

# The inverse Henderson problem in the thermodynamic limit

DISSERTATION

zur Erlangung des Grades  
*Doktor der Naturwissenschaften*

am Fachbereich Physik, Mathematik und Informatik  
der Johannes Gutenberg-Universität  
in Mainz

vorgelegt von

**Fabio Frommer**  
geboren in Münster

Mainz, 30. Januar 2025



# Summary

Particle systems interacting via some Hamiltonian are often described via so-called Gibbs measures. However, in practice it is not possible to measure or calculate these interactions directly from experiments or simulation data. What, however, is possible is to calculate certain statistics of the number and relative positions of particles from spectroscopy or simulations. This of course gives rise to mathematical questions: „Given these point statistics, how do we recover the interaction Hamiltonian and is it unique?“ In an often cited paper Henderson proved, for the case of a finite volume box at fixed density, i.e. a canonical ensemble, that any such solution is necessarily unique and in the '80s Chayes and Chayes showed that under broad conditions a solution exists. However their results do not immediately extend to the thermodynamic limit, i.e. the infinite volume case. The goal of this work is to extend known results for the finite volume to the thermodynamic limit.

The thesis is structured in the following way. We start with a brief introduction to the setting we are working in and some historic background for the problem. In the second chapter we give a rigorous introduction to the mathematical background, i.e. the theory of point processes and in particular Gibbs point processes. Chapters 3 to 5 consist of the main results of this work, namely three peer-reviewed articles.

In Chapter 3 and 4 we discuss the inverse Henderson problem in the thermodynamic limit. We start by discussing Henderson's theorem and proving the thermodynamic limit version of it. Following, we use the Gibbs variational principle to rewrite the inverse Henderson problem as a minimization problem of the specific relative entropy in the thermodynamic limit. Using our version of Henderson's theorem established before, we then show that this functional is strictly convex, and thus that any minimizer has to be unique. We furthermore show that for any given pair potential and any particle density below the close-packing density, there always exists a chemical potential and at least one Gibbs measure for these parameters realizing this density. We can then reformulate the minimization problem in the canonical setting, i.e. work at fixed density. Lastly, on an appropriate space of perturbations, we calculate the derivatives of the relative entropy functional and show the connections to the well-known inverse Monte-Carlo scheme when using Newton's method for the minimization of the relative entropy functional.

In Chapter 5 we take a different approach: Assuming we are given all correlation functions of some Gibbs measure, we construct an expansion for the chemical potential associated to this measure. We show that the conditions for convergence are indeed very mild, as we can show the assumptions are satisfied in a variety of cases. In particular the results also hold for some models with many-body interactions.



# Zusammenfassung

Teilchensysteme, die über einen Hamiltonian interagieren, werden oft durch sogenannte Gibbs-Maße beschrieben. In der Praxis ist es jedoch nicht möglich, diese Wechselwirkungen direkt anhand von Experimenten- oder Simulationsdaten zu berechnen. Was jedoch möglich ist, ist die Berechnung einiger statistischer Größen über die Anzahl und relativen Positionen der Teilchen mit Hilfe von Spektroskopie oder Simulationen. Daraus ergeben sich natürlich die mathematischen Fragen: „Wie bestimme ich aus gegebenen Teilchen-Statistiken den zugrunde liegenden Hamiltonian und ist dieser eindeutig?“ In einer häufig zitierten Arbeit diskutierte Henderson dieses Problem im Fall eines endlichen Volumens mit konstanter Dichte, also eines kanonischen Ensembles, und zeigte, dass eine jede solche Lösung notwendigerweise eindeutig ist und in den 80ern zeigten Chayes und Chayes die Existenz einer Lösung unter relativ allgemeinen Voraussetzungen. Allerdings lassen sich diese Resultate nicht direkt auf den thermodynamischen Limes, also den Fall eines unendlichen Volumens, verallgemeinern. Das Ziel dieser Arbeit ist es, die für den Fall des endlichen Volumens bekannten Resultate auf den thermodynamischen Limes zu verallgemeinern.

Die Thesis ist wie folgt strukturiert. Wir beginnen mit einer kurzen Einführung in das Setting und historische Hintergründe. Im zweiten Kapitel führen wir die mathematischen Hintergründe ein, d.h. die Theorie der Punktprozesse und insbesondere Gibbssche Punktprozesse. Kapitel 3 bis 5 bestehen aus den Hauptresultaten dieser Arbeit, drei Peer Reviewten Artikeln. In Kapiteln 3 und 4 betrachten wir das inverse Henderson Problem im thermodynamischen Limes.

Zuerst führen wir das Henderson Theorem ein und beweisen es für den thermodynamischen Limes. Anschließend nutzen wir das Gibbssche Variationsprinzip um das inverse Henderson Problem in ein Minimierungsproblem für die spezifische relative Entropie im thermodynamischen Limes umzuformulieren. Unter Verwendung unserer Version des Henderson Theorems zeigen wir, dass dieses Funktional strikt konvex ist, und dementsprechend jeder Minimierer eindeutig sein muss. Weiterhin zeigen wir, dass für jede (Teilchen-)dichte unterhalb der dichtesten Kugelpackung immer ein chemisches Potential und mindestens ein Gibbs-Maß zu diesen Parametern existiert, welches diese Dichte realisiert. Damit können wir das Minimierungsproblem für das kanonische Setting neu formulieren, d.h. mit einer festen Dichte arbeiten. Zuletzt berechnen wir, auf einem geeigneten Störungsraum, die Ableitungen des relativen Entropie Funktionals und zeigen die Verbindungen zum bekannten inversen Monte-Carlo iterations Schema, wenn man die Newton Methode für die Minimierung des relativen Entropie Funktionals verwendet.

In Kapitel 5 verfolgen wir einen anderen Ansatz: Unter der Annahme alle Korrelationsfunktionen eines Gibbs-Maßes seien gegeben, konstruieren wir eine Expansion für das chemische Potential des zu Grunde liegenden Maßes. Wir zeigen weiterhin, dass unsere Bedingungen für die Konvergenz sehr schwach sind, da unsere Voraussetzungen in einer Vielzahl von Fällen erfüllt sind. Insbesondere gelten die Ergebnisse auch für einige Modelle mit Mehrkörper-Wechselwirkungen.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Mathematical background</b>	<b>5</b>
2.1	The configuration space and point processes . . . . .	5
2.2	Gibbs measures . . . . .	10
<b>3</b>	<b>A note on the uniqueness result for the Henderson problem</b>	<b>21</b>
3.1	Overview . . . . .	21
3.2	Reprint . . . . .	22
<b>4</b>	<b>A variational framework for the inverse Henderson problem of statistical mechanics</b>	<b>33</b>
4.1	Overview . . . . .	33
4.2	Reprint . . . . .	34
<b>5</b>	<b>An Inverse Cluster Expansion for the Chemical Potential</b>	<b>69</b>
5.1	Overview . . . . .	69
5.2	Reprint . . . . .	70
<b>6</b>	<b>Conclusion and Outlook</b>	<b>87</b>
	<b>Bibliography</b>	<b>92</b>
	<b>Acknowledgements</b>	<b>93</b>



# 1

## Introduction

In classical physics, liquid or gaseous matter at a given temperature is often characterized via the potential energy of the system. This means, the force acting between two (or more) particles is given by the derivative of some Hamiltonian. While one could theoretically simulate systems with high precision by numerically solving the underlying mechanical equations, in practice this is severely time consuming. Thus this approach can only be feasibly used for small systems of up to a few thousand particles and only very small time and length scales. This becomes an even bigger problem as soon as the structure of the system becomes more complicated, i.e. when looking at e.g. molecules or polymers, or when the interaction between the particles involves more than just a simple pair potential, e.g. angle-forces between triplets of particles.

Therefore, to simulate larger or more complicated systems, a *coarse-graining* approach is commonly used, see Peter and Kremer [50] or Noid [48], where in order to reduce computational effort the complicated system is simplified while keeping the relevant properties as close to the original system as possible, cf. Praprotnik, Delle Site, and Kremer [51]. In this work we specifically investigate the case when these structures and higher-order interactions are replaced by point- or bead-like particles interacting with a pair interaction. This mapping from the atomic scale to the „coarse“ beads needs to be designed, of course, in a thermodynamically consistent way.

There are two main approaches: An *ab initio* bottom-up approach where different methods are used to calculate the averaged forces resulting from the mapping to the coarse beads.

The other approach is called top-down coarse-graining. In this case one looks at structural properties of the atomic scale system and asks which interactions between the coarse-grained beads will reproduce them. As the forces at these length scales cannot be measured in a real molecular system, this second ansatz is the inverse problem, cf. Hanke [26], we discuss in detail in this work.

As mentioned, we work with structural information such as the (particle) density or relative positions of particles. These can either be measured via e.g. spectroscopy or evaluated from a simulation of the complicated system with a small number of particles. For simplicity, we will make the assumption that we only have one type of particles and that the interaction between them is translation invariant.

The available data are the so-called *n-point correlation functions*. If the interaction only depends on the distance of the particles, i.e. it is *isotropic*, the quantity of particular interest is the *radial distribution function*  $g$  which is the two-point correlation function normalized with respect to the density  $\rho$ . The radial distribution function is also an isotropic function, and for some  $r > 0$  the value  $g(r)$  gives the likelihood of finding at least one other particle at distance  $r$  from any given particle.

For a canonical or grand-canonical ensemble of particles in a finite volume box that only interact via some isotropic pair potential, this inverse problem becomes: „Given the radial distribution function at fixed temperature and density, determine the corresponding pair potential.“

The first big step in the analysis of this inverse problem was done by Henderson in the '70s. He proved in [31], for a canonical ensemble, that the radial distribution function uniquely determines the pair interaction. This work was widely recognized and hence, we refer to the inverse problem at hand as the *inverse Henderson problem*.

