

While these names are frequently used in graph theory, they may be understood differently in each context. For example, the vague definition of a reducible graph comes from the fact that it is up to the researcher which bonds they want to consider as fundamental for their graph. As one can see especially in the paper Wertheim III³, the correct definition in this matter is crucial to the types of graphs one ends up analyzing. Wertheim III also shows that careful characterizations of articulation points are fundamental to researches in graph theory.

Now, let us return to independent integration variables. Their treatment is easiest understood in a simple exemplary calculation considering a three point singly connected graph. For simplicity, we will consider all properties to be z and all bonds to be f .



$$\begin{aligned}
 &= \int_{\Omega} d\mathbf{r}' \int_{\Omega} d\mathbf{r}'' z(\mathbf{r})z(\mathbf{r}')z(\mathbf{r}'')f(\mathbf{r},\mathbf{r}')f(\mathbf{r},\mathbf{r}'') \\
 &= z(\mathbf{r}) \int_{\Omega} d\mathbf{r}' \int_{\Omega} d\mathbf{r}'' z(\mathbf{r}')f(\mathbf{r},\mathbf{r}')z(\mathbf{r}'')f(\mathbf{r},\mathbf{r}'') \quad (16) \\
 &= z(\mathbf{r}) \left\{ \int_{\Omega} d\mathbf{r}' z(\mathbf{r}')f(\mathbf{r},\mathbf{r}') \right\} \\
 &\quad \times \left\{ \int_{\Omega} d\mathbf{r}'' z(\mathbf{r}'')f(\mathbf{r},\mathbf{r}'') \right\} \\
 &= z(\mathbf{r}) \times \left\{ \text{graph with two points } z \text{ and } z \text{ connected by bond } f \right\}^2.
 \end{aligned}$$

This simple example can be used to illuminate a few ideas. First of all, we used the independence of integration variables in line 3 to reach line 4. On the other hand, the original graph and the resulting graph can be linked with the following procedure if we consider that the labelled point is an articulation point: Separate the graph at the articulation point. Each of the resulting graphs may be computed with the articulation point replaced by a point without properties but with the same label as the labelled point. The original graph is the product of all graphs resulting from the separation times the articulation point with the corresponding property.

Another idea underlying the transformation in equation (16) is the fact that identifying independent subgraphs – independent in a sense that they are only linked by articulation points – can simplify graph computation.

A third remark concerns what simplifications may be made after arriving at the right hand side of equation (16). We did already reduce the number of integrations to be carried out. But we could also go further and define a new property according to

$$z^*(\mathbf{r}) := z(\mathbf{r}) \times \left\{ \text{graph with two points } z \text{ and } z \text{ connected by bond } f \right\}^2, \quad (17)$$

reducing the full graph altogether to

$$\begin{aligned}
 &\text{graph with three points } z, z, z \text{ and bonds } f, f \\
 &= z^*(\mathbf{r}). \quad (18)
 \end{aligned}$$

Systematically used, this simple trick is again capable of reducing large graphs in a straightforward way, and we will in fact

use it in the Mayer cluster expansion.

In the above example, the articulation point was a labelled point. We chose this setting to clearly show the \mathbf{r} -dependence in the reduced graphs. Of course, the same formalism can be used on a field articulation point, but some care has to be taken when treating the integration variable of the field point. If one turned all labelled points in equation (18) into field points, the equality would still hold, while when turning all labelled points in equation (16) into field points, the expressions would not necessarily be equal.

2.5 Exponentials of Graphs

At the end of this section, we give two very powerful lemmas which are crucial for many applications of graph theory. A proof can be found in Morita and Hiroike⁷ (Lemmas 3 and 4), we will just give the basic idea behind it. The first lemma is

Lemma 5

Consider two sets of graphs of field points, $A := \{\Gamma_\alpha\}$ and $B := \{\Lambda_\beta\}$, with corresponding sets of symmetry numbers $\{s_\alpha\}$ and $\{\tilde{s}_\beta\}$. Let B contain all connected subgraphs of A and the single field point graph. If any $\Gamma_\alpha \subset A$ is a product of graphs $\Lambda_\beta \subset B$ (with a graph possibly appearing multiple times or not at all), and any product of graphs $\Lambda_\beta \subset B$ is represented by exactly one $\Gamma_\alpha \subset A$, then

$$1 + \sum_{\alpha \mid \Gamma_\alpha \in A} s_\alpha^{-1} \Gamma_\alpha = \exp \left(\sum_{\beta \mid \Lambda_\beta \in B} \tilde{s}_\beta^{-1} \Lambda_\beta \right). \quad (19)$$

Note that a set A , in order to fulfill the prerequisites of the previous lemma, has to be infinite. This does not apply to the set B .

To understand Lemma 5, we consider the exponential function. If one does not consider factors, the exponential function is 1 plus a sum of all powers of the exponent. The 1 is already on the left hand side of (19), so only the sum remains. Since each graph in A is a product of graphs in B , and A contains all products of graphs in B , summing over all graphs in A must contain the same graphs as the exponential function of a sum of all graphs in B . One can prove that the sums are exactly the same if one examines the factors resulting from the symmetry numbers and the factors $(N!)^{-1}$ in the exponential series.

A similar lemma which uses articulation points is

Lemma 6

Consider two sets of connected graphs $A := \{\Gamma_\alpha\}$ and $B := \{\Lambda_\beta\}$ with corresponding sets of symmetry numbers $\{s_\alpha\}$ and $\{\tilde{s}_\beta\}$. The graphs in A and B do all contain exactly the same labelled points without property and an arbitrary number of field points. Let B contain all the subgraphs of A that stay connected if the labelled points are

removed. If any $\Gamma_\alpha \subset A$ can be uniquely decomposed into (possibly identical) graphs belonging to B by separating Γ_α at the labelled points, and any graph obtained by joining together (possibly identical) graphs from B at the labelled points uniquely belongs to A , then

$$1 + \sum_{\alpha \mid \Gamma_\alpha \in A} s_\alpha^{-1} \Gamma_\alpha = \exp \left(\sum_{\beta \mid \Lambda_\beta \in B} \tilde{s}_\beta^{-1} \Lambda_\beta \right). \quad (20)$$

The proof for this lemma is exactly the same as for Lemma 5, since the process of joining together graphs at labelled points without property is again a simple graph multiplication, so the arguments for the power series in the exponential function and the inclusion of symmetry numbers need not be changed.

An infinite sum of graphs can therefore be expressed by the exponential of a sum of elemental subgraphs, which may be finite. This result will prove useful in the following sections.

3 The Mayer Cluster Expansion

3.1 Fundamentals

The MCE uses graph theory to tackle statistical mechanics. Our task is to acquire thermodynamic properties of a system in thermal equilibrium which contains identical particles moving in the fixed volume V at the fixed temperature T and the fixed chemical potential μ . The system is to be considered uniform, that is all properties have the same value throughout the entire system. The properties of interest may for example be the *pressure* p or the *density* ρ . To that end, we consider the graph theoretical approach to obtain a useful expression for the *grand canonical partition function* Ξ ,

$$\Xi(\mu, V, T) = \sum_{N \geq 0} e^{N\mu/k_B T} Q(N, V, T) \quad (21)$$

with the *Boltzmann constant* k_B and the *canonical partition function* Q . The sum runs over all *numbers of particles* N contained in V . A common replacement is

$$\beta := \frac{1}{k_B T}, \quad (22)$$

and we can define the *fugacity* by

$$\tilde{z} := e^{\beta\mu}. \quad (23)$$

The canonical partition function is given by

$$Q(N, V, T) = \underbrace{\frac{V^N}{N! \lambda_{\text{th}}^{3N}}}_{Q^{(ig)}(N, V, T)} \frac{1}{V^N} \int d\mathbf{r}^N \exp[-\beta U_N(\mathbf{r}^N)], \quad (24)$$

with a *potential* U_N in which the particles move. The prefactor contains the value of the canonical partition function of the ideal gas $Q^{(ig)}(N, V, T)$, with the *thermal wavelength*

$$\lambda_{\text{th}} = \sqrt{\frac{h^2}{2\pi m k_B T}}, \quad (25)$$

where m is a particle's mass and h is *Planck's constant*. Note that the potential U_N does not depend on particle velocity, that is the particles move in a *force field*.

For all that follows, we will only consider two-body interaction potentials of the form

$$U_N = \sum_{i=1}^N \sum_{j>i}^N \phi(ij), \quad (26)$$

$$\phi(ij) := \phi(\mathbf{r}_i, \mathbf{r}_j) = \phi(\mathbf{r}_j, \mathbf{r}_i).$$

For now, we do not need to know any further details on the interaction potential.

In order to adjust the upper equations to their use in graph theory, we make a change to the definition of the fugacity in equation (23) and give it the new form

$$z := \frac{\tilde{z}}{\lambda_{\text{th}}^3}. \quad (27)$$

In all that follows, we will refer to the quantity z when talking about fugacity, ignoring that \tilde{z} and z have different dimensions. In the framework of graph theory, our points carry property functions. We will therefore treat the fugacity as a constant function

$$z(i) \equiv z. \quad (28)$$

We can also use the functions

$$e(ij) := e^{-\beta\phi(ij)}, \quad (29)$$

$$f(ij) := e(ij) - 1.$$

These will be helpful if we insert (26) into the integral in (24). Without factors, the integral then takes the form

$$\begin{aligned} & \int d\mathbf{r}^N \exp(-\beta \sum_{j>i} \phi(ij)) \\ &= \int d\mathbf{r}^N \prod_{j>i} e(ij) = \int d\mathbf{r}^N \prod_{j>i} (f(ij) + 1) \\ &= \int d\mathbf{r}^N [1 + f(12) + \dots + f(12)f(23) + \dots \\ & \quad + f(12)f(23)f(34) + \dots]. \end{aligned} \quad (30)$$

So, we have our functions $f(ij)$ and $z(i)$. Using all reformulations of this chapter, we can start from equation (21) to reach a

simplified expression

$$\begin{aligned} \Xi &= \sum_{N \geq 0} \tilde{z}^N Q(N, V, T) \\ &= \sum_{N \geq 0} \frac{\tilde{z}^N V^N}{N! \lambda_{\text{th}}^{3N}} \frac{1}{V^N} \int d\mathbf{r}^N \exp(-\beta \sum_{j>i} \phi(ij)) \\ &= \sum_{N \geq 0} \frac{\tilde{z}^N}{N!} \int d\mathbf{r}^N \prod_{j>i} (f(ij) + 1) \\ &= \sum_{N \geq 0} \frac{1}{N!} \int d\mathbf{r}^N z(1) \dots z(N) [1 + f(12) + \dots]. \end{aligned} \quad (31)$$

This sum still contains many graphs of equal value. These can be removed by arguments outlined in Section 2.2. We drop the labels and include symmetry numbers for every graph in (31). It is therefore possible to express Ξ in terms of graphs according to

$$\begin{aligned} \Xi &= 1 + \frac{1}{2} \text{---} \bullet \text{---} \bullet + \frac{1}{2} \text{---} \bullet \text{---} \bullet \text{---} \bullet + \frac{1}{6} \text{---} \bullet \text{---} \bullet \text{---} \bullet \text{---} \bullet \\ &+ \frac{1}{2} \text{---} \bullet \text{---} \bullet \text{---} \bullet + \frac{1}{2} \text{---} \bullet \text{---} \bullet \text{---} \bullet + \frac{1}{6} \text{---} \bullet \text{---} \bullet \text{---} \bullet \text{---} \bullet \\ &+ \dots \end{aligned} \quad (32)$$

We can now go on to use the results of the previous section to further simplify the sum given here. To do so, we have a closer look at (32). It has the form of a sum over sums of N -point graphs. Each sum of N -point graphs contains the unconnected N -point graph, a graph with 1 connection, graphs with 2 connections, ..., a graph connecting all N points. All of them are weighted with the corresponding symmetry number. For $N \geq 2$, these graphs are not all connected. But they can all be constructed by combining connected subgraphs. On closer examination, it appears that this does exactly amount to the prerequisite of Lemma 5, where the set B is the set of all connected graphs with $N \geq 1$. Using (19) on (32) yields

$$\begin{aligned} \ln \Xi &= \bullet + \frac{1}{2} \text{---} \bullet \text{---} \bullet + \frac{1}{2} \text{---} \bullet \text{---} \bullet \text{---} \bullet + \frac{1}{6} \text{---} \bullet \text{---} \bullet \text{---} \bullet \text{---} \bullet + \dots \\ &= \{\text{Sum of all connected graphs with } N \geq 1 \\ & \quad \text{points. All points are } z \text{ field points and are} \\ & \quad \text{connected by } f\text{-bonds. (All graphs are weighed} \\ & \quad \text{by symmetry numbers } 1/s.)\} \end{aligned} \quad (33)$$

In the following, we will not mention that graphs in sums are weighted by symmetry numbers, since this will apply to all graphs we will look at. The property $\ln \Xi$ is of great importance to thermodynamic descriptions of our system, since it yields a thermal equation of state (EOS) for the pressure p by

$$\beta p V = \ln \Xi. \quad (34)$$

We do also know from thermodynamics that the *grand canonical potential* Ω is given by

$$\beta\Omega = -\ln \Xi. \quad (35)$$

The differential of Ω is

$$d\Omega = -SdT - pdV - \bar{N}d\mu. \quad (36)$$

Here, S is the *entropy* of the system. We wrote \bar{N} to indicate the expectation value of the number N of particles in the system, which must not be confused with N in the sums of (21) and similar expressions. We will need (36) later on.