The goal of this thesis is to further analyze the inverse Henderson problem in the *thermodynamic limit*, i.e. looking at the limit of a canonical (or grand-canonical) ensemble when growing the box to eventually encapsulate the whole space. In particular, we want to investigate which theoretical arguments from the finite volume case can be extended to the thermodynamic limit and where mathematical difficulties arise. The technique Henderson used goes back to ideas from Hohenberg and Kohn [32] and Mermin [43], in particular he employs a *Gibbs variational principle* [22] which states, roughly speaking, that a system minimizes a certain *relative entropy functional* if and only if the distribution of the particles is given by a Gibbs measure associated to the given thermodynamic parameters. This ansatz has also been used for a similar problem concerning *external potentials* by Chayes, Chayes and Lieb [4, 6], which is a similar but somewhat more easily tractable inverse problem. Later, Shell argued in [58] that this relative entropy approach is fundamental to coarse-graining. The relative entropy Shell used,

$$S_{\text{rel}} = \int \log \left( \frac{dP_*}{dP} \right) dP_*,$$

also known as *Kullback–Leibler divergence*, is a well-known tool in statistics and information theory to measure the difference between two probability distributions  $P$  and  $P_*$ . Here  $P_*$  is the observed distribution, i.e. the distribution of the system we want to coarse-grain, and  $P$  is the ansatz Gibbs distribution with a pair interaction for the coarse-grained particles. In a work building on Shell's argument Murtola, Karttunen and Vattulainen showed in [44] that this relative entropy functional is convex, and that when minimizing  $S_{\text{rel}}$  using Newton's method, the corresponding iteration coincides with the inverse Monte Carlo method (IMC) developed by Lyubartsev and Laaksonen [40].

However, when turning to the thermodynamic limit one needs to be a bit more careful. One has to divide by the volume of the box in the limiting process, and thus the strict inequalities from the finite volume case may not be preserved.

Chapters 3 and 4 are devoted to the analysis of these problems in the thermodynamic limit. We make use of the fact that the re-scaled relative entropy converges to the specific relative entropy. It can be shown that this specific relative entropy satisfies an infinite volume version of the Gibbs variational principle. For continuum particle systems, this was first studied for hard-core interactions in [16] by Gallavotti and Miracle-Sole, and since then this result has been extended to more general interactions, most notably to the case of superstable pair interactions in a series of works by Georgii and Zessin [21, 18, 19]; but also for general hard-core systems by Nguyen and Zessin [46] and general finite-range interactions, see Dereudre [7]. Except for Chapter 5 of this work, we will however restrict ourselves to the case of translation invariant pair interactions.

The rest of the thesis is structured as follows:

In Chapter 2 we give an overview of the mathematical tools used in this work. We review some of the general theory of Point processes, the concept of superstable interactions and finally infinite volume Gibbs measures and their properties.

In Chapter 3 we look at Henderson's Theorem, i.e. we show the mapping of pair interactions to the radial distribution functions is injective. Using the aforementioned Gibbs variational principle of Georgii we can then prove two versions of his theorem. For a grand-canonical perspective, i.e. if we are given the density and pair correlation functions of a Gibbs measure, we show that there can be at most one chemical potential and pair interaction to which they belong. In the second case (more in the spirit of Henderson's version), i.e. when we are given the chemical potential and pair correlation function, we show uniqueness of the pair interaction.

In Chapter 4 we take a closer look at the specific relative entropy. We investigate its strict convexity and

differentiability with respect to the chemical potential and pair interaction, when the target distribution  $P_*$  is fixed. For that we introduce a certain Banach space of admissible perturbations. The idea for this goes back to a series of papers by Hanke [27, 28, 29] where the differentiability of the correlation functions was investigated.

Another approach to calculate interaction potentials from correlation functions stems from so-called *cluster expansions*. This technique comprises of writing certain thermodynamic functions as sums over special types of graphs and is the approach we investigate in Chapter 5. In a review paper from the '60s, Stell [59] collected a multitude of these expansions for different functions. In particular, it includes a formula, developed by Nettleton and Green in [45], to calculate the Hamiltonian of a system from the correlation functions. However, this formula, like most of these cluster expansions, remains (to this day) only a formal equation, as showing convergence of the appearing series, outside of some special cases, has so far not been possible. We use a slightly different approach, namely rewriting the appearing functions as a recursion. In the next step, we use Bell-polynomials to find an upper bound for the expansion. Our method needs only weak decay assumptions on the Hamiltonian and thus we can rigorously prove the Nettleton and Green formula for a wide range of systems, including pair interactions, non-negative many-body interactions of finite range, and the interactions of the *Kirkwood-closure* point process.



## 2

# Mathematical background

In this chapter we will give an overview into the theoretical background needed to understand the results presented in Chapters 3-5. We start with a general introduction to some notions of point processes and then more concretely introduce the specific model of Gibbs measures we will be working with.

## 2.1 The configuration space and point processes

We denote by  $\mathcal{B}_b(\mathbb{R}^d)$  the bounded Borel-subsets of  $\mathbb{R}^d$ . To describe the particles in space we introduce the so-called *configuration space* by

$$\Gamma = \{ \gamma \subset \mathbb{R}^d \mid \Delta \in \mathcal{B}_b(\mathbb{R}^d) \Rightarrow N_\Delta(\gamma) < \infty \},$$

where  $N_\Delta: \Gamma \rightarrow \mathbb{N}_0 \cup \{\infty\}$ ,  $\gamma \mapsto N_\Delta(\gamma) = \#(\gamma \cap \Delta)$ . Sometimes we will also work with the space of finite configurations  $\Gamma_0 := \{ \gamma \in \Gamma \mid \#\gamma < \infty \}$  and the space of configurations constrained to a certain subset of  $\mathbb{R}^d$ , namely for  $\Lambda \subset \mathbb{R}^d$  we define  $\Gamma_\Lambda := \{ \gamma \in \Gamma \mid \gamma \subset \Lambda \}$ . We furthermore introduce the  $\sigma$ -algebra  $\mathcal{F} := \sigma(N_\Delta \mid \Delta \in \mathcal{B}_b(\mathbb{R}^d))$  that is generated by the counting variables.

**Definition 2.1.1** *Any random variable  $\eta$  on some probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  with values in  $(\Gamma, \mathcal{F})$  is called a point process. The distribution  $\mathbb{P} = \mathbb{P} \circ \eta^{-1}$  of  $\eta$  is then a probability measure on  $(\Gamma, \mathcal{F})$ . By some abuse of language we will also call any probability measure on  $(\Gamma, \mathcal{F})$  a point process and in particular the distribution  $\mathbb{P}$  of  $\eta$ .*

If for every  $x \in \mathbb{R}^d$  the distribution of  $\eta$  is the same as for  $\eta - x = \{y \mid y + x \in \eta\}$ , we say  $\eta$  (or  $\mathbb{P}$ ) is a *translation invariant* or *stationary* point process.

**Example 2.1.2** *One of the simplest examples of a point process is the so-called Poisson point process, characterized in the following way:  $\pi$  is a Poisson process with intensity  $\rho > 0$  if*

(i) *For every  $\Lambda \in \mathcal{B}_b(\mathbb{R}^d)$  the distribution of  $N_\Lambda$  is Poisson with parameter  $\rho \mathfrak{L}(\Lambda)$  ( $\mathfrak{L}(\Lambda)$  denotes the volume of the set  $\Lambda$  with respect to the Lebesgue measure), that is to say for every  $k \in \mathbb{N}_0$*

$$\pi(N_\Lambda = k) = \frac{(\rho \mathfrak{L}(\Lambda))^k}{k!} e^{-\rho \mathfrak{L}(\Lambda)}. \quad (2.1)$$

(ii) *For every  $m \in \mathbb{N}$  and all  $\Lambda_1, \dots, \Lambda_m \in \mathcal{B}_b(\mathbb{R}^d)$  such that  $\Lambda_i \cap \Lambda_j = \emptyset$  for all  $i \neq j$ , the random variables  $N_{\Lambda_1}, \dots, N_{\Lambda_m}$  are independent.*

**Remark 2.1.3** *The Poisson process is a probabilistic model for the ideal gas, i.e. point-like particles that do not interact with each other.*

### 2.1.1 Janossy densities and correlation functions

In general, when looking at a non-trivial stationary point process in the whole space, the probability to encounter only finitely many points is zero, see e.g. [38]. In particular, densities of the distribution of the point process with respect to the Lebesgue measure do not exist. However, when looking at a bounded box  $\Lambda$  by definition of the configuration space the number of points will be finite.

**Definition 2.1.4** We say a function  $F: (\mathbb{R}^d)^n \rightarrow \mathbb{R}$  is symmetric, if for any permutation  $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$  we have  $F(x_1, \dots, x_n) = F(x_{\sigma(1)}, \dots, x_{\sigma(n)})$ .

By abuse of notation and the fact that  $\{\#\{x_1, \dots, x_n\} < n\} \subset (\mathbb{R}^d)^n$  is a Lebesgue null set we take  $\mathbf{x}_n = (x_1, \dots, x_n)$  or  $\mathbf{x}_n = \{x_1, \dots, x_n\}$  in the following, depending on context.

**Definition 2.1.5** Let  $\mathbb{P}$  be a probability measure on  $(\Gamma, \mathcal{F})$  and  $(j_\Lambda^{(n)})_{n \geq 0, \Lambda \in \mathcal{B}_b(\mathbb{R}^d)}$ ,  $j_\Lambda^{(n)}: \Lambda^n \rightarrow [0, \infty)$  be a family of symmetric functions. They are called Janossy densities of  $\mathbb{P}$  if for every non-negative function  $F: \Gamma \rightarrow [0, \infty)$  we have

$$\mathbb{E}[F(\eta_\Lambda)] = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\Lambda^n} F(\mathbf{x}_n) j_\Lambda^{(n)}(\mathbf{x}_n) d\mathbf{x}_n. \quad (2.2)$$

Here we write  $\eta_\Lambda = \eta \cap \Lambda$  and  $\int_{\Lambda^0} F(\mathbf{x}_0) j_\Lambda^{(0)}(\mathbf{x}_0) d\mathbf{x}_0 = F(\emptyset) j_\Lambda^{(0)}$ . Note that  $j_\Lambda^{(0)} = \mathbb{P}(N_\Lambda(\eta) = 0)$ .

The Janossy densities are also known as *system of density distributions* as Ruelle introduced them in [56]. Closely related to the Janossy densities are the *correlation functions*.