Another important property is the number density ρ . It is the probability of finding a particle at a given position while ignoring the positions of all other particles. To get a good representation, we can write Ξ by

$$\Xi = \sum_{N=0}^{\infty} \int d\mathbf{r}^N P(N; \mathbf{r}^N) \quad (37)$$

with the probability distribution

$$P(N; \mathbf{r}^N) = \frac{1}{N!} z^N \exp(-\beta V(\mathbf{r}^N)). \quad (38)$$

ρ is then given by the grand canonical average of the delta peak, yielding

$$\begin{aligned} \rho(1) &= \frac{1}{\Xi} \sum_{N=1}^{\infty} N \int d(2) \cdots d(N) P(N; 1, \dots, N) \\ &= z \frac{1}{\Xi} \frac{\partial \Xi}{\partial z} = z \frac{\partial \ln \Xi}{\partial z}. \end{aligned} \quad (39)$$

The factor N in the integral equation takes into account identical, therefore indistinguishable particles. There are higher order number density functions, to which we will come back at the very end of this section. The expectation value of the number N of particles in the system depends on ρ by

$$\bar{N} = \int d(1) \rho(1). \quad (40)$$

We can also express ρ in terms of graphs. Remember that we assigned a constant function to z in (28). We can similarly replace z by $z(1)$ in (39) and replace $\ln \Xi$ by the sum of graphs in (32). If we apply Lemma 4 to each graph, we can express the number density $\rho(1)$ by graphs:

$$\begin{aligned} \rho(1) &= \text{graph 1} + \text{graph 2} + \text{graph 3} + \frac{1}{2} \text{graph 4} \\ &+ \frac{1}{2} \text{graph 5} + \frac{1}{2} \text{graph 6} + \dots \\ &= \{\text{Sum of all graphs obtained from (33) by} \\ &\quad \text{turning one of the field points into a labelled} \\ &\quad \text{point.}\} \end{aligned} \quad (41)$$

The equations (33) and (41) are the two first major results of the MCE. Their further simplification is one of the general goals of graph theory. Wertheim's approach¹ also starts developing its main advantages at this point. However, we will not directly proceed with Wertheim's reasoning but rather introduce a more "classical" way to reduce equation (41), the Mayer formalism. Wertheim's equations closely resemble those obtained by the Mayer formalism, and we can later explain them by mostly analogous arguments.

3.2 The Mayer Formalism and Topological Reduction

A defect of the theory we established so far is the great importance of the fugacity z . This quantity is more difficult to measure than other quantities that are also easier to comprehend by intuition – as for example the number density ρ . Our goal in this subsection is to determine equations for $\ln \Xi$ and the *Helmholtz free energy* A that do not depend on z but on ρ . This would be straightforward if we could replace all occurrences of z by ρ , so it seems a good idea to investigate the connection between those two. We already have equation (39), which can be put as

$$\frac{\rho(1)}{z(1)} = \frac{\partial \ln \Xi}{\partial z(1)}. \quad (42)$$

We will see that reducing the right hand side of this equation leads to the desired relation. Similar to (41), we get

$$\begin{aligned} \frac{\delta \ln \Xi}{\delta z(1)} &= 1 + \text{graph 1} + \text{graph 2} + \frac{1}{2} \text{graph 3} \\ &+ \frac{1}{2} \text{graph 4} + \frac{1}{2} \text{graph 5} + \dots \end{aligned} \quad (43)$$

This sum falls under the prerequisites of Lemma 6: all graphs contained can be uniquely expressed as products of graphs where the labelled point 1 is not an articulation point, and if one joins any graphs where 1 is not an articulation point together, the result is exactly one graph in the sum. We can use (20) and define a new function $c(1)$, which is sometimes referred to as the *correlation function* (also called $h(1)$ by some authors⁶). It is given by

$$c(1) := \ln \left[\frac{\rho(1)}{z(1)} \right], \quad (44)$$

and using the above simplifications, it is

$$\begin{aligned}
 c(1) = & \text{Diagram 1} + \text{Diagram 2} + \frac{1}{2} \text{Diagram 3} + \text{Diagram 4} \\
 & + \frac{1}{2} \text{Diagram 5} + \text{Diagram 6} + \dots \\
 = & \{\text{Sum of all graphs with one labelled point 1} \\
 & \text{and } N_z \geq 1 \text{ } z \text{ field points connected by } f\text{-bonds} \\
 & \text{such that the labelled point is no articulation} \\
 & \text{point.}\}
 \end{aligned} \tag{45}$$

Equation (45) contains all of the graphs with $N \leq 3$ points in $c(1)$ and some of those with $N = 4$ points. We chose them to illustrate the next step in our reduction. We already used the fact that in some graphs of ρ the labelled point is an articulation point. However, there are also numerous graphs where at least one field point is an articulation point. These are still contained in $c(1)$, in (45) they are in the second, forth, fifth and sixth graph. These graphs can all be constructed from an at least doubly connected subgraph containing the labelled point and a graph from ρ in (41) attached to each field point such that the field point is overwritten by the labelled point in ρ . The attached graph may be the one point graph. One can group all the graphs in $c(1)$ into equivalence classes by defining the *maximal connected subgraph* Γ_m for any graph Γ : it is the subgraph with the most field points in the set of subgraphs free of articulation field points that contain the labelled point. Since the labelled point is not an articulation point in the graphs of $c(1)$, Γ_m is unique for any Γ in $c(1)$. If we now look at all the graphs contained in one of the equivalence classes defined by Γ_m , we see that it contains exactly all graphs obtainable by attaching graphs from ρ to the field points in the manner described above. We can therefore use the idea of changing the property of a point we described in Section 2.4 and change the property of the field points from $z(1)$ to $\rho(1)$. Then, we can express $c(1)$ by

$$\begin{aligned}
 c(1) = & \text{Diagram 1} + \frac{1}{2} \text{Diagram 2} + \frac{1}{2} \text{Diagram 3} + \frac{1}{2} \text{Diagram 4} \\
 & + \frac{1}{2} \text{Diagram 5} + \frac{1}{6} \text{Diagram 6} + \dots \\
 = & \{\text{Sum of all graphs with one labelled point 1,} \\
 & N_\rho \geq 1 \text{ } \rho \text{ field points connected by } f\text{-bonds} \\
 & \text{such that none of the points is an articulation} \\
 & \text{point.}\}
 \end{aligned} \tag{46}$$

This equation contains all of the graphs with $N \leq 4$ points in $c(1)$. We use blue points to mark ρ field points. Using the definition of $c(1)$ and the fact that we can express $c(1)$ purely ρ dependent, we can now express $z(1)$ with respect to ρ by inserting (46) into (44), which yields

$$\ln z(1) = \ln \rho(1) - c(1). \tag{47}$$

With these equations, we are able to give a ρ dependent version of $\ln \Xi$. We start with the variation given by (12), which is for $\ln \Xi$

$$\begin{aligned}
 \delta \ln \Xi &= \int d(1) \frac{\delta \ln \Xi}{\delta z(1)} \delta z(1) \\
 &= \int d(1) \rho(1) \frac{\delta z(1)}{z(1)} \\
 &= \int d(1) \rho(1) \delta [\ln z(1)] \\
 &= \int d(1) \rho(1) \delta [\ln \rho(1) - c(1)] \\
 &= \int d(1) [\delta \rho(1) - \delta [\rho(1) c(1)] + c(1) \delta \rho(1)].
 \end{aligned} \tag{48}$$

If there were a function $c^{(0)}$ with the property

$$c(1) = \frac{\delta c^{(0)}}{\delta \rho(1)}, \tag{49}$$

then (48) could be written

$$\delta \ln \Xi = \delta \left[\int d(1) \rho(1) - \int d(1) \rho(1) c(1) + c^{(0)} \right]. \tag{50}$$

This implies

$$\ln \Xi = \int d(1) \rho(1) - \int d(1) \rho(1) c(1) + c^{(0)}, \tag{51}$$

where we ignored the possible addition of a constant, as it is unimportant in a thermodynamic potential. The only chore left is to find the missing graph sum $c^{(0)}$. It is however easy to guess from (46). It must be

$$\begin{aligned}
 c^{(0)} = & \frac{1}{2} \text{Diagram 1} + \frac{1}{6} \text{Diagram 2} + \frac{1}{4} \text{Diagram 3} \\
 & + \frac{1}{8} \text{Diagram 4} + \frac{1}{24} \text{Diagram 5} + \dots \\
 = & \{\text{Sum of all connected graphs with } N \geq 2 \\
 & \rho \text{ field points connected by } f\text{-bonds such} \\
 & \text{that none of the points is an articulation point.}\}
 \end{aligned} \tag{52}$$

Since the symmetry numbers are changed adequately by the process of derivation, $c^{(0)}$ given by (52) satisfies (49).

From an expression for $\ln \Xi$, which gives us an equation of state by the relation (34), we can also derive an expression for the free energy A . It is related to the grand canonical potential Ω by

$$\beta A = \beta \Omega + \beta \mu \bar{N} = -\ln \Xi + \ln(z \lambda_{\text{th}}^3) \bar{N}. \quad (53)$$

We can do some rearrangements by remembering that $z(1) \equiv z$ is constant and using equation (40), (47) and (50), which leads to

$$\begin{aligned} \beta A &= - \int d(1) \rho(1) + \int d(1) \rho(1) c(1) - c^{(0)} \\ &\quad + \int d(1) \rho(1) \ln(z(1) \lambda_{\text{th}}^3) \\ &= \int d(1) \rho(1) [\ln(\rho(1) \lambda_{\text{th}}^3) - 1] - c^{(0)}. \end{aligned} \quad (54)$$

We have therefore reached the goal of this section and derived equations for $\ln \Xi$ and βA that do not depend on the possibly disadvantageous property z but on the number density ρ . Before proceeding, we want to introduce a different class of functions that are often of interest to thermodynamic models. One may consider them a generalization of $c(1)$ to multiple labelled points and they, too, are called correlation functions.

3.3 s -Particle Distribution and Correlation Functions

The function $\rho(1)$ describes the particle density at location 1 (possibly including orientation). Obviously, there may be functions $\rho(1 \cdots s)$ that describe the density of s particles – or, in that case more accurately, their distribution – at the locations $1, \dots, s$. They are described by a more general form of (39),

$$\rho(1 \cdots s) = \sum_{N=s}^{\infty} \int d(s) \cdots d(N) \frac{N!}{(N-s)!} P(N; \mathbf{r}^N), \quad (55)$$

which again is the result for calculating the canonical average of delta peaks. We can simplify (55) to

$$\rho(1 \cdots s) = \frac{1}{\Xi} \prod_{i=1}^s z(i) \frac{\delta^s \Xi}{\prod_{i=1}^s \delta z(i)}. \quad (56)$$

This time, we can not jump straight to the logarithm for which we have a representation by connected graphs. However, we can define the *Ursell functions*

$$u(1 \cdots s) := \prod_{i=1}^s z(i) \frac{\delta^s \ln \Xi}{\prod_{i=1}^s \delta z(i)} \quad (57)$$

that are in close relation to $\rho(1 \cdots s)$. To investigate this relation, we define for $s \geq 2$ the *correlation functions*

$$h(1 \cdots s) = \frac{u(1 \cdots s)}{\rho(1) \cdots \rho(s)} \quad (58)$$

and

$$g(1 \cdots s) = \frac{\rho(1 \cdots s)}{\rho(1) \cdots \rho(s)}. \quad (59)$$

In the special case of $s = 1$, we already met the function $h(1)$, which is the same as $c(1)$ in (44). We can now consider the relationship between the $u(1 \cdots s)$ functions and the $\rho(1 \cdots s)$ functions, starting with the obvious case of

$$u(1) = \rho(1). \quad (60)$$

For $s = 2$, we find

$$\begin{aligned} u(12) &= z(1)z(2) \frac{\delta}{z(2)} \frac{1}{\Xi} \frac{\partial \Xi}{\partial z(1)} \\ &= z(1)z(2) \left[\frac{1}{\Xi} \frac{\delta^2 \Xi}{\delta z(1) \delta z(2)} - \frac{1}{\Xi^2} \frac{\partial \Xi}{\partial z(1)} \frac{\partial \Xi}{\partial z(2)} \right] \\ &= \rho(12) - \rho(1)\rho(2). \end{aligned} \quad (61)$$

It becomes evident that by an analogous procedure any $\rho(1 \cdots s)$ can be described by $u(1 \cdots r)$ with $r \leq s$. Since the $u(1 \cdots s)$ are all sums of graphs obtained from $\ln \Xi$ by turning s field points into labelled points, $\rho(1 \cdots s)$ will have a similar form. We will explain this with the example of the two-particle correlation functions $g(12)$ and $h(12)$. Let us look closer at $u(12)$. It contains

$$u(12) = \text{graph 1} + \text{graph 2} + \text{graph 3} + \text{graph 4} + \text{graph 5} + \text{graph 6} + \dots \quad (62)$$