**Definition 2.1.6** Let  $\mathbb{P}$  be a probability measure on  $(\Gamma, \mathcal{F})$  and  $(\rho^{(n)})_{n \geq 1}$  be a family of symmetric non-negative functions. These are called correlation functions of  $\mathbb{P}$  if for every  $F: (\mathbb{R}^d)^n \rightarrow [0, \infty)$  we have

$$\mathbb{E} \left[ \sum_{\substack{x_1, \dots, x_n \in \eta, \\ x_i \neq x_j}} F(\mathbf{x}_n) \right] = \int_{(\mathbb{R}^d)^n} F(\mathbf{x}_n) \rho^{(n)}(\mathbf{x}_n) d\mathbf{x}_n. \quad (2.3)$$

**Remark 2.1.7** If  $\mathbb{P}$  is a stationary point process with correlation functions  $(\rho^{(n)})_{n \geq 1}$ , then these are translation invariant as well, i.e.

$$\rho^{(n)}(x_1 + a, \dots, x_n + a) = \rho^{(n)}(x_1, \dots, x_n) \quad \text{for all } a \in \mathbb{R}^d.$$

In particular, we have

$$\rho^{(n)}(x_1, \dots, x_n) = \rho^{(n)}(0, x_2 - x_1, \dots, x_n - x_1) \quad (2.4)$$

and thus, the one-particle correlation function is constant, i.e.  $\rho^{(1)} \equiv \rho$  and is called the density or (counting) intensity of a point process. In general we can factor the density out of the correlation functions and write  $\rho^{(n)} = \rho^n g^{(n)}$  for appropriate functions  $g^{(n)}$ . The only one of interest in this thesis being the so-called radial distribution function  $g^{(2)} = g = \rho^{(2)}/\rho^2$ , which by (2.4) satisfies  $g(x_1, x_2) = g(x_1 - x_2, 0)$ . By abuse of notation we therefore write  $g(x_1 - x_2)$  below.

**Example 2.1.8** If  $F = \prod_{i=1}^n \mathbb{1}_\Lambda$  for some  $\Lambda \in \mathcal{B}_b(\mathbb{R}^d)$ , then

$$\mathbb{E} \left[ \sum_{\substack{x_1, \dots, x_n \in \eta, \\ x_i \neq x_j}} \prod_{i=1}^n \mathbb{1}_\Lambda(x_i) \right] = \mathbb{E}[N_\Lambda(\eta)(N_\Lambda(\eta) - 1) \cdots (N_\Lambda(\eta) - (n - 1))] = \int_{\Lambda^n} \rho^{(n)}(\mathbf{x}_n) d\mathbf{x}_n \quad (2.5)$$

and in particular  $\mathbb{E}[N_\Lambda(\eta)] = \int_\Lambda \rho^{(1)}(x) dx$  which for a stationary point process reduces further to  $\mathbb{E}[N_\Lambda(\eta)] = \mathfrak{K}(\Lambda)\rho$ .

**Remark 2.1.9** *The property (2.5) in particular is what makes the correlation functions a „measurable“ quantity, as they can be estimated from realizations of the point process by only having to count points in subsets.*

Both the correlation functions and the Janossy densities, if they exist, determine the associated probability measure uniquely due to the definition of  $\mathcal{F}$ , cf. [33]. Furthermore, using (2.2) and (2.3), we get the following relation.

**Proposition 2.1.10 ([33])** *Let  $\mathbb{P}$  be a point process admitting Janossy densities, then the correlation functions of  $\mathbb{P}$  exist and there holds for every  $n \in \mathbb{N}$  and  $\Lambda \in \mathcal{B}_b(\mathbb{R}^d)$*

$$\rho^{(n)}(\mathbf{x}_n) = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\Lambda^n} j_{\Lambda}^{(n+k)}(\mathbf{x}_n, \mathbf{y}_k) d\mathbf{y}_k, \quad \mathbf{x}_n \in \Lambda^n. \quad (2.6)$$

Here we write  $j_{\Lambda}^{(n+k)}(\mathbf{x}_n, \mathbf{y}_k) = j_{\Lambda}^{(n+k)}(x_1, \dots, x_n, y_1, \dots, y_k)$  for brevity.

**Remark 2.1.11** *Note that the correlation functions are defined on the whole space whereas the Janossy densities always are connected to some bounded set  $\Lambda$ . However, they contain the averaged contributions of the whole space by definition and thus the right-hand side above is independent of  $\Lambda$ , as long as  $\mathbf{x}_n \in \Lambda^n$ .*

**Example 2.1.12** *The correlation functions of the Poisson process with intensity  $\rho$  exist and are given by  $\rho^{(n)}(\mathbf{x}_n) \equiv \rho^n$ , see e.g. [33].*

In general calculating the correlation functions  $(\rho^{(n)})_{n \geq 1}$  for a given system is computationally very intensive for  $n > 2$ . Thus, in computational physics the so-called *Kirkwood-superposition* or *Kirkwood-closure* has been used, i.e.

$$\rho^{(n)}(\mathbf{x}_n) \approx \rho^n \prod_{1 \leq i < j \leq n} g(x_i - x_j)$$

where  $\rho$  is the density and  $g$  the radial distribution function of some given point process. This approximation was first suggested by Kirkwood in [34], see also [30], and has been used to approximate in particular the 3-point correlation function. However, in general it is unclear whether these approximated correlation functions still belong to some point process.

**Definition 2.1.13** *Let  $\rho > 0$  and  $g: \mathbb{R}^d \rightarrow [0, \infty)$  be an even function, a point process  $\mathbb{K}_{\rho, g}$  with correlation functions  $(\rho^{(n)})_{n \geq 1}$  is the Kirkwood-closure process if there holds*

$$\rho^{(n)}(\mathbf{x}_n) = \rho^n \prod_{1 \leq i < j \leq n} g(x_i - x_j). \quad (2.7)$$

Existence of the Kirkwood-closure process was first proved by Ambartzumian and Sukiasian in [1] for the case  $g \leq 1$  and later generalized in [37] by Kuna, Lebowitz and Speer.

**Theorem 2.1.14 ([37])** *Let  $g: \mathbb{R}^d \rightarrow [0, \infty)$  be an even function on  $\mathbb{R}^d$  such that*

1.  $C(g) = \int_{\mathbb{R}^d} |g(x) - 1| dx < \infty$ ;
2. *There exists a  $1 \leq b < \infty$ , such that for all  $n \geq 1$*

$$\prod_{i=1}^n g(x_0 - x_i) \leq b$$

*whenever  $x_0, x_1, \dots, x_n$  satisfy  $\prod_{i < j} g(x_i - x_j) > 0$ .*

Then for every  $\rho$  satisfying

$$0 \leq \rho \leq (ebC(g))^{-1}$$

the Kirkwood-closure process  $K_{\rho,g}$  exists.

We use the Kirkwood-closure process in Chapter 4, i.e. [14], to show the ill-posedness of the inverse Henderson problem in the thermodynamic limit.

### 2.1.2 The Ruelle bound

**Definition 2.1.15** *The correlation functions  $(\rho^{(n)})_{n \geq 1}$  of a point process  $P$  satisfy a so-called Ruelle bound if there is a  $\xi > 0$  such that*

$$\rho^{(n)}(\mathbf{x}_n) \leq \xi^n. \quad (\mathcal{R}_\xi)$$

We also say  $P$  satisfies a Ruelle condition if  $P$  admits correlation functions that satisfy a Ruelle bound.

**Example 2.1.16** *The Poisson process trivially satisfies a Ruelle condition with  $\xi = \rho$  and the Kirkwood-closure process from [37] satisfies a Ruelle condition with  $\xi = \rho b^{1/2}$ .*

For point processes satisfying a Ruelle condition, we get an inverse formula to (2.6).

**Theorem 2.1.17 ([36])** *Let  $P$  be a point process that satisfies a Ruelle condition, then  $P$  admits Janossy densities and for any  $\Lambda \in \mathcal{B}_b(\mathbb{R}^d)$  there holds*

$$j_\Lambda^{(n)}(\mathbf{x}_n) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{\Lambda^k} \rho^{(n+k)}(\mathbf{x}_n, \mathbf{y}_k) d\mathbf{y}_k \quad (2.8)$$

where in the case  $n = 0$  we formally set  $\rho^{(0)} = 1$ .

The above formula is the ansatz we use in Chapter 5, i.e. [13], to find an expansion of the so-called chemical potential.

### 2.1.3 Truncated correlation functions and exponential representation

Lastly, we introduce the truncated version correlation functions and show how they are connected to an exponential representation for the Janossy densities.

**Definition 2.1.18** *Let  $(\rho^{(n)})_{n \geq 1}$  be the correlation functions of some point process  $P$ , the truncated correlation functions  $(\rho_T^{(n)})_{n \geq 1}$  are recursively defined by  $\rho_T^{(1)}(x) = \rho^{(1)}(x)$  and for  $n \geq 2$  as*

$$\rho_T^{(n)}(\mathbf{x}_n) = \rho^{(n)}(\mathbf{x}_n) - \sum_{k=2}^n \sum_{\pi \in \Pi_k(\mathbf{x}_n)} \prod_{i=1}^k \rho_T^{(\kappa_i)}(\pi_i) \quad (2.9)$$

where  $\Pi_k(\mathbf{x}_n)$  is the set of partitions  $\pi = (\pi_1, \dots, \pi_k)$  of  $\{x_1, \dots, x_n\}$  into  $k$  non-empty sets  $\pi_i$  and  $\kappa_i = \#\pi_i$ .

**Remark 2.1.19** *The truncated correlation functions are sometimes (depending on whether the work is more from the point process or mathematical physics approach) also called cluster functions, see [55], or Ursell functions, e.g. [59]. In [14] we have used the term cluster functions in place of truncated correlation functions and denote them by  $\omega^{(n)}$  as is done in [55] and [29]. The term Ursell functions has also been used for another object defined by a relation of the type (2.9), see e.g. [55]. Since there is no other naming convention for this object, we will reserve the term Ursell functions for the latter object (see (2.42)).*

**Remark 2.1.20** While the correlation functions of a point process are typically only bounded functions, i.e.  $\rho^{(n)} \in L^\infty((\mathbb{R}^d)^n)$ , but not integrable over the whole space, the truncated correlation functions are (in many examples we look at) integrable  $n - 1$  times in the sense that

$$\int_{(\mathbb{R}^d)^{n-1}} \left| \rho_T^{(n)}(0, \mathbf{x}_{n-1}) \right| d\mathbf{x}_{n-1} < \infty.$$

In fact, functions related to one another like the correlation functions, truncated correlation functions and the Janossy densities satisfy a type of exponential relation that appears quite naturally in statistical physics. The following Theorem is well-known, we use the version proved by Boluh and Rebenko.

**Theorem 2.1.21 ([3])** Let  $(F_n)_{n \geq 1}$  be a family of symmetric functions with  $F_n: (\mathbb{R}^d)^n \rightarrow \mathbb{R}$  such that there exist  $0 < c < 1/2$  and  $D > 0$  such that for every bounded  $\Lambda \subset \mathbb{R}^d$  there holds

$$\int_{\Lambda^n} |F_n(\mathbf{x}_n)| d\mathbf{x}_n \leq \mathfrak{K}(\Lambda) n! D c^n \quad (2.10)$$

for every  $n \in \mathbb{N}$ . Then the function  $\Phi: \Gamma_0 \rightarrow \mathbb{R}$  defined by  $\Phi(\emptyset) = 1$  and

$$\Phi(\eta) = \sum_{k=1}^{N(\eta)} \sum_{\pi \in \Pi_k(\eta)} \prod_{i=1}^k F_{\kappa_i}(\pi_i), \quad (2.11)$$

satisfies

$$\sum_{n=0}^{\infty} \frac{1}{n!} \int_{\Lambda^n} \Phi(\mathbf{x}_n) d\mathbf{x}_n = \exp \left( \sum_{n=1}^{\infty} \frac{1}{n!} \int_{\Lambda^n} F_n(\mathbf{x}_n) d\mathbf{x}_n \right). \quad (2.12)$$

**Remark 2.1.22** Note that for any function  $a \in L^\infty(\mathbb{R}^d)$  with  $\|a\|_\infty = a_0$  and any family  $(F_n)_{n \geq 1}$  satisfying the assumptions of Theorem 2.1.21 the family  $(F_n^a)_{n \geq 1}$  defined by

$$F_n^a(\mathbf{x}_n) = F_n(\mathbf{x}_n) \prod_{i=1}^n a(x_i)$$

also satisfies the assumptions of Theorem 2.1.21 when making the restriction  $0 < c < 1/(2a_0)$ .

Theorem 2.1.21 is the second tool we use in Chapter 5. We introduce the assumption:

**Assumption A** There are constants  $M > 0$  and  $C_\rho, D_\rho > 0$  such that for any  $\Lambda \subset \mathbb{R}^d$  and any  $\mathbf{x}_n \in \Lambda^n$

$$\int_{\Lambda^k} \left| \rho_T^{(n+k)}(\mathbf{x}_n, \mathbf{y}_k) \right| d\mathbf{y}_k \leq (n+k-1)! M C_\rho^{n-1} D_\rho^k. \quad (2.13)$$

**Remark 2.1.23** It is well-known that Assumption A is satisfied for Gibbs measures of superstable and regular pair interactions (introduced in the next section), see e.g. [55] and [2]. Furthermore, the truncated correlation functions of the Kirkwood-closure process also satisfy Assumption A, cf. [37].

**Example 2.1.24** For any point process satisfying Assumption A with  $D_\rho$  small enough we can use (2.8) and Theorem 2.1.21 to write

$$j_\Lambda^{(0)} = \exp \left( \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{\Lambda^k} \rho_T^{(k)}(\mathbf{y}_k) d\mathbf{y}_k \right). \quad (2.14)$$

## 2.2 Gibbs measures

An important type of point processes are so-called *Gibbs point processes* or *Gibbs measures* after Josiah Willard Gibbs who first formulated many of the ideas behind statistical mechanics, see e.g. [22]. The heuristic is the following: When describing a system of interacting particles in equilibrium via a probability measure, a specific configuration of particles should be more likely when it is energetically favorable. This energy is calculated via a Hamiltonian that depends on the specific interaction of the model. For most of this work we will restrict ourselves to the case where the Hamiltonian only consist of a pair interaction that depends on the relative positions of the points.

### 2.2.1 Assumptions on the interaction

To be able to rigorously formulate our results we need to restrict ourselves to interactions that behave in a somewhat nice way, so-called (*super-*)*stable* interactions. This concept covers a variety of interactions such as the Lennard-Jones interaction or particles with a hard-core. In general we consider a *Hamiltonian*  $H$  to be a measurable map  $H: \Gamma_0 \rightarrow \mathbb{R}^d \cup \{\infty\}$  such that

- $H(\emptyset) = 0$ ,
- If  $H(\eta) < \infty$ , then  $H(\eta \setminus \{x\}) < \infty$  for all  $x \in \eta$ .

These conditions are called *non-degeneracy* and *heredity*. For simplicity we will also assume that  $H$  does not contain any self-interaction, i.e.  $H(\{x\}) = 0$  for any  $x \in \mathbb{R}^d$ . We will treat this self-interaction separately and denote it by  $\mu$ , this is usually called the *chemical potential*.

**Definition 2.2.1** A *Hamiltonian* is called *stable* if there exists some  $B \geq 0$  such that

$$H(\eta) \geq -BN(\eta),$$

and  $H$  is called *superstable* if there also exists a constant  $A > 0$  such that for any bounded set  $\Lambda \subset \mathbb{R}^d$  and any configuration  $\eta \subset \Lambda$  we have

$$H(\eta) \geq \frac{A}{\lambda(\Lambda)} N(\eta)^2 - BN(\eta). \quad (2.15)$$

We say the Hamiltonian  $H$  consists only of pair interactions if there is some even function  $u: \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\}$  such that

$$H(\eta) = \frac{1}{2} \sum_{x \neq y \in \eta} u(x - y), \quad \eta \in \Gamma_0. \quad (2.16)$$

**Definition 2.2.2** The *interaction* or *interaction energy*  $W: \Gamma_0 \times \Gamma_0 \rightarrow \mathbb{R} \cup \{\infty\}$  between two configurations is defined as

$$W(\eta, \gamma) := \begin{cases} \sum_{x \in \eta, y \in \gamma} u(x - y), & \text{if } \sum_{x \in \eta, y \in \gamma} |u(x - y)| < \infty \\ \infty, & \text{otherwise.} \end{cases} \quad (2.17)$$

In particular, if both  $\eta, \gamma \in \Gamma_0$ , we have

$$H(\eta \cup \gamma) = H(\eta) + W(\eta, \gamma) + H(\gamma). \quad (2.18)$$

Lastly, we want to introduce the *Mayer-function*, which is related to the pair interaction and often times easier to work with.

**Definition 2.2.3** Let  $\beta > 0$  and  $u: \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\}$  be an even function that is bounded from below, then the Mayer-function  $f_\beta$  associated to  $u$  and  $\beta$  is defined by

$$f_\beta(x) = e^{-\beta u(x)} - 1. \quad (2.19)$$

If  $f_\beta \in L^1(\mathbb{R}^d)$  for every  $\beta > 0$ , we say that  $u$  is regular and we denote  $C(u, \beta) := \int_{\mathbb{R}^d} |f_\beta(x)| dx$ . When working at fixed  $\beta$  we will often write  $f$  in place of  $f_\beta$ .

There have been a multitude of works to find sufficient conditions on  $u$  so that the Hamiltonian associated to  $u$  by (2.16) is stable (resp. superstable) and regular, see e.g. [12] or [10] and [56].

**Definition 2.2.4** We denote by  $\mathcal{U}$  the set of even functions  $u: \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\}$  such that there exists  $r_0 > 0$  and decreasing positive functions  $\varphi: (0, r_0) \rightarrow \mathbb{R}_0^+$  and  $\psi: [0, \infty) \rightarrow \mathbb{R}^+$  with

$$\int_0^{r_0} r^{d-1} \varphi(r) dr = \infty \quad \text{and} \quad \int_0^\infty r^{d-1} \psi(r) dr < \infty, \quad (2.20)$$

such that

$$\begin{aligned} u(x) &\geq \varphi(|x|) && \text{for } 0 < |x| < r_0, \\ |u(x)| &\leq \psi(|x|) && \text{for } |x| \geq r_0, \end{aligned} \quad (2.21)$$

holds true almost everywhere; for our convenience we take  $\psi$  to be a bounded and decreasing function defined for all  $r \geq 0$ . We denote by  $\mathcal{U}_0$  the subset of the above pair potentials, which also belong to  $L_{\text{loc}}^\infty(\mathbb{R}^d \setminus \{0\})$ , and thus  $\mathcal{U}$  is the union of  $\mathcal{U}_0$  with the hard-core potentials  $\mathcal{U}_{\text{hc}}$ , which satisfy the condition above with  $\varphi$  replaced by  $\infty$ . In this case  $r_0$  is called the hard-core radius. If  $u$  satisfies (2.20) and (2.21) with a minorant and majorant of the forms

$$\varphi(r) = cr^{-\alpha} \quad \text{and} \quad \psi(r) = C(1 + r^2)^{-\alpha/2} \quad (2.22)$$

for some  $C > c > 0$  and  $\alpha > d$  we say that  $u$  is of Lennard-Jones type. We denote the subset of Lennard-Jones type potentials by  $\mathcal{U}_{\text{LJ}}$ .

It is quite easy to see, cf. Chapter 4, that the sets  $\mathcal{U}$ ,  $\mathcal{U}_0$ ,  $\mathcal{U}_{\text{hc}}$  and  $\mathcal{U}_{\text{LJ}}$  are convex. In Chapter 4 we will investigate differentiability properties of certain thermodynamic functions. As the set  $\mathcal{U}$  lacks a natural topology and the conditions (2.20) and (2.21) cannot hold under a change of sign, we instead introduce a corresponding Banach space  $\mathcal{V}_{u_*}$  of perturbations for a given potential  $u_* \in \mathcal{U}$ , following [27].  $\mathcal{V}_{u_*}$  is the space consisting of all even measurable functions  $v: \mathbb{R}^d \rightarrow \mathbb{R}$ , for which the associated norm

$$\|v\|_{\mathcal{V}_{u_*}} = \sup_{x \in \mathbb{R}^d} \frac{|v(x)|}{\psi_*(|x|)} \quad (2.23)$$

is finite. Here,  $\psi_*$  is the majorant  $\psi$  of (2.20) associated with  $u_*$ . Note that the norm (2.23) is somewhat stronger than the one employed in [27] because it is more restrictive in the core region  $0 < r < r_0$ , and hence, the resulting space of perturbations is smaller. However, it can easily be seen that for any  $v \in \mathcal{V}_{u_*}$  the sum  $u_* + v$  belongs to  $\mathcal{U}$ , see Chapter 4.