We can see that any graph in $u(12)$ can be constructed from a subgraph in which none of the labelled points is an articulation point (the field points may still be articulation points) and a graph from $\rho(1)$ appended (i.e. multiplied) to the point 1 and a graph from $\rho(2)$ to the point 2. Note that the subgraph is thus unique. The sum over all these subgraphs is the correlation

function $h(12)$, given by

$$\begin{aligned}
 h(12) &= \text{---}_1\text{---}_2 + \text{---}_1\text{---}_2 + \text{---}_1\text{---}_2 \\
 &+ \text{---}_1\text{---}_2 + \text{---}_1\text{---}_2 + \dots \\
 &= \{\text{Sum of all connected graphs consisting of} \\
 &\quad \text{labelled points 1 and 2, } z\text{-field points and} \\
 &\quad \text{f-bonds. No labelled point is an articulation} \\
 &\quad \text{point.}\}
 \end{aligned} \tag{63}$$

where dashed points are again points without property. Alongside (61), this leads to

$$\rho(12) = \rho(1)h(12)\rho(2) + \rho(1)\rho(2) = \rho(1)g(12)\rho(2). \tag{64}$$

Here, g is simply

$$g(12) = h(12) + 1. \tag{65}$$

1 is also the value of the graph containing two unbonded labelled points without property. If we add it to (63), we see that for any graph in $g(12)$ where there is an f -bond between the labelled points, there is also a graph in $g(12)$ where there is no bond between the two, and vice versa. The sum of an f -bond and no bond is an e -bond. We can therefore write

$$\begin{aligned}
 g(12) &= \text{---}_1\text{---}_2 + \text{---}_1\text{---}_2 + \text{---}_1\text{---}_2 + \dots \\
 &= \{\text{Sum of all connected graphs consisting of} \\
 &\quad \text{labelled points 1 and 2 connected by an} \\
 &\quad \text{e-bond and a network of } z \text{ field points and} \\
 &\quad \text{f-bonds. No labelled point is an articulation} \\
 &\quad \text{point.}\}
 \end{aligned} \tag{66}$$

Dashed bonds are e -bonds. It is possible to further reduce $g(12)$ by the same arguments we used to reduce $c(1)$. Since the field points may still be field articulation points, the sum in $g(12)$ contains maximal articulation point free subgraphs and all graphs resulting from appending some graph of $\rho(1)$ to each field point. We can therefore replace z field points by ρ field points in (66), arriving at

$$\begin{aligned}
 g(12) &= \text{---}_1\text{---}_2 + \text{---}_1\text{---}_2 + \text{---}_1\text{---}_2 + \dots \\
 &= \{\text{Sum of all at least doubly connected graphs} \\
 &\quad \text{consisting of labelled points 1 and 2 connected} \\
 &\quad \text{by an e-bond and a network of } \rho \text{ field} \\
 &\quad \text{points and f-bonds.}\}
 \end{aligned} \tag{67}$$

This concludes the remarks on correlation functions for now.

4 Wertheim's Multiple Density Formalism

The approach presented in the previous section is rather general, since we have not committed to any forms of the interaction potential $\phi(12)$. Of course, when choosing a specific class of $\phi(12)$, it may be advisable to also undertake variations to the process of graph reduction. One of these variations was presented by M. S. Wertheim in 1983¹, and this section will describe its main ideas.

Most of what follows is part of Wertheim's first paper on the topic¹, which we will simply call I or Wertheim I. We also consider his generalization presented in III³. The two other subsequent publications II² and IV⁴ will only be briefly visited. Wertheim submitted an additional paper on thermodynamic perturbation theory⁵, which we will especially consider in the next section. We will refer to it by (Wertheim) TPT.

4.1 Revisiting the Fundamentals

As we mentioned at the end of the last section, the Mayer Cluster Expansion in its generality may not be the most adequate way to treat special forms of the interaction potential $\phi(12)$. The form Wertheim chose to analyze is

$$\phi(12) = \phi_R(12) + \sum_{A,B \in \Gamma} \phi_{AB}(|\mathbf{r}_2 + \mathbf{d}_B(\Omega_2) - \mathbf{r}_1 - \mathbf{d}_A(\Omega_1)|). \tag{68}$$

A and B are labels for interaction sites on each particle and belong to a set of sites Γ . The position of A and B is described by $\mathbf{d}(\Omega_i)$, dependent on the orientation Ω_i of each particle. The potentials ϕ_{AB} are considered to be purely attractive, that is

$$\phi_{AB} \leq 0 \quad \forall A, B. \tag{69}$$

The rest of the interaction is purely repulsive, preferably strong enough to make a lot of graphs vanish later in the expansion. A classical example is the *hard sphere potential* ϕ_R^{HS} , which strictly forbids core overlap by

$$\phi_R^{HS}(12) = \begin{cases} \infty, & r_{12} \leq D, \\ 0, & r_{12} > D, \end{cases} \tag{70}$$

where D is the hard sphere diameter.

Wertheim I focusses on the case that there is only one interaction site per particle, turning (68) into

$$\phi(12) = \phi_R(12) + \phi_A(|\mathbf{r}_2 + \mathbf{d}_A(\Omega_2) - \mathbf{r}_1 - \mathbf{d}_A(\Omega_1)|). \tag{71}$$

The binding site A is situated at the edge of the repulsive core. In the hard sphere example, this may be of the form

$$\phi_A(x) = \begin{cases} < 0 & x < a, \\ = 0 & x > a. \end{cases} \quad (72)$$

The position of the site defined by $\mathbf{d}_A(\Omega_i)$ has to be chosen to ensure that the site is within the particle but close enough to the edge to avoid that the attractive region is cancelled out by the infinite core repulsion. A case of special interest would be short-ranged interaction leading to glue spots such that a site can only interact attractively with one other site. In that case, a would have to be much smaller than D .

Potentials of the form (68) can obviously be clearly separated into their attractive and their repulsive contributions. So it seems adequate to maintain this separation when considering $e(12)$ and $f(12)$. We can define

$$\left. \begin{aligned} e_i(12) &:= \exp(-\beta\phi_i(12)) \\ f_i(12) &= e_i(12) - 1 \end{aligned} \right\} i \in \{R, A\}, \quad (73)$$

leading to

$$e(12) = e_R(12)e_A(12), \quad (74)$$

which means for $f(12)$ as in (29) that

$$\begin{aligned} f(12) &= e_R(12)(e_A(12) - 1 + 1) - 1 \\ &= [e_R(12) - 1] + e_R(12)[e_A(12) - 1] \\ &=: f_R(12) + F(12). \end{aligned} \quad (75)$$

We can use this new definition to change (33) and (41). The resulting graphs should contain either f_R -bonds and F -bonds. However, it is more convenient to define a subset of graphs with the help of which all graphs in $\ln \Xi$ can be constructed. Wertheim introduces the s -mer graphs, which are the graphs of s points in which all points are connected by F -bonds. Note that there may also be f -bonds present. Because of this, the set of pure s -mer graphs can be further simplified, which we can illustrate by the example of all trimer graphs, which would be in a sum

$$\begin{aligned} & \frac{1}{2} \text{ (graph with 3 points, 2 zigzag bonds, 1 straight bond)} + \frac{1}{2} \text{ (graph with 3 points, 2 zigzag bonds, 1 straight bond)} + \frac{1}{6} \text{ (graph with 3 points, 3 zigzag bonds)} \\ &= \frac{1}{2} \text{ (graph with 3 points, 2 zigzag bonds, 1 straight bond)} + \frac{1}{6} \text{ (graph with 3 points, 3 zigzag bonds)} \end{aligned} \quad (76)$$

The zigzag connections are here (and in the following) F -bonds, the straight continuous lines are f_R -bonds and the dashed lines are e_R -bonds. It is clear that when replacing all f -bonds by $F + f_R$, a replacement like in (76) will be possible for any graph already connected by F -bonds, since the sum includes F -connected graph and no other bond and the F -connected graph plus an f_R -bond, plus two f_R -bonds etc. The

replacement procedure will also not change the symmetry numbers. We do not give a clean mathematical explanation here, but it is at least evident that one can imagine the replacement to take place "symmetrically" on all bonds. Therefore, we state

Definition 7

A (pure) s -mer graph is an s -point graph of z field points where all points are connected by F -bonds. All pairs of points not directly connected by an F -bond are connected by an e_R -bond.

This leads us straight to an expression for

$$\ln \Xi = \{ \text{Sum of all connected graphs consisting of } s\text{-mer graphs, } s \geq 1, \text{ and } f_R\text{-bonds between pairs of points in distinct } s\text{-mer graphs.} \} \quad (77)$$

$\rho(1)$ can be derived from $\ln \Xi$ in the same way as before, it is still the sum over all graphs obtained from $\ln \Xi$ by turning a field point into a point labelled 1.

The e_R -bonds within s -mers are considered irreducible, that is it is not allowed to exchange an e_R -bond by the sum of an f_R -bond and no bond. This has the important consequence that a pure s -mer graph does not contain articulation points of any sorts. This means that neither can an s -mer graph be broken down into two or more independent subgraphs at a labelled point nor can an s - and an s' -mer graph with $s, s' \geq 2$ be joined together into an $(s + s' - 1)$ -mer graph at a labelled point since there would be at least one e_R -bond missing. You could say that by Definition 7, all s -mer graphs are closed in a sense that it is impossible to obtain a higher s -mer graph by appending a new graph to a single point.

This last remark makes the following approach seem interesting: it could be useful to separate the graphs in $\rho(1)$ into those where an s -mer with $s \geq 2$ can be attached to the labelled point and those where this is not possible. This idea and its consequences are the next section's focus.

4.2 The Two Density Approach

We decompose $\rho(1)$ by

$$\rho(1) = \rho_0(1) + \rho_1(1). \quad (78)$$

Here, $\rho_0(1)$ contains all graphs where the labelled point is in a monomer, that is it has no adjacent F -bonds. $\rho_1(1)$ contains all the other graphs, i.e. the graphs where the labelled point is in an s -mer with $s \geq 2$. We can call $\rho_0(1)$ the *monomer density*. Like in Section 3.2, $\rho_0(1)/z$ is the exponential of the subgraphs of ρ_0/z that have no articulation labelled point. In analogy, we define

$$c_0(1) := \ln \left[\frac{\rho_0(1)}{z(1)} \right]. \quad (79)$$

Since an arbitrary number of f_R -bonds can be attached to the labelled point (in fact, any point) in $\rho_1(1)$, it must contain $\rho_0(1)$ as a factor. It is therefore convenient to define

$$c_1(1) := \frac{\rho_1(1)}{\rho_0(1)}, \quad (80)$$

the sum of all graphs with a property-free labelled point 1 in an s -mer ($s \geq 2$) in which 1 is not an articulation point.

In the Mayer formalism, the next step would be to change the property of the field points from z to ρ to get an articulation point free representation of $c(1)$ and a ρ dependent expression for z . The Wertheim approach continues analogously, but there is a slight variation necessary.

We did already elaborate in the remarks on Definition 7 that we can not add F -bonds to a field point that is already in an s -mer with $s \geq 2$. Therefore, those points within doubly connected graphs in $c_0(1)$ and $c_1(1)$ that do belong to such an s -mer can only carry graphs from $\rho_0(1)$. For monomeric points within the doubly connected graphs it is still possible to attach any graph from $\rho(1)$ to the point. Therefore, we change the property from $z(1)$ to either $\rho(1)$ or $\rho_0(1)$. We consider only the at least doubly connected graphs in $c_0(1)$ and $c_1(1)$, where points in monomers carry the property $\rho(1)$ and points in s -mers for $s \geq 2$ carry the property $\rho_0(1)$. For example,

$$c_0(1) = \underbrace{\text{graph}}_1 + \frac{1}{2} \underbrace{\text{graph}}_1 + \frac{1}{2} \underbrace{\text{graph}}_1 + \frac{1}{2} \underbrace{\text{graph}}_1 + \dots \quad (81)$$

Here, the blue points are still ρ field points and the green points are ρ_0 field points.

So what happens to $\ln \Xi$? We can use (79) and then (80) to find for the variation of $\ln \Xi$ that

$$\begin{aligned} \delta \ln \Xi &= \int d(1) \rho(1) \delta [\ln z(1)] \\ &= \int d(1) \rho(1) \delta [\ln \rho_0(1) - c_0(1)] \\ &= \int d(1) \frac{\rho(1)}{\rho_0(1)} \delta \rho_0(1) - \int d(1) \delta [\rho(1) c_0(1)] \\ &\quad + \int d(1) c_0(1) \delta \rho(1) \\ &= \int d(1) \delta \rho_0(1) - \int d(1) \delta [\rho(1) c_0(1)] \\ &\quad + \int d(1) [c_0(1) \delta \rho(1) + c_1(1) \delta \rho_0(1)]. \end{aligned} \quad (82)$$

Again, the similarity to (48) is strong. This time, we need a

function $c^{(0)}$ that satisfies

$$c_0(1) = \frac{\delta c^{(0)}}{\delta \rho(1)}, \quad c_1(1) = \frac{\delta c^{(0)}}{\delta \rho_0(1)}. \quad (83)$$

Then, we have

$$\delta c^{(0)} = \int d(1) [c_0(1) \delta \rho(1) + c_1(1) \delta \rho_0(1)], \quad (84)$$

and therefore, again ignoring addition of a constant,

$$\ln \Xi = \int d(1) \rho_0(1) - \int d(1) \rho(1) c_0(1) + c^{(0)}. \quad (85)$$

, or the equivalent

$$\ln \Xi = \int d(1) [\rho(1) - c_1(1) \rho_0(1) - c_0(1) \rho(1)] + c^{(0)}. \quad (86)$$

Finding $c^{(0)}$ is fairly obvious from $c_0(1)$ and $c_1(1)$. It should be the sum of the graphs in $c_0(1)$ and $c_1(1)$ where the labelled point is turned into a ρ field point or a ρ_0 field point, respectively. All graphs resulting from this manipulation are only counted once (despite the fact that the derivative of the graph may appear in both $c_0(1)$ and $c_1(1)$) and weighted with the correct symmetry number. The straightforward definition can be found in Wertheim I.