**Proposition 2.2.5 ([55, 56])** For any  $u \in \mathcal{U}$  the Hamiltonian associated to  $u$  by (2.16) is superstable (and thus in particular stable) and regular.

**Example 2.2.6** While all of the pair potentials in  $\mathcal{U}$  are "physical", there are some superstable potentials which are not included in the set  $\mathcal{U}$ . However, these potentials do not arise from any natural physical

model. We give a simple example. For some  $r_0 > 0$  let  $u_1 = \infty \mathbf{1}_{B_{r_0}(0)}$ , i.e. the pure hard-core potential. Then  $u$  induces a superstable Hamiltonian (2.16) with  $A = B/\rho_{\text{cp}}$  and any  $B > 0$  where

$$\rho_{\text{cp}} := \sup \left\{ \frac{N_\Lambda(\eta)}{\mathfrak{A}(\Lambda)} \mid \Lambda \in \mathcal{B}_b(\mathbb{R}^d), \eta \in \Gamma \text{ with } |x - y| \geq r_0 \text{ for all } x \neq y \in \eta \right\}$$

is the close-packing density.  $H(\eta) = \infty$  for any  $\eta \subset \Lambda$  with  $N_\Lambda(\eta) > \rho_{\text{cp}}\mathfrak{A}(\Lambda)$  and the lower bound (2.15) is negative in any other case. We thus make the following ansatz:

We define an even function  $u: \mathbb{R}^d \rightarrow \mathbb{R}_0^+$  by

$$u := u_1 + u_2 \tag{2.24}$$

where  $u_2 = \sum_{n=1}^{\infty} \mathbf{1}_{A_n}$  with  $A_n \subset B_{(n+1)r_0}(0) \setminus B_{nr_0}(0)$  is a sequence of sets such that  $\sum_{n=1}^{\infty} \mathfrak{A}(A_n) < \infty$ , then it is obvious that  $u \notin \mathcal{U}$ , but  $u$  induces a superstable Hamiltonian as sum of a superstable and a stable potential ( $u_2$  is stable as any non-negative potential is stable), cf. [56]. Furthermore

$$\begin{aligned} \int_{\mathbb{R}^d} |e^{-\beta u(x)} - 1| dx &= \int_{\mathbb{R}^d} |e^{-\beta(u_1+u_2)(x)} - e^{-\beta u_2(x)} + e^{-\beta u_2(x)} - 1| dx \\ &\leq \int_{\mathbb{R}^d} e^{-\beta u_2(x)} |e^{-\beta u_1(x)} - 1| dx + \int_{\mathbb{R}^d} |e^{-\beta u_2(x)} - 1| dx < \infty. \end{aligned}$$

The first term is finite since  $e^{-\beta u_2(x)} \leq 1$  and  $u_1$  is regular, and the second is finite as  $\int_{\mathbb{R}^d} |e^{-\beta u_2(x)} - 1| dx = (1 - e^{-\beta}) \sum_{n=1}^{\infty} \mathfrak{A}(A_n) < \infty$ , thus  $u$  is also regular.

## 2.2.2 The grand-canonical ensemble

In light of the previous section's discussion we can now define Gibbs measures in finite volume. Usually Gibbs measures are written to be dependent on three parameters, in the grand-canonical ensemble (which we mostly work with) these are the *chemical potential*  $\mu$ , the pair potential  $u$  and the *inverse temperature*  $\beta > 0$ .

**Definition 2.2.7** For  $\beta > 0$ ,  $\mu \in \mathbb{R}$ ,  $u \in \mathcal{U}$  and some  $\Lambda \in \mathcal{B}_b(\mathbb{R}^d)$ , the grand-canonical partition function is given by

$$\Xi_\Lambda(\beta, \mu, u) = \sum_{n=0}^{\infty} \frac{e^{\beta\mu n}}{n!} \int_{\Lambda^n} e^{-\beta H(\mathbf{x}_n)} d\mathbf{x}_n \tag{2.25}$$

where the summand for  $n = 0$  is taken to be equal to one. The grand-canonical  $(\beta, \mu, u)$ -Gibbs measure in  $\Lambda$  is defined as the probability measure  $\mathbb{G}_\Lambda$  on  $\Gamma_\Lambda$  with

$$\int_{\Gamma_\Lambda} F(\gamma) d\mathbb{G}_\Lambda(\gamma) = \frac{1}{\Xi_\Lambda} \sum_{n=0}^{\infty} \frac{e^{\beta\mu n}}{n!} \int_{\Lambda^n} F(\mathbf{x}_n) e^{-\beta H(\mathbf{x}_n)} d\mathbf{x}_n \tag{2.26}$$

for all  $F: \Gamma_\Lambda \rightarrow [0, \infty)$ .

**Remark 2.2.8** Another approach to Gibbs measures in a finite volume is the so-called canonical ensemble, instead of prescribing a chemical potential, the number of particles is fixed. For  $\beta > 0$ ,  $N \in \mathbb{N}$ ,  $u \in \mathcal{U}$  and some  $\Lambda \in \mathcal{B}_b(\mathbb{R}^d)$  the canonical ensemble  $\mathbb{C}_\Lambda$  is a probability measure on  $\Lambda^N$  with

$$\int_{\Lambda^N} F(\mathbf{x}_N) d\mathbb{C}_\Lambda(\mathbf{x}_N) = \frac{1}{Z_\Lambda(\beta, N, u)} \int_{\Lambda^N} F(\mathbf{x}_N) e^{-\beta H(\mathbf{x}_N)} d\mathbf{x}_N \tag{2.27}$$

for every  $F: \Lambda^N \rightarrow [0, \infty)$ . Here  $Z_\Lambda(\beta, N, u)$  is the canonical partition function defined by

$$Z_\Lambda(\beta, N, u) = \int_{\Lambda^N} e^{-\beta H(\mathbf{x}_N)} d\mathbf{x}_N.$$

The canonical ensemble is commonly used in computational physics or chemistry, cf. [30], for simulations due to the much simpler implementation compared to the grand-canonical ensemble.

Note that since the Hamiltonian  $H$  associated to  $u$  is stable, we have that

$$1 \leq \Xi_\Lambda \leq \exp\left(e^{\beta(\mu+B)} \mathfrak{K}(\Lambda)\right) < \infty$$

and thus the grand-canonical partition function is always well-defined. Furthermore, for any  $\Delta \subset \Lambda$  and  $F: \Gamma_\Delta \rightarrow [0, \infty)$  using (2.18) and Fubini's Theorem we can write

$$\int_{\Gamma_\Delta} F(\gamma) dG_\Delta(\gamma) = \frac{1}{\Xi_\Delta} \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \frac{e^{\beta(n+k)\mu}}{k!n!} \int_{\Delta^n} \int_{(\Lambda \setminus \Delta)^k} F(\mathbf{x}_n \cup \mathbf{y}_k) e^{-\beta H(\mathbf{x}_n) - \beta W(\mathbf{x}_n | \mathbf{y}_k) - \beta H(\mathbf{y}_k)} d\mathbf{y}_k d\mathbf{x}_n.$$

Using (2.26) we can reformulate the above equation as

$$\int_{\Gamma_\Delta} F(\gamma) dG_\Delta(\gamma) = \sum_{n=0}^{\infty} \frac{e^{\beta n \mu}}{n!} \int_{\Delta^n} \int_{\Gamma_{\Lambda \setminus \Delta}} F(\mathbf{x}_n \cup \gamma) e^{-\beta H(\mathbf{x}_n) - \beta W(\mathbf{x}_n | \gamma)} dG_\Delta(\gamma) d\mathbf{x}_n. \quad (2.28)$$

**Remark 2.2.9** From (2.28) it follows that the Janossy densities in  $\Delta \subset \Lambda$  of the grand-canonical Gibbs measure  $G_\Delta$  on  $\Lambda$  are given by

$$j_{\Delta, \Lambda}^{(n)}(\mathbf{x}_n) = \frac{1}{\Xi_\Delta} \sum_{k=0}^{\infty} \frac{e^{\beta(n+k)\mu}}{k!} \int_{(\Lambda \setminus \Delta)^k} e^{-\beta H(\mathbf{x}_n, \mathbf{y}_k)} d\mathbf{y}_k, \quad \mathbf{x}_n \in \Delta^n. \quad (2.29)$$

In particular  $j_{\Delta, \Lambda}^{(n)}(\mathbf{x}_n) = e^{\beta n \mu - \beta H(\mathbf{x}_n)} / \Xi_\Delta$ .