After finding $\ln \Xi$, we can again find βA by (54),

$$\begin{aligned} \beta A &= - \int d(1) \rho_0(1) + \int d(1) \rho(1) c_0(1) - c^{(0)} \\ &\quad + \int d(1) \rho(1) \ln (z(1) \lambda_{\text{th}}^3) \\ &= \int d(1) [\rho(1) \ln (\rho(1) \lambda_{\text{th}}^3) - \rho_0(1)] - c^{(0)}. \end{aligned} \quad (87)$$

The two density formalism stays therefore very close to the Mayer formalism.

4.3 Multiple Densities

The two density formalism is designed for cases where the attractive interaction between particles is only governed by a single site per particle. While the creation of s -meric compounds with arbitrary values for s is possible with potentials of that kind, the structure of the compounds has very little degrees of freedom. For example, a system containing chains of particles can not be modelled by this approach, the chains would require at least two sites per particle. To meet this new challenge, it is possible to generalize the two density formalism to a multiple density formalism which can deal with an arbitrary number of sites per particle. This section shall briefly outline the main ideas of this generalization. We revisit the four steps we covered for the two density case, namely to define a potential, to define the topology including s -mer graphs, to introduce

new densities ρ_α and to find reduced funtions c_α , and finally to express the EOS and βA in terms of the definitions we made.

New potential. We now consider potentials of the form (68) without the restriction of (71). The set of all sites on one particle shall be denoted by Γ . We still use the functions $e_R(12)$ and $f_R(12)$ for the repulsive interaction and for the attractive interaction

$$\left. \begin{aligned} e_{AB}(12) &:= \exp(-\beta\phi_{AB}(12)) \\ f_{AB}(12) &= e_{AB}(12) - 1 \end{aligned} \right\} A, B \in \Gamma. \quad (88)$$

This means for the full $e(12)$ and $f(12)$ that

$$\begin{aligned} e(12) &= e_R(12) \prod_{A,B} e_{AB}(12), \\ f(12) &= f_R(12) + e_R(12) \left\{ \prod_{A,B} [f_{AB}(12) + 1] - 1 \right\}. \end{aligned} \quad (89)$$

If we set $\phi_{AA} = 0$ for all $A \in \Gamma$, then $f(12)$ is of the form

$$\begin{aligned} f(12) &= f_R(12) + e_R(12) \{ f_{AB}(12) + f_{AC}(12) + \dots \\ &\quad + f_{AB}(12)f_{AC}(12) + f_{AC}(12)f_{BC}(12) + \dots \}. \end{aligned} \quad (90)$$

Introducing bonds of type $F(12) = e_R(12)f_{AB}(12)$ only made sense if there were a maximum of one bond of this type between two particles, in the situation with more than one site per particle and therefore the possibility of multiple attractive bonds between particles, this would lead to higher powers of $e_R(12)$, which would not be true to (89). Indeed, it is now important to clearly distinguish bonds between (attractive) interaction sites from bonds between (repulsive) particles. There may be more than one bond $f_{AB}(12)$ between two particles while there can still only be one bond of type $e_R(12)$ or $f_R(12)$. As in Wertheim I, the attractive f_{AB} -bonds between sites are still paired with repulsive e_R -bonds between particles, so it remains possible to use the advantage of identifying s -mers by their interconnection with e_R -bonds. However, the question of the additional e_R -bonds resulting from combining the graph with an f_R -bond between a pair of points and the graph where there is no bond present should be discussed again.

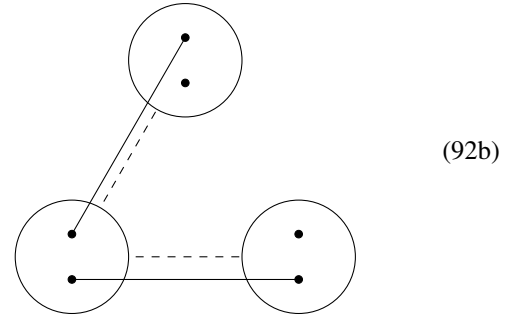
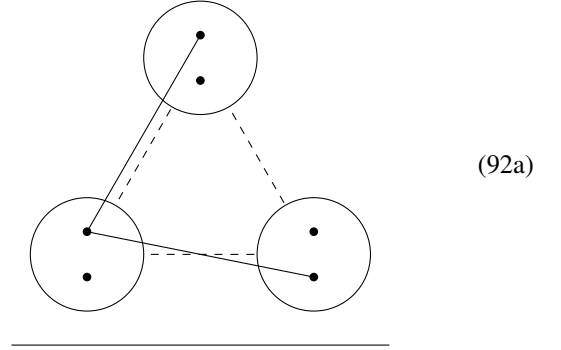
Toplogy and s -mers. We need to introduce some additional terminology. This is best done with some explanatory graphs. We can draw a point with 2 sites by

$$\begin{array}{c} \text{---} \circ \text{---} \\ \bullet \\ \bullet \end{array} \quad \text{or} \quad \begin{array}{c} \text{---} \bigcirc \text{---} \\ \bullet \\ \bullet \end{array} \quad (91)$$

Adopting Wertheim's definition, we call the big circle a *hyperpoint* and the small black circles are the sites. We will additionally introduce the convention that a hyperpoint with a thick

circle line is a field hyperpoint (right point in (91)) while a thin circle means a labelled hyperpoint (left point in (91)).

Now, consider how the following graphs differ in connection:



The e_R -bonds already anticipate Wertheim's generalization, which we will now explain.

One would consider both graphs to be connected. The connection in (92a) is clearer, the sites are connected by direct attraction bonds. The connection in (92b) comes from attraction bonds and the fact that sites are connected because they belong to the same hyperpoint. Wertheim calls the sites that are connected by a path of attraction bonds and sites *bond-connected*, while he calls sites belonging to the same hyperpoint *constraint-connected*. The hyperpoints of bond-connected sites are all additionally interconnected by e_R -bonds, while in the case of constraint-connection no further e_R -bonds are added.

An s -mer graph would therefore be a graph in which all hyperpoints are connected by e_R -bonds. There may be pairs of points without an e_R -bond. Those may carry either no bond or an f_R -bond, which may not be reduced to an e_R -bond. Therefore, the e_R -connection may not be as interwoven as in the two density case. It is possible, as for example at the left hyperpoint in (92b), that a hyperpoint is an articulation point. We call such a point a *constraint-articulation point* (c-AP). A hyperpoint is a c-AP if deleting the constraint connection of the sites within the hyperpoint and all incident f_R and e_R -bonds leads to two or more separate graphs.

In terms of f_R -bonds, Wertheim introduces the *bare* s -mer graphs as those s -mer graphs where no f_R -bonds are present

and the *hindered* s -mer graphs as those where at least one f_R -bond is present. The two sets are disjoint and the set of all pure s -mer graphs is the union of both.

Finally, a quick remark on the purpose of using the third e_R -bond in (92a) and not in (92b). Especially when progressing to TPT, Wertheim uses steric incompatibilities between hyperpoints – that is hard core overlap – to further treat the equations this section is concerned with. In (92a) one site is connected to two sites in different hyperpoints. If one identifies sites with valence electrons, this type of bonding becomes physically dubious. To ensure the cancellation of graphs of this kind, Wertheim demands the e_R -bond. The graph (92a) will always vanish: either at least one of the f_{AB} -bonds is zero because the pertaining sites are too far apart, or if both the attractive bonds are non-vanishing, the hard core overlap between the hyperpoints forbids this structure, which mathematically means that the e_R -bond between the upper and the right hyperpoint in (92a) becomes zero. Including the bond and considering it irreducible will ensure that it stays intact after the process of topological reduction and that therefore no information about steric incompatibility is lost. One can later use it to further simplify the results.

We did now introduce a variety of new definitions and conventions into the former framework of the two density formalism. The new definitions do not change the main form for $\ln \Xi$ as stated by (77), where only the transition from points to hyperpoints has to be kept in mind. $\rho(1)$ is again obtained by turning field hyperpoints in $\ln \Xi$ into labelled hyperpoints.

Defining densities. While we had ρ_0 for the graphs with no site-site interaction at the labelled point and ρ_1 for those with at least one, we can now generally introduce ρ_α for the graphs where the labelled point has (possibly multiple) attraction bonds at all sites in the set α . If we call the set of all sites per particle Γ , this leads to

$$\rho(1) = \sum_{\alpha \subseteq \Gamma} \rho_\alpha(1). \quad (93)$$

The set $\rho_0(1)$ stays basically the same, so, in the spirit of (79), we can define

$$c_0(1) := \ln \left[\frac{\rho_0(1)}{z(1)} \right]. \quad (94)$$

$c_0(1)$ contains all graphs of $\rho_0(1)/z$ where 1 is not a c-AP. Again, any $\rho_\alpha(1)$ must contain $\rho_0(1)$ as a factor. We can define for $\alpha \neq \emptyset$

$$c_\alpha(1) := \{\text{Sum of all graphs of } \rho_\alpha(1)/\rho_0(1) \text{ such that 1 is not a c-AP.}\} \quad (95)$$

$\rho_\alpha(1)/\rho_0(1)$ does not only contain $c_\alpha(1)$, but also products of $c_\gamma(1)$, where γ is part of partitions of α . An example: When turning the left point in (92b) into a labelled point, the resulting graph belongs to ρ_{AB} . However, it does not belong to $\rho_0(1)c_{AB}(1)$ but to $\rho_0(1)c_A(1)c_B(1)$ since 1 is a c-AP.

This leads to

$$\rho_\alpha(1) = \rho_0(1) \sum_{\{\gamma\} \in P(\alpha)} \prod_\gamma c_\gamma(1). \quad (96)$$

$P(\alpha)$ is the partition of α into nonempty subsets. Especially, $\{\alpha\} \in P(\alpha)$. (96) is simply the mathematical illustration that any graph belonging to $\rho_\alpha(1)$ is constructed by repulsive bonds between the hyperpoint 1 and other s -mers ($\rho_0(1)$) and combinations of irreducible s -mers such that exactly all sites in α are bonded at the hyperpoint 1 (product over $c_\gamma(1)$).

As before, we can reduce all the $c_\gamma(1)$ by considering which graphs can still be attached to the hyperpoints of all irreducible subgraphs and their sites without violating the s -mer construction conditions dictated by e_R -bonds. For two densities, either $\rho_0(1)$ or the full $\rho(1)$ could be attached. The density $\rho_1(1)$ is only indirectly included since $\rho(1) = \rho_0(1) + \rho_1(1)$. And indeed, whenever a graph of $\rho_1(1)$ can be appended to 1, then it must also be possible to append a graph of $\rho_0(1)$ to it. This can be expanded to a case of multiple sites: if it is possible to add a graph from some $\rho_\alpha(1)$ to a point 1 – that is, a graph in which all sites of α are bonded in at least one s -mer – then it must be possible to append a graph from any $\rho_\gamma(1)$ with $\gamma \subseteq \alpha$ to 1. We define

$$\sigma_\alpha(1) := \sum_{\gamma \subseteq \alpha} \rho_\gamma(1), \quad (97)$$

especially

$$\sigma_0(1) = \rho_0(1), \quad \sigma_\Gamma(1) = \rho(1). \quad (98)$$

By $\sigma_\alpha(1)$ we can reduce the graphs $c_\alpha(1)$ as we did in the previous section. This yields

$$c^{(0)} = \{\text{Sum of all at least doubly connected graphs constructed by } s\text{-mer graphs of field hyperpoints interconnected by } f_R\text{-bonds. If } \alpha \text{ is the set of all attraction bonded sites of a hyperpoint, then the hyperpoint carries the property } \sigma_{\Gamma-\alpha}.\} \quad (99)$$

The properties $c_\alpha(1)$ are again obtained by differentiation via

$$c_\alpha(1) = \frac{\delta c^{(0)}}{\delta \sigma_{\Gamma-\alpha}}. \quad (100)$$

Finding the EOS. The remainder is just a longer calculation. We break it into smaller pieces, starting from the very begin-

ning with

$$\begin{aligned}
\delta \ln \Xi &= \int d(1) \rho(1) \delta [\ln z(1)] \\
&\stackrel{(94)}{=} \int d(1) \frac{\rho(1)}{\rho_0(1)} \delta \rho_0(1) - \int d(1) \rho(1) \delta c_0(1) \\
&= \underbrace{\int d(1) \frac{\rho(1)}{\rho_0(1)} \delta \rho_0(1) - \int d(1) \delta \rho(1)}_{(A)} \\
&\quad - \underbrace{\int d(1) \rho(1) \delta c_0(1)}_{(B)} + \underbrace{\int d(1) \delta \rho(1)}_{(C)}.
\end{aligned} \tag{101a}$$

We inserted the term (C) to get closer to the variation of (86). We can leave (C) and (B) as they are and focus on the term (A). Also, to keep equations slim, we drop the arguments in the integrand.

$$\begin{aligned}
(A) &= \int \left\{ \frac{\rho}{\rho_0} \delta \rho_0 - \delta \rho \right\} \\
&\stackrel{(93)}{=} \int \left\{ \left[1 + \sum_{\gamma \subseteq^* \Gamma} \frac{\rho_\gamma}{\rho_0} \right] \delta \rho_0 - \delta \rho \right\} \\
&= \int \left\{ \sum_{\gamma \subseteq^* \Gamma} \frac{\rho_\gamma}{\rho_0} - \delta [\rho - \rho_0] \right\} \\
&= \int \sum_{\gamma \subseteq^* \Gamma} \left\{ \frac{\rho_\gamma}{\rho_0} - \delta \rho_\gamma \right\}.
\end{aligned} \tag{101b}$$

The notation $\gamma \subseteq^* \Gamma$ shall indicate that only the nonempty subsets of Γ are summed over. For the quotient ρ_γ/ρ_0 we can use (96) to find

$$\begin{aligned}
(A) &= \int \sum_{\gamma \subseteq^* \Gamma} \sum_{M \in P(\gamma)} \left\{ \left[\prod_{\alpha \in M} c_\alpha \right] \delta \rho_0 - \delta \left[\rho_0 \prod_{\alpha \in M} c_\alpha \right] \right\} \\
&= \int \sum_{\gamma \subseteq^* \Gamma} \sum_{M \in P(\gamma)} \left\{ \left[\prod_{\alpha \in M} c_\alpha \right] \delta \rho_0 - \left[\prod_{\alpha \in M} c_\alpha \right] \delta \rho_0 \right. \\
&\quad \left. - \rho_0 \delta \left[\prod_{\alpha \in M} c_\alpha \right] \right\} \\
&= - \int \sum_{\gamma \subseteq^* \Gamma} \sum_{M \in P(\gamma)} \sum_{\alpha \in M} \rho_0 \left[\prod_{\alpha' \in M \setminus \{\alpha\}} c_{\alpha'} \right] \delta c_\alpha.
\end{aligned} \tag{101c}$$

This long equation will reduce to a rather simple one by a rearrangement of sums. We want to get the sum over α to the very left. We first have to exchange it with the sum over M . Since M counts all the partitions of γ into subsets and α runs over all subsets contained in M , we can express the sum over all subsets contained in all partitions of γ as the sum over all

nonempty subsets $\alpha \subseteq^* \gamma$ and the partitions M of γ containing the subset. That is,

$$(A) = - \int \sum_{\gamma \subseteq^* \Gamma} \sum_{\alpha \subseteq^* \gamma} \sum_{\substack{M \in P(\gamma) \\ M \ni \alpha}} \rho_0 \left[\prod_{\alpha' \in M \setminus \{\alpha\}} c_{\alpha'} \right] \delta c_\alpha. \tag{101d}$$

We can investigate the sets M appearing in (101d) further. Taking all partitions of γ and then selecting those containing α yields the same set as taking all partitions of $\gamma \setminus \alpha$ and uniting them with $\{\alpha\}$, or expressed mathematically:

$$P(\gamma) \cap \{M \mid \alpha \in M\} = P(\gamma \setminus \alpha) \cup \{\alpha\}. \tag{101e}$$

The appearance of both the set α and the set containing α , $\{\alpha\}$, is due to the fact that M is a set of subsets of γ while γ and α are simply sets of binding sites, making the separation necessary. For all that follows, we will use the simple terminology $\gamma - \alpha$ whenever we want to express that the set α is not included. Since we can construct the sets M from $P(\gamma - \alpha)$ as in (101e), we do not need the partition of the full set γ in (101d). This is especially useful since the product over α' can then be simplified via (96). We calculate

$$\begin{aligned}
(A) &= - \int \sum_{\gamma \subseteq^* \Gamma} \sum_{\alpha \subseteq^* \gamma} \sum_{M \in P(\gamma - \alpha)} \rho_0 \left[\prod_{\alpha' \in M} c_{\alpha'} \right] \delta c_\alpha \\
&= - \int \sum_{\gamma \subseteq^* \Gamma} \sum_{\alpha \subseteq^* \gamma} \rho_{\gamma - \alpha} \delta c_\alpha \\
&= - \int \sum_{\alpha \subseteq^* \Gamma} \sum_{\substack{\gamma \subseteq^* \Gamma \\ \gamma \supseteq \alpha}} \rho_{\gamma - \alpha} \delta c_\alpha.
\end{aligned} \tag{101f}$$

On the last line, we swapped the sums over α and γ , where we now sum over all the subsets of Γ and then all sets containing these subsets. If we consider all summands multiplied by c_α for some fixed α , we find sums of the form

$$\begin{aligned}
&\rho_{\alpha - \alpha} + \rho_{\alpha + A - \alpha} + \rho_{\alpha + B - \alpha} + \dots + \rho_{\alpha + A + B - \alpha} + \dots \\
&= \rho_0 + \rho_A + \rho_B + \dots + \rho_{AB} + \dots,
\end{aligned} \tag{101g}$$

with $A, B \in (\Gamma - \alpha)$. This is exactly $\sigma_{\Gamma - \alpha}$ by (97). Therefore, we find when including (B) and (C) again

$$\begin{aligned}
\delta \ln \Xi &= - \int \sum_{\alpha \subseteq^* \Gamma} \sigma_{\Gamma - \alpha} \delta c_\alpha - \int \sigma_\Gamma \delta c_0 + \int \delta \rho \\
&= - \int \sum_{\alpha \subseteq \Gamma} \sigma_{\Gamma - \alpha} \delta c_\alpha + \int \delta \rho \\
&= - \int \sum_{\alpha \subseteq \Gamma} \delta [\sigma_{\Gamma - \alpha} c_\alpha] + \int c_\alpha \delta \sigma_{\Gamma - \alpha} + \int \delta \rho \\
&\stackrel{(100)}{=} \delta \left\{ \int \rho - \int \sum_{\alpha \subseteq \Gamma} \sigma_{\Gamma - \alpha} c_\alpha + c^{(0)} \right\}.
\end{aligned} \tag{101h}$$

If we ignore an additive constant and properly spell the integrals, this yields the desired equation

$$\ln \Xi = \int d(1) \rho(1) - \int d(1) \sum_{\alpha \subseteq \Gamma} \sigma_{\Gamma-\alpha}(1) c_{\alpha}(1) + c^{(0)}. \quad (102)$$

This can be paired with (54) to find an expression for

$$\begin{aligned} \beta A &= \int d(1) \ln(z(1) \lambda_{\text{th}}^3) \rho(1) - \int d(1) \rho(1) \\ &\quad + \int d(1) \sum_{\alpha \subseteq \Gamma} \sigma_{\Gamma-\alpha}(1) c_{\alpha}(1) - c^{(0)} \\ &= \int d(1) [\ln(\rho_0(1) \lambda_{\text{th}}^3) - 1] \rho(1) \\ &\quad + \int d(1) \sum_{\substack{\alpha \subseteq \Gamma \\ \alpha \neq \emptyset}} \sigma_{\Gamma-\alpha}(1) c_{\alpha}(1) - c^{(0)}. \end{aligned} \quad (103)$$

There is also a representation of A entirely independent of $c_{\alpha}(1)$. It can be derived from (103) by inserting a reformulation of c_{α} in terms of the σ_{γ} for $\gamma \subseteq \alpha$. The approach prominently features Clifford's dual numbers, which favor some rather elegant manipulations of graph sums, yet they are more technical than what we used so far. We will not go into the details here and simply name the final result that

$$\beta A = \int d(1) [\sigma_{\Gamma}(1) \ln(\sigma_0(1) \lambda_{\text{th}}^3) + Q(1)] - c^{(0)}, \quad (104)$$

with

$$\begin{aligned} Q(1) &= - \sum_{A \in \Gamma} \sigma_{\Gamma-A}(1) \\ &\quad + \sigma_0(1) \sum_{\substack{\{\gamma\} \in P(\Gamma) \\ \#\{\gamma\} \geq 2}} (-1)^M (M-2)! \prod_{\gamma} \frac{\sigma_{\gamma}(1)}{\sigma_0(1)}. \end{aligned} \quad (105)$$

Here, $\#\{\gamma\}$ is the cardinality of the set $\{\gamma\}$, so the improper partition $\{\gamma\} = \{\Gamma\}$ is not included. With (104) it is possible to express $\ln \Xi$ independent of $c_{\alpha}(1)$ as well. We can use the *Gibbs free energy* G , which is related to A by

$$G = E - TS + pV = A + pV = \mu N. \quad (106)$$

With the differential of A ,

$$dA = -SdT - pdV + \mu dN, \quad (107)$$

we arrive at

$$\left. \frac{\partial A}{\partial N} \right|_{T,V} = \mu. \quad (108)$$

This also means

$$G = N \left. \frac{\partial A}{\partial N} \right|_{T,V} = \rho \left. \frac{\partial A}{\partial \rho} \right|_{T,V}, \quad (109)$$

leading to a graph representation of (106) by

$$\ln \Xi = \beta G - \beta A = \int d(1) \sigma_{\Gamma}(1) \frac{\delta(\beta A)}{\delta \sigma_{\Gamma}(1)} - \beta A. \quad (110)$$

We consider the variables $\sigma_{\alpha}(1)$ to be independent from one another, which yields

$$\ln \Xi = - \int d(1) Q(1) + c^{(0)} - \int d(1) \sigma_{\Gamma}(1) \frac{\delta c^{(0)}}{\delta \sigma_{\Gamma}(1)}. \quad (111)$$

(102) and (103) or (111) and (104) are the two fundamental equations of the multiple density formalism. Of course, the two density formalism remains a special case. The advantages of multiple densities over the one density approach of the Mayer formalism are discussed in Wertheim I and III. Especially the topic of steric incompatibility is approached systematically.

Instead, we are now interested in putting the equations obtained so far straight to use.

5 Thermodynamic Perturbation Theory

5.1 TPT Equations for the Multiple Density Approach

For the following, we will consider a system with the potential (68) with only two sites A and B per particle and with the additional condition

$$\phi_{AA} = \phi_{BB} = 0. \quad (112)$$

The repulsive part of the potential will be of spherical form, as given by (70) with the hard sphere diameter D . (112) ensures that only one unique type of attractive bond f_{AB} can link two particles. Furthermore, the system is in a grand canonical ensemble within the volume V at temperature T and density ρ . We call this system the *real system* and are interested in calculating its pressure p .

We also impose some more constraints to which graphs may be non-vanishing. For geometric reasons (i.e. steric incompatibility, cf. Wertheim IV⁴) it is impossible that one site is bonded to more than one other site in a different particle. Additionally, there may only be one attraction bond between two particles, so we do not allow for any double bonds. These restrictions are reasonable for attraction sites not too close to each other, that is with a sufficiently large angle between them. We will also ignore rings in our considerations.

In TPT, we are interested in systematically approaching the calculation of properties like the pressure p or the Helmholtz free energy A – for which we have expressions with infinite sums of graphs – by equations containing only a finite amount of

graphs. This usually means that the remainder of the infinite sum is considered negligible in the system being examined. For example, when decreasing the density of our real system, graphs of more than two hyperpoints will soon contribute only very insignificantly to $\ln \Xi$, and when further decreasing the density, all site-site interaction will become completely negligible. The hyperpoint-hyperpoint interaction remains important, since it reduces the effective volume the particles are contained in. This means that a hard sphere fluid will eventually describe the system accurately.

Since the hard sphere fluid is easy to treat and there are indeed sophisticated analytical expressions for its various properties, we can use it as a *reference system* in the study at hand. It is described by the purely repulsive potential of the reference system without any attraction. It has the same particle density ρ and Temperature T as the real system. Its pressure p_R can be reasonably well described by the *Carnahan-Starling* EOS⁸ for the *compressibility factor* Z_R of the reference system

$$Z_R := \frac{\beta p_R}{\bar{\rho}} = \frac{1 + \eta + \eta^2 - \eta^3}{(1 - \eta)^3} \quad (113)$$

with the *packing density*

$$\eta = \frac{1}{6} \pi \bar{\rho} D^3 \quad (114)$$

and

$$\bar{\rho} = \rho \Omega, \quad (115)$$

with Ω being the integral over angular orientations of the particle. The motivation of choosing this value will arise at a later point.