**Definition 2.2.10** The finite volume grand-canonical pressure is defined by

$$p_\Lambda(\mu, u) = \frac{1}{\beta \mathfrak{K}(\Lambda)} \log \Xi_\Delta(\beta, \mu, u). \quad (2.30)$$

**Proposition 2.2.11** ([20]) The finite volume pressure  $p_\Lambda$  is strictly convex in  $(\mu, u)$ .

**Definition 2.2.12** For any bounded set  $\Lambda \subset \mathbb{R}^d$  the entropy of a point process  $P$  in  $\Lambda$  is defined as

$$S_\Lambda(P) = \begin{cases} \sum_{n=0}^{\infty} \int_{\Lambda^n} j_\Lambda^{(n)}(\mathbf{x}_n) \log j_\Lambda^{(n)}(\mathbf{x}_n) d\mathbf{x}_n, & \text{if } P \text{ admits Janossy densities in } \Lambda, \\ \infty, & \text{else.} \end{cases} \quad (2.31)$$

If  $Q$  is another point process, then the relative entropy of  $P$  with respect to  $Q$  is given by

$$S_\Lambda(P | Q) = \begin{cases} \sum_{n=0}^{\infty} \int_{\Lambda^n} j_{P, \Lambda}^{(n)}(\mathbf{x}_n) \log \left( \frac{j_{P, \Lambda}^{(n)}(\mathbf{x}_n)}{j_{Q, \Lambda}^{(n)}(\mathbf{x}_n)} \right) d\mathbf{x}_n, & \text{if } P \text{ and } Q \text{ admit Janossy densities in } \Lambda, \\ \infty, & \text{else.} \end{cases} \quad (2.32)$$

**Remark 2.2.13** *It is well-known that  $S_\Lambda(\mathbb{P} \mid \mathbb{Q}) \geq 0$  with equality if and only if  $\mathbb{P} = \mathbb{Q}$ , see e.g. [20]. Depending on authors sometimes the entropy (resp. relative entropy) is defined with the opposite sign to this work.*

Using the relative entropy gives us another way of characterizing Gibbs measures. For given  $\Lambda \in \mathcal{B}_b(\mathbb{R}^d)$  we choose  $\mathbb{Q} = \mathbb{G}_\Lambda$  and we get, by (2.26), (2.29) and (2.30)

$$S_\Lambda(\mathbb{P} \mid \mathbb{G}_\Lambda) = S_\Lambda(\mathbb{P}) - \beta \mu \mathbb{E}_\mathbb{P} [N_\Lambda(\eta)] + \beta \mathbb{E}_\mathbb{P} \left[ \frac{1}{2} \sum_{x \neq y \in \eta_\Lambda} u(x - y) \right] + \beta p_\Lambda(\beta, \mu, u) \mathfrak{K}(\Lambda).$$

By Remark 2.2.13 we know  $S_\Lambda(\mathbb{P} \mid \mathbb{G}_\Lambda) \geq 0$  with again equality if and only if  $\mathbb{P} = \mathbb{G}_\Lambda$ . In this case it follows that

$$p_\Lambda(\beta, \mu, u) \mathfrak{K}(\Lambda) = -\frac{1}{\beta} S_\Lambda(\mathbb{P}) + \mu \mathbb{E}_\mathbb{P} [N_\Lambda(\eta)] - \mathbb{E}_\mathbb{P} \left[ \frac{1}{2} \sum_{x \neq y \in \eta_\Lambda} u(x - y) \right]. \quad (2.33)$$

The right-hand side of the above equation is known in statistical physics as the *grand potential*, see e.g. [30], which is known to equal  $p_\Lambda(\beta, \mu, u) \mathfrak{K}(\Lambda)$  for homogeneous systems. This already shows a problem when trying to define Gibbs measures in an infinite volume, the left-hand side above obviously diverges. The solution to this is to divide everything by the volume of the box  $\Lambda$  and then taking the limit, this is called the *thermodynamic limit*.

### 2.2.3 Gibbs measures in infinite volume

The previous section leaves open how to define Gibbs measures in an infinite volume. However, close inspection of (2.28) shows that both sides can be well-defined in an infinite volume when replacing  $\mathbb{G}_\Lambda$  with some probability measure  $\mathbb{P}$  on  $\Gamma$  instead of  $\Gamma_\Lambda$ . The resulting equation is the so-called *Ruelle-equation*, which was first introduced by Ruelle in [56], wherein it was called *system of equilibrium equations*.

**Definition 2.2.14** *We say a point process  $\mathbb{P}$  satisfies the Ruelle-equation for some chemical potential  $\mu \in \mathbb{R}$ , a pair interaction  $u \in \mathcal{U}$  and  $\beta > 0$  if for every non-negative function  $F: \Gamma \rightarrow [0, \infty]$  and every bounded set  $\Lambda \subset \mathbb{R}^d$ , there holds*

$$\int_\Gamma F(\eta) d\mathbb{P}(\eta) = \sum_{n=0}^{\infty} \frac{e^{n\beta\mu}}{n!} \int_{\Lambda^n} \int_{\Gamma_{\Lambda^c}} F(\mathbf{x}_n \cup \eta) e^{-\beta H(\mathbf{x}_n) - \beta W(\mathbf{x}_n | \eta)} d\mathbb{P}(\eta) d\mathbf{x}_n. \quad (\text{R})$$

From our definition of the interaction  $W$  (2.17) it follows that  $W$  is well-defined on  $\Gamma_{\Lambda^c}$ . However, to ensure that the measures satisfying (R) have nice properties, i.e. satisfy a Ruelle condition, we need to make the additional assumption that the measure  $\mathbb{P}$  is supported on a set of „nice“ configurations, cf. [36].

**Definition 2.2.15** *We define the set of tempered configurations as*

$$\Gamma_* := \bigcup_{M \geq 1} \bigcap_{n \geq 1} \{\eta \in \Gamma \mid N_{\Delta_n}(\eta) \leq M \mathfrak{K}(\Delta_n)\}$$

where  $\Delta_n = B_{n+1}(0) \setminus B_n(0)$  for  $n \geq 2$  and  $\Delta_1 = B_1(0)$ .

**Definition 2.2.16** *Any point process  $\mathbb{P}$  with  $\mathbb{P}(\Gamma_*) = 1$  that satisfies the Ruelle-equation (R) for some  $\beta > 0$ ,  $\mu \in \mathbb{R}$  and  $u \in \mathcal{U}$  is called a tempered  $(\beta, \mu, u)$ -Gibbs measure. We denote by  $\mathcal{G}(\beta, \mu, u)$  the set of all tempered  $(\beta, \mu, u)$ -Gibbs measures. Since we will only be looking at tempered Gibbs measures in this work, we often leave out the term tempered.*

Ruelle also showed existence of Gibbs measures with superstable pair interactions:

**Theorem 2.2.17** ([56]) *Let  $\beta > 0$ ,  $\mu \in \mathbb{R}$  and  $u \in \mathcal{U}$ , then  $\mathcal{G}(\beta, \mu, u) \neq \emptyset$ .*

**Remark 2.2.18** *The question of whether (or when) there exists more than one Gibbs measure for a given set of parameters is largely unsolved as of writing this thesis. It can be shown that for  $\mu$  small enough there is only one. However, in general showing the existence of more than one Gibbs measure (taken as the existence of a phase transition in statistical physics) has so far only been proved in models with higher-order interactions such as the Widom-Rowlinson model, see e.g. [57, 5], a four-body Kac-type interaction, cf. [39], or a Quermass interaction, e.g. [9].*

**Remark 2.2.19** *The Ruelle-equation immediately characterizes the Janossy densities of a Gibbs measure: Replacing  $F$  by  $F(\cdot_\Lambda)$  for any  $F: \Gamma \rightarrow [0, \infty)$  we conclude from (2.2) that*

$$j_\Lambda^{(n)}(\mathbf{x}_n) = \int_{\Gamma_{\Lambda^c}} e^{\beta n \mu - \beta H(\mathbf{x}_n) - \beta W(\mathbf{x}_n | \eta)} d\mathbf{P}(\eta), \quad n \in \mathbb{N}_0. \quad (2.34)$$

An equivalent characterization of Gibbs measures is the *GNZ-equation* named for Georgii, Nguyen and Zessin, who introduced the equation separately in [17] and [47].

**Definition 2.2.20** *We say a point process  $\mathbf{P}$  satisfies the GNZ-equation (Georgii-Ngyuen-Zessin) for some chemical potential  $\mu \in \mathbb{R}$ , a pair interaction  $u \in \mathcal{U}$  and  $\beta > 0$  if for every non-negative function  $F: \mathbb{R}^d \times \Gamma \rightarrow [0, \infty]$ , there holds*

$$\int_\Gamma \sum_{x \in \eta} F(x, \eta) d\mathbf{P}(\eta) = \int_{\mathbb{R}^d} \int_\Gamma F(x, \eta \cup \{x\}) e^{\beta \mu - \beta W(\{x\} | \eta)} d\mathbf{P}(\eta) dx. \quad (\text{GNZ})$$

**Theorem 2.2.21** ([47]) *Any point process  $\mathbf{P}$  satisfies (R) if and only if it satisfies (GNZ).*

**Proposition 2.2.22** ([47]) *Let  $\mathbf{P}$  satisfy (GNZ) for some chemical potential  $\mu \in \mathbb{R}$ , a pair interaction  $u \in \mathcal{U}$  and  $\beta > 0$ , then  $\mathbf{P}$  also satisfies the multivariate GNZ-equation (MGNZ), i.e. for every  $m \in \mathbb{N}$  and every non-negative function  $F: (\mathbb{R}^d)^m \times \Gamma \rightarrow [0, \infty]$ , there holds*

$$\int_\Gamma \sum_{\substack{x_1, \dots, x_m \in \eta \\ x_j \neq x_i}} F(\mathbf{x}_m, \eta) d\mathbf{P}(\eta) = \int_{(\mathbb{R}^d)^m} \int_\Gamma F(\mathbf{x}_m, \eta \cup \mathbf{x}_m) e^{n\beta \mu - \beta H(\mathbf{x}_m) - \beta W(\mathbf{x}_m | \eta)} d\mathbf{P}(\eta) dx. \quad (\text{MGNZ})$$

**Remark 2.2.23** *From (MGNZ) and (2.3) it immediately follows that the correlation functions  $(\rho^{(n)})_{n \geq 1}$  of  $\mathbf{P} \in \mathcal{G}(\beta, \mu, u)$  exist and for  $n \in \mathbb{N}$  are given by*

$$\rho^{(n)}(\mathbf{x}_n) = \int_\Gamma e^{\beta n \mu - \beta H(\mathbf{x}_n) - \beta W(\mathbf{x}_n | \eta)} d\mathbf{P}(\eta). \quad (2.35)$$

Furthermore it can be shown, see [56], that in this case the correlation functions satisfy a Ruelle bound  $(\mathcal{R}_\xi)$  with

$$\xi = e^{\beta \mu} \Upsilon(u) \quad (2.36)$$

where  $\Upsilon(u) > 1$  is a constant only depending on the pair interaction  $u \in \mathcal{U}$ .

From Ruelle's equation (R) and (MGNZ) there also follow two ways of characterizing Gibbs measures via integral equations for their correlation functions.