The pressure in the real system is described by (111), with the Since they both obey (111) (with $Q_R(1) = 0$), we can quantify the difference between the real system and the reference system by

$$\begin{aligned} \beta(p - p_R)V = & \int d(1) \left[-\sigma_\Gamma(1) + \sigma_A(1) + \sigma_B(1) \right. \\ & \left. - \frac{\sigma_A(1)\sigma_B(1)}{\sigma_0(1)} \right] + \Delta c^{(0)} \\ & - \int d(1) \sigma_\Gamma \frac{\delta \Delta c^{(0)}}{\delta \sigma_\Gamma(1)}, \end{aligned} \quad (116)$$

where we inserted the result for $Q(1)$ with two sites and (112). Obviously, a new quantity of interest is

$$\Delta c^{(0)} = c^{(0)} - c_R^{(0)}. \quad (117)$$

$c_R^{(0)}$ is the graph sum we defined in (52). The $\Delta c^{(0)}$ is the sum of all graphs in $c^{(0)}$ not purely interacting by repulsion. Equation (116) contains three unknowns, σ_0 , σ_A and σ_B along with

$\Delta c^{(0)}$, which is itself a function of the latter. This function contains infinite sums of graphs, so we are not able to obtain results without simplifications. To that end, let us group the graphs in $\Delta c^{(0)}$ hierarchically. To that end, we will from now on only consider the graphs in $\Delta c^{(0)}$ with a single *s-meric chain* for $s \geq 2$. By chain we mean singly and not more than singly connected *s-mers*. Ignoring the graphs with more than one chain and possible attractive intraconnection of *s-mers* will not be reasonable when self interaction and chain-chain interactions play an important role in modelling the fluid. With this approximation, we can isolate all remaining graphs containing a dimer, trimer etc. in $\Delta c^{(0)}$. We call these subclasses κ_s . Obviously,

$$\Delta c_{\text{TPT}}^{(0)} = \sum_{s \geq 2} \kappa_s. \quad (118)$$

We wrote $\Delta c_{\text{TPT}}^{(0)}$ to clarify that we will from now on deal with approximations as opposed to physically reasonable assumptions. We can truncate the above series at some S , which we will call the $S - 1$ -th order of TPT. First order TPT only retains κ_2 , which is

$$\begin{aligned} \kappa_2 = & \text{Diagram 1} + \text{Diagram 2} \\ & + \text{Diagram 3} + \text{Diagram 4} \\ & + \frac{1}{2} \text{Diagram 5} + \frac{1}{2} \text{Diagram 6} \\ & + \dots \end{aligned} \quad (119)$$

We can compare this result to what we found for the hard sphere correlation function $g_R(12)$ in (67), then we have

$$\begin{aligned} \kappa_2 = & \int d(1)d(2) \text{Diagram 1} \times g_R(12) \\ = & \int d(1)d(2) \sigma_A(1) f_{AB}(12) \sigma_B(2) g_R(12) \end{aligned} \quad (120)$$

$g_R(12)$ reduces the infinite graph sum κ_2 to a single graph including a g_R -bond. Similar arguments can be adjusted to higher

order single- s -mer graph sums. Since we will also treat second order TPT, we will later need

$$\kappa_3 = \int d(1)d(2)d(3) \sigma_B(1)\sigma_0(2)\sigma_A(3) \times f_{AB}(12)f_{AB}(23)[g_R(123) - g_R(12)g_R(23)]. \quad (121)$$

In the chain-free approximation, we can also simplify the expression for the excess pressure (116). Independent of the TPT order, all graphs in $\Delta c_{\text{TPT}}^{(0)}$ will only contain exactly one point with property σ_A and one with property σ_B . Therefore, derivation of $\Delta c_{\text{TPT}}^{(0)}$ with respect to either σ_A or σ_B will not change the symmetry number and we can recover $\Delta c_{\text{TPT}}^{(0)}$ by

$$\Delta c_{\text{TPT}}^{(0)} = \int d(1)\sigma_B(1) \frac{\delta \Delta c_{\text{TPT}}^{(0)}}{\delta \sigma_B(1)}. \quad (122)$$

But since

$$\frac{\delta \Delta c_{\text{TPT}}^{(0)}}{\delta \sigma_B(1)} = c_A(1) = \frac{\sigma_A(1) - \sigma_0(1)}{\sigma_0(1)}, \quad (123)$$

(116) reduces to

$$\beta(p - p_R)V = \int d(1) \left[-\sigma_\Gamma(1) + \sigma_A(1) \right] - \int d(1)\sigma_\Gamma \frac{\delta \Delta c_{\text{TPT}}^{(0)}}{\delta \sigma_\Gamma(1)}. \quad (124)$$

An analogous formula could be derived when replacing all appearances of σ_A by σ_B .

Before we proceed to establish the first and second order TPT, we will motivate the approach we will take. Since we only consider graphs containing single chains, σ_A and σ_B will appear exactly once in each graph of κ_s . This will always lead to the mathematical conclusion $\sigma_G := \sigma_A = \sigma_B$, which is also physically sensible: since $\sigma_A = \rho_0 + \rho_A$, it contains monomers and the density of particles that are bonded at A . The latter only appear in chains, but every chain containing a particle bonded at A must also contain a particle bonded at B , and since branching is impossible, there is exactly one particle bonded at A and one bonded at B . Simply put, a chain beginning at an A must end at a B . We can therefore understand ρ_G as the density of chains and σ_G as the density of what we shall call *units*, that is either monomers or chains. We can also use that the space-dependency of the functions σ_G and ρ_G was only an artifact of the auxiliary space-dependency of z . Therefore, the integrals in (124) are just constant factors and we get

$$\beta(p - p_R)V = -\sigma_\Gamma V\Omega + \sigma_A V\Omega - \sigma_\Gamma \int d(1) \frac{\delta \Delta c_{\text{TPT}}^{(0)}}{\delta \sigma_\Gamma(1)}. \quad (125)$$

Ω is again the integral over all particle orientations. It will appear often, therefore we define

$$\left. \begin{aligned} \bar{\rho}_\alpha &:= \rho_\alpha \Omega \\ \bar{\sigma}_\alpha &:= \sigma_\alpha \Omega \end{aligned} \right\} \forall \alpha \in \{0, A, B, AB\} \quad (126)$$

With the system at hand, we can therefore easily access the compressibility factor by

$$\frac{\beta(p - p_R)}{\bar{\rho}} = -1 + \frac{\bar{\sigma}_G}{\bar{\rho}} - \frac{1}{V} \int d(1) \frac{\delta \Delta c_{\text{TPT}}^{(0)}}{\delta \bar{\sigma}_\Gamma(1)}. \quad (127)$$

When using the values $\bar{\sigma}_\alpha$ instead of σ_α , we have to keep in mind their appearance in κ_s , which means that the integrals we introduced have to be divided by the correct power of Ω . This can be viewed as a normalization process, we will use it in the next section.

Now, one quantity of special interest is

$$\nu := \frac{\bar{\rho}}{\bar{\sigma}_G} = \frac{\bar{\sigma}_{AB}}{\bar{\sigma}_G}. \quad (128)$$

It is the ratio between the total density and the density of units. This could be understood as the mean length of units (the mean chain length) in the system. Therefore, necessarily $\nu \geq 1$. We can use another ratio, namely the ration between the unit density and the monomer density,

$$\tau := \frac{\bar{\sigma}_G}{\bar{\sigma}_0}. \quad (129)$$

It will be of help in finding the value of ν . First, we need to know

$$\begin{aligned} \bar{\sigma}_{AB} &= \bar{\rho}_0 + \bar{\rho}_A + \bar{\rho}_B + \bar{\rho}_{AB} \\ &= \bar{\rho}_0 + \bar{\rho}_A + \bar{\rho}_B + \bar{\rho}_0 c_{ACB} + \bar{\rho}_0 c_{CAB} \\ &= \frac{1}{\bar{\rho}_0} \left(\bar{\rho}_0 + \bar{\rho}_A + \bar{\rho}_B + \frac{\bar{\rho}_A \bar{\rho}_B}{\bar{\rho}_0^2} \right) + \bar{\rho}_0 c_{AB} \\ &= \frac{1}{\bar{\rho}_0} (\bar{\rho}_0 + \bar{\rho}_A) (\bar{\rho}_0 + \bar{\rho}_B) + \bar{\rho}_0 c_{AB} \\ &= \frac{\bar{\sigma}_A \bar{\sigma}_B}{\bar{\sigma}_0} + \bar{\sigma}_0 c_{AB} = \frac{\bar{\sigma}_G^2}{\bar{\sigma}_0} + \bar{\sigma}_0 c_{AB}. \end{aligned} \quad (130)$$

This means for the difference

$$\begin{aligned} \nu - \tau &= \frac{\bar{\sigma}_{AB}}{\bar{\sigma}_G} - \frac{\bar{\sigma}_G}{\bar{\sigma}_0} = \frac{\bar{\sigma}_0 \bar{\sigma}_{AB} - \bar{\sigma}_G^2}{\bar{\sigma}_0 \bar{\sigma}_G} \\ &= \frac{\bar{\sigma}_0^2 c_{AB}}{\bar{\sigma}_0 \bar{\sigma}_G} = \frac{\bar{\sigma}_0 c_{AB}}{\bar{\sigma}_G} = \frac{c_{AB}}{\tau}, \end{aligned} \quad (131)$$

while we know for τ itself

$$\begin{aligned} \tau &= \frac{\bar{\sigma}_G}{\bar{\sigma}_0} = \frac{\bar{\rho}_0 + \bar{\rho}_G}{\bar{\rho}_0} \\ &= 1 + \frac{\bar{\rho}_G}{\bar{\rho}_0} = 1 + c_G. \end{aligned} \quad (132)$$

We found equations for v and τ containing $\bar{\sigma}_G$ and $\bar{\sigma}_0$ and simple derivatives of $\Delta c_{\text{TPT}}^{(0)}$. The relationship between $\Delta c_{\text{TPT}}^{(0)}$ and c_G was already explored in (122) and (123). From there, we can conclude by the calculation we just did that

$$\Delta c_{\text{TPT}}^{(0)} = \int d(1) \bar{\sigma}_G \frac{\delta \Delta c_{\text{TPT}}^{(0)}}{\delta \bar{\sigma}_G} = \bar{\sigma}_G \int d(1) (\tau - 1). \quad (133)$$

The functional derivative with respect to $\bar{\sigma}_{AB}$ appearing in (125) can simply be considered to be

$$\frac{\delta}{\delta \bar{\sigma}_{AB}} \int d(1) (\tau - 1) = \frac{\delta(\tau - 1)}{\delta \bar{\sigma}_{AB}}. \quad (134)$$

With all this knowledge, we can rearrange (127)

$$\frac{\beta(p - p_R)}{\bar{\rho}} = -1 + v^{-1} - \bar{\sigma}_G \frac{\partial(\tau - 1)}{\partial \bar{\sigma}_{AB}}. \quad (135)$$

We can conclude that solving the equations for $\tau - 1$ and $v - \tau$ yields the desired TPT approximation to the excess pressure or more accurately the excess compressibility factor. How they are computed and to what final expressions we can get will be explained in the following two examples of first and second order TPT.

5.2 First Order TPT

Now, we know that

$$\Delta c_{\text{TPT1}}^{(0)} = \kappa_2. \quad (136)$$

We can derive c_α , $\alpha \neq 0$, via (100) to be

$$\begin{aligned} c_A &= \bar{\sigma}_A I_1, \\ c_B &= \bar{\sigma}_B I_1, \\ c_{AB} &= 0. \end{aligned} \quad (137)$$

With the integral

$$I_1 = \frac{1}{\Omega} \int d(2) f_{AB}(12) g_R(12). \quad (138)$$

We ignore the artificial spatial dependencies of our arguments and treat them as constants. By (96), we know about the densities

$$\begin{aligned} \bar{\rho}_A &= \bar{\rho}_0 \bar{\sigma}_A I_1, \\ \bar{\rho}_B &= \bar{\rho}_0 \bar{\sigma}_B I_1, \\ \bar{\rho}_{AB} &= \bar{\rho}_0 \bar{\sigma}_A \bar{\sigma}_B I_1^2 = \frac{\bar{\rho}_A \bar{\rho}_B}{\bar{\rho}_0}. \end{aligned} \quad (139)$$

(98) links ρ_α to $\bar{\sigma}_\alpha$, leaving us with equations for $\bar{\sigma}_A$, $\bar{\sigma}_B$ and $\bar{\sigma}_0$ and the known $\bar{\sigma}_{AB}$ and I_1 ,

$$\begin{aligned} \frac{\bar{\sigma}_A - \bar{\sigma}_0}{\bar{\sigma}_A} &= \bar{\sigma}_0 I_1, \\ \frac{\bar{\sigma}_B - \bar{\sigma}_0}{\bar{\sigma}_B} &= \bar{\sigma}_0 I_1, \\ \bar{\sigma}_{AB} \bar{\sigma}_0 &= \bar{\sigma}_A \bar{\sigma}_B. \end{aligned} \quad (140)$$

The first two of these lead to the identity

$$1 - \frac{\bar{\sigma}_0}{\bar{\sigma}_A} = 1 - \frac{\bar{\sigma}_0}{\bar{\sigma}_B} \iff \bar{\sigma}_A = \bar{\sigma}_B =: \bar{\sigma}_G, \quad (141)$$

which we already discussed.