**Definition and Theorem 2.2.24 ([56])** For any  $P \in \mathcal{G}(\beta, \mu, u)$  the correlation functions  $(\rho^{(n)})_{n \geq 1}$  of  $P$  satisfy the Kirkwood-Salsburg equations, i.e. for every  $n \in \mathbb{N}_0$  and  $x_0, x_1, \dots, x_n \in \mathbb{R}^d$  there holds

$$\rho^{(n+1)}(x_0, \mathbf{x}_n) = e^{\beta\mu} e^{-\beta W(x_0|\mathbf{x}_n)} \sum_{k=0}^{\infty} \frac{1}{k!} \int_{(\mathbb{R}^d)^k} \prod_{i=1}^k (e^{-\beta u(x_0-y_i)} - 1) \rho^{(n+k)}(\mathbf{x}_n, \mathbf{y}_k) d\mathbf{y}_k, \quad (2.37)$$

and the Mayer-Montroll equations, i.e. for every  $n \in \mathbb{N}$  and  $x_1, \dots, x_n \in \mathbb{R}^d$  there holds

$$\rho^{(n)}(\mathbf{x}_n) = e^{n\beta\mu} e^{-\beta H(\mathbf{x}_n)} \left( 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \int_{(\mathbb{R}^d)^k} \prod_{i=1}^k (e^{-\beta W(\mathbf{x}_n|y_i)} - 1) \rho^{(k)}(\mathbf{y}_k) d\mathbf{y}_k \right) \quad (2.38)$$

where we use the convention  $\rho^{(0)} = 1$ .

The Mayer-Montroll equations were first introduced in [42] and later generalized by Mayer in [41] as the *Mayer-equations*. However, Kirkwood and Salsburg analyzed a special case of the Mayer-equations, namely the Kirkwood-Salsburg equations, further in [35] and showed that due to their simpler form they are easier to work with. In fact, an easy way of showing existence of Gibbs measures can be achieved via the Kirkwood-Salsburg equations. Furthermore, when writing them as a fixed point equation it can be shown that there is exactly one Gibbs measure if  $\mu$  is small enough. More precisely, for

$$e^{\beta\mu} < \frac{1}{C(u, \beta) e^{2B+1}} \quad (2.39)$$

the set  $\mathcal{G}(\beta, \mu, u)$  has exactly one Element. In statistical physics this is called the *gas phase*. In Chapter 4 we also use the Kirkwood-Salsburg equations to prove a continuity type property of Gibbs measures. The last characterization we want to introduce is via the so-called *Gibbs variational principle*, basically the infinite volume version of (2.33). Such a variational principle was proved for the case of a hard-core interaction by Galavotti and Miracle-Sole in [16] and later further discussed by Zessin and Nguyen in [46]. For general superstable pair interactions a variational principle was proved in a series of papers by Georgii and Zessin, cf. [21, 18, 19]. This is the version we work with for our result in Chapter 3, i.e. [15]. For completeness we also mention a variational principle proved by Dereudre for the case of finite-range interactions, see [7].

When taking the thermodynamic limit of (2.33) all the appearing objects have well-defined limits under certain conditions. The existence of the *grand-canonical pressure*, i.e. the limit

$$p(\beta, \mu, u) = \lim_{\Lambda \nearrow \mathbb{R}^d} p_{\Lambda}(\beta, \mu, u),$$

was first shown for potentials bounded from above in [54] and later by Dobrushin and Minlos [11] for potentials which are non-integrably divergent at the origin, and by Ginibre [23] for positive and hard-core interactions. Later, all these approaches were combined by Ruelle in [56] under the umbrella of superstable interactions. Here by writing  $\Lambda \nearrow \mathbb{R}^d$ , we mean taking a sequence of increasing sets  $\Lambda_n \subset \Lambda_{n+1}$  such that for any  $\Delta \in \mathcal{B}_b(\mathbb{R}^d)$  there is a  $n_{\Delta}$  with  $\Delta \subset \Lambda_{n_{\Delta}}$ . The limit of the entropy  $S_{\Lambda}(P)$  was first discussed in the continuum case by Robinson and Ruelle in [53], where they showed that the *specific entropy* of  $P$  defined by

$$S(P) = \lim_{\Lambda \nearrow \mathbb{R}^d} \frac{1}{\mathfrak{K}(\Lambda)} S_{\Lambda}(P) \quad (2.40)$$

is well-defined with  $S(P) \in [-1, \infty]$ . The limit of the expectation of the number of particles trivially has a limit for stationary point processes with correlation functions as by Example 2.1.8 we have

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \frac{1}{\mathfrak{K}(\Lambda)} \mu \mathbb{E}_P [N_{\Lambda}(\eta)] = \lim_{\Lambda \nearrow \mathbb{R}^d} \frac{1}{\mathfrak{K}(\Lambda)} \mu \mathfrak{K}(\Lambda) \rho = \mu \rho.$$

Lastly, the limit of the expectation of the energy per volume is known to exist in this case as well and we have

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \frac{1}{\mathfrak{K}(\Lambda)} \mathbb{E}_{\mathbf{P}} \left[ \frac{1}{2} \sum_{x \neq y \in \eta_\Lambda} u(x-y) \right] = \frac{1}{2} \int_{\mathbb{R}^d} u(x) \rho^{(2)}(0, x) dx,$$

cf. [19] and Chapter 3.

**Definition 2.2.25** *Let  $\mathbf{P}$  be a translation invariant point process with density  $\rho$  and pair correlation function  $\rho^{(2)}$  and  $\beta > 0$ ,  $\mu \in \mathbb{R}$  and  $u \in \mathcal{U}$  be given, then the specific relative entropy of  $\mathbf{P}$  with respect to  $(\beta, \mu, u)$ -Gibbs measures is defined by*

$$S(\mathbf{P} \mid \beta, \mu, u) := \frac{1}{\beta} S(\mathbf{P}) + p(\beta, \mu, u) - \mu\rho + \frac{1}{2} \int_{\mathbb{R}^d} u(x) \rho^{(2)}(0, x) dx. \quad (2.41)$$

When working at fixed  $\beta > 0$  we will also write  $S(\mathbf{P} \mid \mu, u)$  in place of  $S(\mathbf{P} \mid \beta, \mu, u)$ .

**Theorem 2.2.26 (Gibbs variational principle [19])** *For any translation invariant point process with correlation functions there holds*

$$S(\mathbf{P} \mid \beta, \mu, u) \geq 0,$$

with equality if and only if  $\mathbf{P} \in \mathcal{G}(\beta, \mu, u)$ .

**Remark 2.2.27** *Note, in the thermodynamic limit  $S(\mathbf{P} \mid \beta, \mu, u) = 0$  and  $S(\mathbf{Q} \mid \beta, \mu, u) = 0$  does not imply  $\mathbf{P} = \mathbf{Q}$ .*

This version of the variational principle is one of our main tools in Chapters 3 and 4 where we use it to prove Henderson's Theorem in the thermodynamic limit and show strict convexity of the pressure and relative entropy functional.

## 2.2.4 Cluster Expansions for pair interactions

Lastly, we want to introduce so-called *cluster expansions*, which are the main tool behind our results in Chapter 5. Let  $H$  be given by (2.16) for some function  $u \in \mathcal{U}$ . Recall that by (2.30) the finite volume grand-canonical pressure is given by

$$p_\Lambda(\beta, \mu, u) = \frac{1}{\beta \mathfrak{K}(\Lambda)} \log \Xi_\Lambda(\beta, \mu, u)$$

and since  $\Xi_\Lambda(\beta, \mu, u)$  is a power series of  $e^{\beta\mu}$  ( $z = e^{\beta\mu}$  is called *activity*), the pressure must also have such a power series expansion for sufficiently small  $\mu$ . We make use of Theorem 2.1.21 to find this expansion.

**Definition 2.2.28** *For a chemical potential  $\mu$  and a Hamiltonian  $H$  the Ursell functions  $(\omega^{(n)})_{n \geq 1}$  are recursively defined as  $\omega^{(1)} \equiv e^{\beta\mu}$  and for  $n \geq 2$  as*

$$\omega^{(n)}(\mathbf{x}_n) = (e^{\beta\mu})^n e^{-\beta H(\mathbf{x}_n)} - \sum_{k=2}^n \sum_{\pi \in \Pi_k(\mathbf{x}_n)} \prod_{i=1}^k \omega^{(\kappa_i)}(\pi_i). \quad (2.42)$$

Note that this definition of the Ursell functions does not rest on the Hamiltonian  $H$  only consisting of pair interactions, we will now derive an easier formula for the case it does. We first note that we can write

$$e^{-\beta H(\mathbf{x}_n)} = \prod_{i < j} (1 + f_{ij}) = \sum_{G \in \mathcal{G}_n} \prod_{\{i,j\} \in E(G)} f_{ij} \quad (2.43)$$

where  $f_{ij} = f(x_i - x_j) = e^{-\beta u(x_i - x_j)} - 1$ ,  $\mathcal{G}_n$  is the set of graphs on  $n$  vertices and  $E(G)$  is the set of edges of the graph  $G$ . The idea is to find appropriate functions  $(F_n)_{n \geq 1}$  to write  $e^{-H(\mathbf{x}_n)}$  in the form (2.11). For that we notice that every graph  $G$  in  $\mathcal{G}_n$  splits into its connected components  $G_1, \dots, G_r$  and there obviously holds

$$\prod_{\{i,j\} \in E(G)} f_{ij} = \prod_{l=1}^r \prod_{\{i,j\} \in E(G_l)} f_{ij}.$$

Defining

$$\tilde{\omega}^{(n)}(\mathbf{x}_n) := (e^{\beta\mu})^n \sum_{C \in \mathcal{C}_n} \prod_{\{i,j\} \in E(C)} f_{ij}, \quad (2.44)$$

where  $\mathcal{C}_n$  is the set of connected graphs on  $n$  vertices (or nodes), we get

$$(e^{\beta\mu})^n e^{-\beta H(\mathbf{x}_n)} = \sum_{k=1}^n \sum_{\pi \in \Pi_k(\mathbf{x}_n)} \prod_{i=1}^k \tilde{\omega}^{(\kappa_i)}(\pi_i)$$

and thus  $\tilde{\omega}^{(n)} = \omega^{(n)}$ . As we want to use Theorem 2.1.21, we need to show that the Ursell functions  $(\omega^{(n)})_{n \geq 1}$  satisfy a bound of the type (2.10). However, since the number of connected graphs grows like  $2^{\binom{n}{2}}$ , finding an upper bound is not straight forward. The most well-known approach is the so-called *tree graph inequality* first found by Groeneveld in [24, 25]. Over the years several improvements of the tree-graph inequality have been made by choosing specific ways to construct trees out of the connected graphs. The version presented here was first found by Procacci in [52].