The only remaining unknowns are now $\bar{\sigma}_0$ and $\bar{\sigma}_G$. To proceed, we now introduce v and τ . We can use (140) to calculate for TPT1

$$v = \frac{\bar{\sigma}_G^2}{\bar{\sigma}_0 \bar{\sigma}_G} = \frac{\bar{\sigma}_G}{\bar{\sigma}_0} = \tau, \quad (142)$$

in agreement with (131),

$$v - \tau = \frac{\bar{\sigma}_0 c_{AB}}{\bar{\sigma}_G} = 0. \quad (143)$$

We are left with the single unknown quantity v . We can reexpress $\bar{\sigma}_0$ and $\bar{\sigma}_G$ by v and $\bar{\rho}$ with (128) and (142)

$$\begin{aligned} \bar{\sigma}_G &= \bar{\rho} v^{-1}, \\ \bar{\sigma}_0 &= \bar{\sigma}_G v^{-1} = \bar{\rho} v^{-2}. \end{aligned} \quad (144)$$

We can make insertions in one of the first two equations of (140) to obtain

$$1 - v^{-1} = \bar{\rho} v^{-2} I_1. \quad (145)$$

This equation has two solutions, of which we drop the negative one to arrive at

$$v = \frac{1}{2} + \sqrt{\frac{1}{4} + \bar{\rho} I_1}. \quad (146)$$

While this equation is in principle useful to determine v from a given system, it will not be used here. The reason is that v can be easily given as a parameter to a simulation, which is not that simple for the value of the integral I_1 , which depends for a given reference system on the function f_{AB} . So, treating v as a given variable, we can use (132) to reduce (135) to

$$\frac{\beta(p - p_R)}{\bar{\rho}} = -1 + v^{-1} - \bar{\sigma}_G^2 I_1'. \quad (147)$$

which is now only dependent on I_1' , v and $\bar{\rho}$. The prime denotes differentiation of I_1 with respect to $\bar{\rho}$. An important case is rigid chain formation (Wertheim calls this the *glue spot limit*), where f_{AB} is just a constant times a δ peak centered at some point at the particle's boundary and therefore

$$I_1 = \frac{1}{\Omega} \int d(2) K \delta(2 - [\mathbf{r}_1 + \mathbf{d}(\Omega_1)]) g(2) = K g(D). \quad (148)$$

An equation ignoring this constant is desirable. And since

$$\begin{aligned} (1 - v^{-1}) \bar{\rho} &= \bar{\rho} - \bar{\sigma}_G = \bar{\rho}_G + \bar{\rho}_{AB} = \bar{\sigma}_G \bar{\rho}_0 I_1 + \frac{\bar{\rho}_G^2}{\bar{\rho}_0} I_1 \\ &= \bar{\sigma}_G (\bar{\rho}_0 + \bar{\rho}_G) I_1 = \bar{\sigma}_G^2 I_1, \end{aligned} \quad (149)$$

we can write

$$\frac{\beta(p - p_R)}{\bar{\rho}} = (-1 + v^{-1}) \left(1 + \bar{\rho} \frac{I'_1}{I_1} \right). \quad (150)$$

Since the Carnahan-Starling EOS depends on η instead of $\bar{\rho}$, we are more interested in an equation also depending on η . By (114), the variables are linearly dependent, therefore we get

$$\frac{\beta(p - p_R)}{\bar{\rho}} = (-1 + v^{-1}) \left(1 + \eta \frac{I'_1}{I_1} \right), \quad (151)$$

where the prime now denotes differentiation with respect to η . In the glue spot limit, we can use

$$g(D) = \frac{1 - \frac{1}{2}\eta}{(1 - \eta)^3}, \quad (152)$$

which means alongside (148) that

$$\frac{\beta(p - p_R)}{\bar{\rho}} = (-1 + v^{-1}) \left(1 - \frac{\eta}{2 - \eta} + \frac{3\eta}{1 - \eta} \right). \quad (153)$$

This is the first order TPT result. It required some argumentation, but at that point we can see that the Wertheim formalism combined with the Carnahan Starling EOS reduced infinite sums of graphs to a comparably simple equation in the two variables v and η . Note, however, that these results are not analytical even for systems that only contain dimers and monomers because we ignored all dimer-dimer interaction.

5.3 Second Order TPT

For the second order of TPT,

$$\Delta c_{\text{TPT2}}^{(0)} = \kappa_2 + \kappa_3. \quad (154)$$

We face another integral of the form

$$I_2 := \frac{1}{\Omega^2} \int d(2)d(3) f_{AB}(12) f_{AB}(23) \times [g_R(123) - g_R(12)g_R(23)]. \quad (155)$$

It can be shown that

$$\frac{1}{\bar{\sigma}_A \bar{\sigma}_B} \frac{\delta \kappa_3}{\delta \bar{\sigma}_0} = \frac{1}{\bar{\sigma}_B \bar{\sigma}_0} \frac{\delta \kappa_3}{\delta \bar{\sigma}_A} = \frac{1}{\bar{\sigma}_A \bar{\sigma}_0} \frac{\delta \kappa_3}{\delta \bar{\sigma}_B} = I_2. \quad (156)$$

With that, we find

$$\begin{aligned} c_A &= \bar{\sigma}_A I_1 + \bar{\sigma}_0 \bar{\sigma}_A I_2, \\ c_B &= \bar{\sigma}_B I_1 + \bar{\sigma}_0 \bar{\sigma}_B I_2, \\ c_{AB} &= \bar{\sigma}_A \bar{\sigma}_B I_2. \end{aligned} \quad (157)$$

This means

$$\begin{aligned} \bar{\rho}_A &= \bar{\rho}_0 (\bar{\sigma}_A I_1 + \bar{\sigma}_0 \bar{\sigma}_A I_2), \\ \bar{\rho}_B &= \bar{\rho}_0 (\bar{\sigma}_B I_1 + \bar{\sigma}_0 \bar{\sigma}_B I_2), \\ \bar{\rho}_{AB} &= \frac{\bar{\rho}_A \bar{\rho}_B}{\bar{\rho}_0} + \bar{\rho}_0 \bar{\sigma}_A \bar{\sigma}_B I_2 \end{aligned} \quad (158)$$

and finally

$$\begin{aligned} \frac{\bar{\sigma}_A - \bar{\sigma}_0}{\bar{\sigma}_A} &= \bar{\sigma}_0 I_1 + \bar{\sigma}_0^2 I_2, \\ \frac{\bar{\sigma}_B - \bar{\sigma}_0}{\bar{\sigma}_B} &= \bar{\sigma}_0 I_1 + \bar{\sigma}_0^2 I_2, \\ \bar{\sigma}_{AB} \bar{\sigma}_0 &= \bar{\sigma}_A \bar{\sigma}_B + \bar{\sigma}_0^2 \bar{\sigma}_A \bar{\sigma}_B I_2. \end{aligned} \quad (159)$$

Again, mathematics yields the physical condition that $\bar{\sigma}_A = \bar{\sigma}_B$, so we will use $\bar{\sigma}_G$ again. We switch to v and τ to find

$$\begin{aligned} v - \tau &= \bar{\sigma}_0 \bar{\sigma}_G I_2, \\ \tau - 1 &= \bar{\sigma}_G I_1 + \bar{\sigma}_0 \bar{\sigma}_G I_2. \end{aligned} \quad (160)$$

This means for the compressibility

$$\frac{\beta(p - p_R)}{\bar{\rho}} = -1 + v^{-1} - \bar{\sigma}_G^2 I'_1 + \bar{\sigma}_0 \bar{\sigma}_G^2 I'_2 \quad (161)$$

This equation is not quite the one we want to end up with. For once, it contains I'_1 and I'_2 , which will depend on powers of some constant in the glue spot limit. And even more troublesome, I_2 and thus I'_2 contain $g_R(123)$, of which there is only very poor knowledge. The appearance of $g_R(123)$ can not entirely be mended, but we will look for a form in which we may be able to approximate this unknown value. We do so using the system of equations (160) and the definition of τ in (129), yielding

$$\begin{aligned} 2\tau - v - 1 &= \bar{\sigma}_G I_1, \\ \tau(v - \tau) &= \bar{\sigma}_G^2 I_2. \end{aligned} \quad (162)$$

In this equation, we can eliminate $\bar{\sigma}_G$ to find

$$(2\tau - v - 1)^2 \lambda = \tau(v - \tau). \quad (163)$$

The new factor

$$\lambda := \frac{I_2}{I_1^2} \quad (164)$$

will prove very valuable. It is especially independent of any constants arising in the integration of I_1 and I_2 . As we do not know much about I_2 , it makes sense considering it an unknown, which means that λ can be viewed as an unknown.

We now try to apply the simplification we used in the TPT1 case to introduce I'_1/I_1 into the compressibility expression. Again, we calculate

$$\begin{aligned} (1 - v^{-1}) \bar{\rho} &= \bar{\rho}_G + \bar{\rho}_{AB} = \bar{\rho}_G + \frac{\bar{\rho}_G^2}{\bar{\rho}_0} + \bar{\sigma}_0 \bar{\sigma}_G^2 I_2 \\ &= \frac{\bar{\rho}_G}{\bar{\rho}_0} \bar{\sigma}_G + \bar{\sigma}_0 \bar{\sigma}_G^2 I_2 = \bar{\sigma}_G^2 I_1 + 2\bar{\sigma}_0 \bar{\sigma}_G^2 I_2. \end{aligned} \quad (165)$$

On the other hand, we have by (131)

$$\bar{\sigma}_0 \bar{\sigma}_G^2 I_2 = \bar{\sigma}_0 c_{AB} = (v - \tau) \bar{\sigma}_0 \tau = \frac{v - \tau}{v} \bar{\rho}. \quad (166)$$

The derivatives of I_2 appear in λ , and we find

$$\lambda' = \left(\frac{I_2}{I_1^2} \right)' = I_2' \frac{1}{I_1^2} - 2I_1' \frac{I_2}{I_1^3} = \lambda \left(\frac{I_2'}{I_2} - 2 \frac{I_1'}{I_1} \right). \quad (167)$$

Now, putting all these calculations together, and going from $\bar{\rho}$ differentiation to η differentiation, we obtain

$$\frac{\beta(p - p_R)}{\bar{\rho}} = (-1 + v^{-1}) \left(1 + \eta \frac{I_1'}{I_1} \right) + \frac{v - \tau}{v} \frac{\eta \lambda'}{\lambda}. \quad (168)$$

We will use this equation for the excess compressibility factor. It avoids the problems we discussed about equation (161) as good as possible, though now the value τ is explicitly present. However, by (163), τ is the solution of a polynomial equation of second order. It can be given dependent on v and λ by

$$\tau = v \left[\frac{1}{2} + \frac{2\lambda}{(1 + 4\lambda)v} + \frac{\sqrt{1 + 4\lambda - 4\lambda v^{-2}}}{2(1 + 4\lambda)} \right], \quad (169)$$

which means

$$\frac{v - \tau}{v} = \frac{1}{2} - \frac{2\lambda}{(1 + 4\lambda)v} - \frac{\sqrt{1 + 4\lambda - 4\lambda v^{-2}}}{2(1 + 4\lambda)}. \quad (170)$$

For small λ , we can use a Taylor expansion of this expression. Then, the excess compressibility factor is given in a first order approximation by

$$\begin{aligned} \frac{\beta(p - p_R)}{\bar{\rho}} &= (-1 + v^{-1}) \left(1 + \eta \frac{I_1'}{I_1} \right) \\ &+ (-1 + v^{-1})^2 \eta \lambda' + O(\lambda' \lambda). \end{aligned} \quad (171)$$

Wertheim states that λ is sufficiently small to truncate the series at this. In the glue spot limit, which will be our concern for the next section, we can reuse the TPT1 equations and drop the residual to obtain

$$\begin{aligned} \frac{\beta(p - p_R)}{\bar{\rho}} &= (-1 + v^{-1}) \left(1 - \frac{\eta}{2 - \eta} + \frac{3\eta}{1 - \eta} \right) \\ &+ (-1 + v^{-1})^2 \eta \lambda'. \end{aligned} \quad (172)$$

5.3.1 Looking for λ

λ' is the last brick missing in the wall to solve (172). Wertheim himself proposes the expression

$$\lambda_{WH} = 0.234\eta(1 + 1.482\eta). \quad (173)$$

It comes from an analytical calculation he performed in the limit $\eta = 0$ and a fitting this result to data of Uehara *et al.*⁹

The approach we chose is, admittedly, less rigorous. We considered system with chain lengths $v \in \{4, 8, 16\}$. For each

chain length v , we performed $K^v \in \mathbb{N}$ simulations¹⁰. The temperature and volume of the systems were kept fix for all simulations, they only differ in a set $\{\eta_i^v \mid 1 \leq i \leq K^v\}$ of packing densities. We therefore obtained pairs $(\eta_i^v, Z_{SIM}(v, \eta_i^v))$. The values for the packing density were always $0 \leq \eta_i^v < 4$, where the maximum value for η_i^v decreased with increasing chain length. For most of what follows, we drop the index i again and consider the equations to be understood as pointwise. There should not be too much ambiguity arising from that, and notation is shorter.

With the simulated Z_{SIM} , we can calculate the optimal λ'_{opt} , that is the λ' reproducing the simulation results in the TPT2 approach.

$$\lambda'_{opt} = \frac{Z_{SIM} - Z_R}{\eta(v^{-1} - 1)^2} - \frac{1}{v^{-1} - 1} \left(\frac{1}{\eta} - \frac{1}{2 - \eta} + \frac{3}{1 - \eta} \right), \quad (174)$$

in which Z_R is given by the Carnahan-Starling EOS (113) as usual. The remaining task was to fit a function for λ' to (174). This idea has the advantage of – potentially – approximating the simulation data very accurately. One of the key disadvantages is that it demands that the TPT2 order should be well converged to the actual result, so it is not improvable by systematically including higher order TPT corrections, and indeed its demand of accurate TPT2 is physically wrong.

Nevertheless, if physical insight is not of highest priority, fitting to λ'_{opt} should be qualified to yield good results. We are also not interested in λ itself, so we do not have to be able to recover λ from the function λ' we choose to approximate λ'_{opt} .

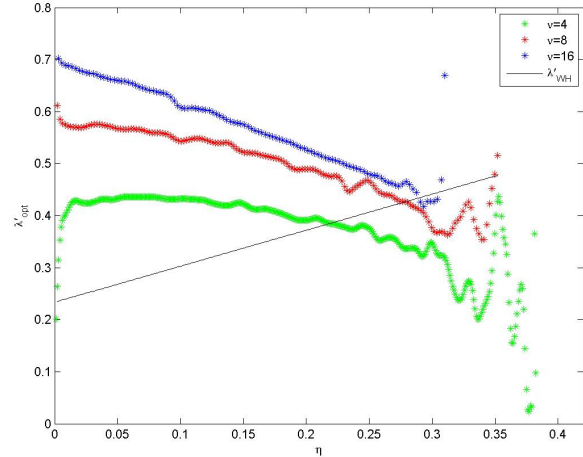


Figure 1 The optimal λ'_{opt} values for different chain lengths by (174) and the Wertheim approximation λ'_{WH} by (173).

There are four important aspects that have to be considered before one can properly tackle the problem of approximation:

v dependence. There are two main classes of approximations: those which, like the actual λ and λ' , only depend on η ;

and those which also depend on v . We call the first class v fix functions. It will contain λ'_4 , λ'_8 and λ'_{16} , which are approximations to the λ'_{opt} for $v = 4, 8$ and 16 , respectively. We also consider λ'_{all} , which is a v independent approximation, but fitted not only to one but to all three λ'_{opt} at once. And then there is the function that contains a v independence, λ'_v . We call it simply the v dependent function. Since it has the most degrees of freedom, the v dependent function must be potentially superior to the v fix functions in an overall context, while the v fix functions will obviously be better for the v value they were fitted to (which excludes λ'_{all}).

Data selection. Consider Figure 1. The stars are simulation data points. We can see that for each v the simulation results have increasing noise towards high η . For $v = 8$ and especially $v = 16$ they also change strongly at low η . We therefore deem it sensible to restrict the fitting efforts to the areas where λ'_{opt} is well-behaved and to neglect data points where η is too high and also those with very small η .

Residual definition. We want to find the optimum λ' with respect to λ'_{opt} . But the optimum needs further definition. There are of course numerous possibilities for the fitting function, of which we chose a simple *least squares fit*, that is we try to minimize the *root mean square deviation* (RMS)

$$\Delta_{RMS}^v[\lambda'] := \sqrt{\frac{1}{K^v} \sum_{i=1}^{K^v} (\lambda'_{opt}(\eta_i^v, v) - \lambda'(\eta_i^v, v))^2}, \quad (175a)$$

$$\Delta_{RMS}[\lambda'] := \sqrt{\sum_v \frac{1}{K^v} \sum_{i=1}^{K^v} (\lambda'_{opt}(\eta_i^v, v) - \lambda'(\eta_i^v, v))^2}. \quad (175b)$$

The two definitions are either for a single v or for all v . We will need both. It would also be possible to include a weighting function in (175), but since we do not consider any η interval to be of higher or lower importance than another (except for those we already neglected), we have no reason to use such a function.

Functional form. The approximation potential does of course highly depend on the functional form we allow our approximative functions λ' to take. Wertheim chose a first order polynomial. One can see in Figure 1 that his coefficients do not reproduce λ'_{opt} very accurately (which was never his intention). But even with better coefficients, the curvature of the λ'_{opt} will not be ideally reproduced. However, a fit of this type is useful. We call them first order polynomial fit and define

$$\lambda'_{p1} := c_1^{p1} + c_2^{p1} \eta. \quad (176)$$

For a v dependent λ the coefficients will be v dependent. Analogously, we can use a second order polynomial

$$\lambda'_{p2} := c_1^{p2} + c_2^{p2} \eta + c_3^{p2} \eta^2. \quad (177)$$

Finally, we tried a third approach with a Gaussian bell curve. The motivation to use this function was mostly that the $v = 4$

and $v = 8$ curves for λ'_{opt} seemed to increase in decay toward high η values, and that especially for high η the polynomial functions would increase towards infinite values, possible destroying the decent results obtained by TPT1, while the exponential curve just returns the TPT1 compressibility factors for high η . The exponential curve is given by the parameters

$$\lambda'_{ex} := c_1^{ex} \exp\left(-\frac{(\eta - c_3^{ex})^2}{c_2^{ex}}\right). \quad (178)$$

Now we have gathered the proper definitions and data selections to find approximating functions. Just a quick word on nomenclature: we will call the v fix functions $\lambda'_{v,F}$ with $v \in \{4, 8, 16, all\}$ and $F \in \{p1, p2, ex\}$. The v dependent functions will be called $\lambda'_{v,F}$. We will look at each functional form separately, first establishing the four v free approximations and guessing from them the v dependent approximation, which will then be fitted accordingly. MATLAB[®]¹¹ functions `fminunc` and `fmincon` are used to find the fitting function.

v fix fits. We collected the data for the first order polynomial, second order polynomial and Gaussian fit in Tables 1, 2 and 3, respectively. The RMS values are calculated by (175a) for $v = 4, 8, 16$ and by (175b) for $v = all$.

Table 1 Coefficients and RMS deviation for the polynomial fit of first order (176).

	c_1^{p1}	c_2^{p1}	RMS
$v = 4$	0.467	-0.441	0.0251
$v = 8$	0.608	-0.643	0.0159
$v = 16$	0.705	-0.897	0.0052
$v = all$	0.614	-0.687	0.2932

For the first order polynomial, the case $v = 16$ can already be well approached. As we can see in Figure 1, λ'_{opt} is close to linear in this case. $v = 4$, however, is more problematic. It is possible that this also amounts from the increasing oscillations toward higher η . When we consider the coefficients, we can see that they show monotonous behavior. They also appear to be close to linear, that is, since we increased the number v exponentially, one could maybe describe them rather well by a logarithmic function. We will come back to that later.

For the second order polynomial, we can only slightly improve the convergence of the $v = 16$ fit while especially the $v = 4$ fit strongly improves. There is also some improvement in the $v = 8$ fit. Again, the coefficients follow a monotonous pattern, which is again close to linear (or rather, logarithmic).

The relation between the second order polynomial and the Gaussian fit function is more interesting. They are both superior to the first order polynomial, but the Gaussian function

Table 2 Coefficients and RMS deviation for the polynomial fit of second order (177).

	c_1^{p2}	c_2^{p2}	c_3^{p2}	RMS
$v = 4$	0.416	0.423	-2.555	0.0106
$v = 8$	0.581	-0.156	-1.472	0.0107
$v = 16$	0.699	-0.754	-0.475	0.0041
$v = all$	0.587	-0.204	-1.436	0.1830

Table 3 Coefficients and RMS deviation for the Gaussian fit function (178).

	c_1^{ex}	c_2^{ex}	c_3^{ex}	RMS
$v = 4$	0.435	0.142	0.086	0.0116
$v = 8$	0.577	0.261	-0.007	0.0105
$v = 16$	0.779	0.480	-0.233	0.0042
$v = all$	0.584	0.264	-0.019	0.1832

describes the case $v = 8$ slightly better while it is not as good for the $v = 4$ case. They are equally good for describing the case $v = 16$ and also have similar results for $v = all$, at least when only comparing them in RMS value. Considering the Gaussian fit alone, we find again monotonous behavior in the coefficients. However, this case seems closer to an actual linear relation between the coefficients and v instead of $\ln[v]$.

The case $v = all$ is only poorly covered. The worst RMS deviation for a single v is still more than an order of magnitude smaller than the overall RMS for each approximation. We are interested in a fit that may interpolate λ_{opt} for v values between 4 and 16 (or even extrapolate), so a v dependent fit seems to be the only option. We will lose in accuracy for the single v cases, but we hope to improve the fit for the $v = all$ case.

v dependent fits. After considering the various v fix functions, we want to find v dependent versions. We already noted for the two polynomial cases that the coefficients seem to behave linearly on the logarithmic scale for v . We therefore assume

$$\begin{aligned}\lambda'_{v,p1} &= c_1^{p1}(v) + c_2^{p1}(v)\eta, \\ c_i^{p1}(v) &:= c_{i,1}^{p1} + c_{i,2}^{p1} \ln[v]\end{aligned}\quad (179)$$

for the first order polynomial. After fitting, we find for the coefficients

$$\begin{aligned}c_1^{p1}(v) &= 0.2446 + 0.1665 \ln[v], \\ c_2^{p1}(v) &= 0.0187 - 0.3276 \ln[v].\end{aligned}\quad (180)$$

The RMS values can be found in Table 4, as for the other two fits. We will compare them later.

We have analogously for the second order polynomial

$$\begin{aligned}\lambda'_{v,p2} &= c_1^{p2}(v) + c_2^{p2}(v)\eta + c_3^{p2}(v)\eta^2, \\ c_i^{p2}(v) &:= c_{i,1}^{p2} + c_{i,2}^{p2} \ln[v].\end{aligned}\quad (181)$$

The coefficients can be fitted to arrive at

$$\begin{aligned}c_1^{p2}(v) &= 0.1423 + 0.2013 \ln[v], \\ c_2^{p2}(v) &= 1.6104 - 0.8606 \ln[v], \\ c_3^{p2}(v) &= -4.6831 + 1.5520 \ln[v].\end{aligned}\quad (182)$$

In the Gaussian case, the coefficients do not seem to be that well-described with logarithmic functions. We will rather use linear relations between the c^{ex} and v , as already noted above. This means

$$\begin{aligned}\lambda'_{v,p2} &= c_1^{ex}(v) \exp\left(-\frac{(\eta - c_3^{ex}(v))^2}{c_2^{ex}(v)}\right), \\ c_i^{ex}(v) &:= c_{i,1}^{ex} + c_{i,2}^{ex} v.\end{aligned}\quad (183)$$

Our coefficient suggestion is then

$$\begin{aligned}c_1^{ex}(v) &= 0.2319 + 0.0576v, \\ c_2^{ex}(v) &= 0.4216 + 0.0425v, \\ c_3^{ex}(v) &= 0.0732 - 0.0503v.\end{aligned}\quad (184)$$

Table 4 RMS deviations for all v dependent fits. RMS values are given for sets of only one value for v as well as all values of v in the row $v = all$.

	$\lambda_{v,p1}$	$\lambda_{v,p2}$	$\lambda_{v,ex}$
$v = 4$	0.0244	0.0120	0.0201
$v = 8$	0.0258	0.0255	0.0146
$v = 16$	0.0057	0.0047	0.0062
$v = all$	0.0560	0.0422	0.0410

Now, we succeeded in finding v dependent fits. Comparison of Table 4 to Tables 1, 2 and 3 shows that the loss in accuracy for each isolated v is well outweighed by the gain in accuracy for the overall RMS deviation. The overall deviation is still greater than each single deviation, but the difference is now roughly a factor of 2 from the worst single v approximation as compared to a factor greater than 10.

So, finally, there is only the question of which fit is best. The answer, however, is not that simple. We neither have further experimental nor simulation data to compare the results to. We try to give a few arguments: First of all, since none of the three fits is particularly complex, the slightly lesser degree of complexity in the first order polynomial is not a real advantage. The other two fits are slightly more accurate in RMS deviation, therefore

we recommend one of them. When only considering the overall RMS, the Gaussian fit offers a slight advantage over the second order polynomial.

Of course, the maximum degree of accuracy depends on the functional form we offered for the coefficients. We chose rather vague guesses, due to the fact that we only simulated three distinct v .

6 Conclusion and Outlook

We introduced the framework of graph theory and the Mayer formalism. From there, we followed Wertheim’s reasoning to find more convenient expressions for the grand canonical EOS and the free energy in a system governed by highly directional attractive forces. By means of thermodynamic perturbation theory, we derived equations of state of first and second TPT order, where we fitted a parameter in the latter to simulation data. The results are three good approximations within the range of our data. We did not check how well the approximation is for different v values or how strongly we may depend on β and V . This may be a topic for further research.

Within our approximation, it would be interesting to closer observe the v dependency of the fitting function coefficients. With more data points, more elaborate guesses could be possible, maybe also for the fitting functions themselves. It could also be interesting to find expressions for λ' that are closer to Wertheim’s analytical value for λ in the case $\eta \rightarrow 0$. By that, a compromise between pure mathematical approximation and physical reasoning may be possible.

Acknowledgment

I want to thank Professor J. Groß for his support and Mr. Wasilios Zmpitas for his introduction to the topic and everything following that.

References

- [1] M. S. Wertheim, “Fluids with Highly Directional Attractive Forces. I. Statistical Thermodynamics,” *J Stat Phys*, vol. 35, pp. 19–34, Apr 1984.
- [2] M. S. Wertheim, “Fluids with Highly Directional Attractive Forces. II. Thermodynamic Perturbation Theory and Integral Equations,” *J Stat Phys*, vol. 35, pp. 35–47, Apr 1984.
- [3] M. S. Wertheim, “Fluids with Highly Directional Attractive Forces. III. Multiple Attraction Sites,” *J Stat Phys*, vol. 42, pp. 459–476, Feb 1986.
- [4] M. S. Wertheim, “Fluids with Highly Directional Attractive Forces. IV. Equilibrium Polymerization,” *J Stat Phys*, vol. 42, pp. 477–492, Feb 1986.
- [5] M. S. Wertheim, “Thermodynamic Perturbation Theory of Polymerization,” *J. Chem. Phys.*, vol. 87, no. 12, p. 7323, 1987.
- [6] G. Stell, H. L. Frisch, and J. L. Lebowitz, “The Equilibrium Theory of Classical Fluids,” *Benjamin, New York*, 1964.
- [7] T. Morita and K. Hiroike, “A New Approach to the Theory of Classical Fluids. III,” *Progress of Theoretical Physics*, vol. 25, pp. 537–578, Apr 1961.
- [8] N. F. Carnahan and K. E. Starling, “Equation of State for Nonattracting Rigid Spheres,” *J. Chem. Phys.*, vol. 51, no. 2, p. 635, 1969.
- [9] Y. Uehara, Y.-T. Lee, T. Ree, and F. H. Ree, “Triplet distribution functions for hard spheres and hard disks,” *J. Chem. Phys.*, vol. 70, no. 4, p. 1884, 1979.
- [10] T. van Westen, B. Oyarzún, T. J. Vlucht, and J. Gross, “An equation of state for the isotropic phase of linear, partially flexible and fully flexible tangent hard-sphere chain fluids,” *Molecular Physics*, vol. 112, pp. 919–928, Jul 2013.
- [11] MATLAB, *version 8.0.0.783 (R2012b)*. Natick, Massachusetts: The MathWorks Inc., 2012.