**Theorem 2.2.29 (Tree-graph inequality [52])** *Let  $u \in \mathcal{U}$ , then for all  $n \in \mathbb{N}$  and  $x_1, \dots, x_n \in \mathbb{R}^d$  there holds*

$$\left| \sum_{C \in \mathcal{C}_n} \prod_{\{i,j\} \in E(C)} f_{ij} \right| \leq e^{\beta B n} \sum_{T \in \mathcal{T}_n} \prod_{\{i,j\} \in E(T)} (1 - e^{-\beta|u(x_i - x_j)|}) \leq e^{\beta B n} \sum_{T \in \mathcal{T}_n} \prod_{\{i,j\} \in E(T)} |f_{ij}|. \quad (2.45)$$

Since every tree has exactly  $n - 1$  edges, it is well-known that the number of trees on  $n$  vertices is  $n^{n-2}$ , and  $n^{n-2} \leq (n - 2)!e^n$ , we can conclude from (2.45) that

$$\int_{\Lambda^k} |\omega^{(n)}(\mathbf{x}_n)| d\mathbf{x}_n \leq n! \mathfrak{A}(\Lambda) e^{(\beta\mu + \beta B + 1)n} C(\beta, u)^{n-1}. \quad (2.46)$$

In particular, for  $\mu$  small enough we can use Theorem 2.1.21 to rewrite the grand-canonical partition function as

$$\Xi_\Lambda = \exp \left( \sum_{n=1}^{\infty} \frac{(e^{\beta\mu})^n}{n!} \int_{\Lambda^n} \sum_{C \in \mathcal{C}_n} \prod_{\{i,j\} \in E(C)} f_{ij} d\mathbf{x}_n \right) \quad (2.47)$$

and use Theorem 2.1.21 for  $\mu$  small enough to get

$$p_\Lambda(\beta, \mu, u) = \frac{1}{\beta \mathfrak{K}(\Lambda)} \sum_{n=1}^{\infty} \frac{(e^{\beta\mu})^n}{n!} \int_{\Lambda^n} \sum_{C \in \mathcal{C}_n} \prod_{\{ij\} \in E(C)} f_{ij} \, d\mathbf{x}_n.$$

It can be shown, e.g. [49], that in the thermodynamic limit there holds

$$p(\beta, \mu, u) = \frac{1}{\beta} \sum_{n=0}^{\infty} \frac{(e^{\beta\mu})^{n+1}}{(n+1)!} \int_{(\mathbb{R}^d)^n} \sum_{C \in \mathcal{C}_{n+1}} \prod_{\{ij\} \in E(C)} f_{ij} \, d\mathbf{x}_n. \quad (2.48)$$

**Remark 2.2.30** *Expansions of the type (2.48) are so-called cluster expansions. These expansions have long been of big interest in statistical physics, see e.g. the review [59] where many of these have been collected. However, proving convergence of these expansions has only been possible for a small selection of them, e.g. the activity expansion of the pressure (2.48) and the virial expansion, which is an expansion of the pressure in terms of the density  $\rho$  and the Mayer-function  $f$ .*

To conclude this chapter, we want to mention that similar to the pressure we can get an expansion of the truncated correlation functions of a Gibbs measure, namely

$$\rho_T^{(n)}(\mathbf{x}_n) = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{(\mathbb{R}^d)^k} \omega^{(n+k)}(\mathbf{x}_n, \mathbf{y}_k) \, d\mathbf{y}_k. \quad (2.49)$$

From this expansion it is also possible to get different decay properties of the correlation function, see Chapter 4.



# 3

## A note on the uniqueness result for the Henderson problem

### 3.1 Overview

As mentioned in Remark 2.1.9, contrary to the chemical potential or pair interaction, the correlation functions, in particular the density and the pair correlation function, can be extracted from simulation or experiment data, and while there are some formulas to calculate the correlation functions for given chemical potential and pair interaction, e.g. (2.35) and (2.49), the inverse is not true (if only given the density and pair correlation function).

This leads to the so-called *inverse Henderson problem*: Given the density  $\rho$  and pair correlation function  $\rho^{(2)}$ , determine the chemical potential  $\mu$  and pair potential  $u$ . In this section we are specifically concerned with the question whether the *Henderson operator*

$$\begin{aligned} F: \mathbb{R} \times \mathcal{U} &\rightarrow \mathbb{R} \times L^\infty(\mathbb{R}^d) \\ (\mu, u) &\mapsto (\rho, \rho^{(2)}) \end{aligned} \tag{3.1}$$

is injective. In other words, can there be two pairs  $(\mu, u)$  and  $(\nu, v)$  with  $\mu, \nu \in \mathbb{R}$  and  $u, v \in \mathcal{U}$  and Gibbs-measures  $P \in \mathcal{G}(\beta, \mu, u)$  and  $P' \in \mathcal{G}(\beta, \nu, v)$  such that  $\rho_P = \rho_{P'}$  and  $\rho_P^{(2)} = \rho_{P'}^{(2)}$ .

In the framework of the canonical ensemble, see Remark 2.2.8, the density  $\rho$  of the system is prescribed since the number of particles is fixed and thus we only look at the pair interaction  $u$  and the radial distribution function  $g$ , i.e. at the *canonical Henderson operator*

$$\begin{aligned} F_C: \mathcal{U} &\rightarrow L^\infty(\mathbb{R}^d) \\ u &\mapsto g. \end{aligned} \tag{3.2}$$

This case was looked at by Henderson in [31], where he showed that the canonical Henderson operator is injective; this result is known as *Henderson's theorem*. He also investigates the thermodynamic limit, however, the variational principle used by him does not extend to this case and his argument lacks rigor. Using the Gibbs variational principle established by Georgii (Theorem 2.2.26) we can prove the following results.

**Theorem 3.1.1** *Let  $u, v \in \mathcal{U}$ ,  $\beta > 0$ , and  $\mu, \nu \in \mathbb{R}$  be given, and assume that some  $P \in \mathcal{G}(\beta, \mu, u)$  and  $P' \in \mathcal{G}(\beta, \nu, v)$  admit the same density  $\rho$  and the same pair correlation function  $\rho^{(2)}$ . Then  $\mu = \nu$  and  $u = v$  almost everywhere.*

**Theorem 3.1.2** *Let  $u, v \in \mathcal{U}$ ,  $\beta > 0$ , and  $\mu \in \mathbb{R}$  be given, and assume that some  $P \in \mathcal{G}(\beta, \mu, u)$  and  $P' \in \mathcal{G}(\beta, \mu, v)$  admit the same pair correlation function  $\rho^{(2)}$ . Then  $u = v$  almost everywhere.*

**Remark 3.1.3** *Note that to ensure this uniqueness we need to restrict ourselves to Gibbs measures with pair interactions. Consider the following, let  $\beta > 0$ ,  $\mu \in \mathbb{R}$  and  $u$  be non-negative or with a hard-core and  $\rho$  and  $\rho^{(2)}$  correspond to some  $P \in \mathcal{G}(\beta, \mu, u)$ . Then the Kirkwood-closure process  $K_{\rho, g}$  with the same density and pair correlation function exists, but in general  $K_{\rho, g} \notin \mathcal{G}(\beta, \mu, u)$ .*

**Remark 3.1.4** *In principle, it is also possible to extend the result from Theorem 3.1.1 to special cases of higher-order interactions and higher-order correlation functions, using the Gibbs variational principles by Dereudre, cf. [7], or Zessin, see [46]. In the finite volume case this has been done by Wang, Stillinger and Torquato in [60].*

## 3.2 Reprint

The article is reprinted by permission from Springer Nature: Journal of Mathematical Physics: A note on the uniqueness result for the inverse Henderson problem, Fabio Frommer, Martin Hanke, Sabine Jansen, 2019, 60, 9, 27 September 2019 Springer Science+Business Media (doi.org/10.1063/1.5112137).

RESEARCH ARTICLE | SEPTEMBER 27 2019

## A note on the uniqueness result for the inverse Henderson problem

F. Frommer ; M. Hanke ; S. Jansen

Check for updates

*J. Math. Phys.* 60, 093303 (2019)  
<https://doi.org/10.1063/1.5112137>



### Articles You May Be Interested In

On the analyticity of the pressure for a non-ideal gas with high density boundary conditions

*J. Math. Phys.* (May 2023)

Uniqueness and absence of percolation of classical gases

*J. Math. Phys.* (February 2023)

Revisiting Groeneveld's approach to the virial expansion

*J. Math. Phys.* (February 2021)

16 October 2024 14:39:42



## Special Topics Open for Submissions

[Learn More](#)

# A note on the uniqueness result for the inverse Henderson problem

Cite as: J. Math. Phys. 60, 093303 (2019); doi: 10.1063/1.5112137

Submitted: 3 June 2019 • Accepted: 4 September 2019 •

Published Online: 27 September 2019



F. Frommer,<sup>1,a)</sup>  M. Hanke,<sup>1,b)</sup>  and S. Jansen<sup>2,c)</sup> 

## AFFILIATIONS

<sup>1</sup>Institut für Mathematik, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany

<sup>2</sup>Institut für Mathematik, Ludwig-Maximilians Universität München, Theresienstr. 39, 80333 München, Germany

<sup>a)</sup>fafromme@uni-mainz.de

<sup>b)</sup>hanke@math.uni-mainz.de

<sup>c)</sup>jansen@math.lmu.de

## ABSTRACT

The inverse Henderson problem of statistical mechanics is the theoretical foundation for many bottom-up coarse-graining techniques for the numerical simulation of complex soft matter physics. This inverse problem concerns classical particles in continuous space which interact according to a pair potential depending on the distance of the particles. Roughly stated, it asks for the interaction potential given the equilibrium pair correlation function of the system. In 1974, Henderson proved that this potential is uniquely determined in a canonical ensemble and he claimed the same result for the thermodynamical limit of the physical system. Here, we provide a rigorous proof of a slightly more general version of the latter statement using Georgii's variant of the Gibbs variational principle.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5112137>

## I. INTRODUCTION

The numerical simulation of soft matter requires a large variety of multiscale techniques—even on high-performance computers of the latest generation—to get close to relevant time and length scales, see the survey of Potestio, Peters, and Kremer,<sup>37</sup> or the special issue<sup>7</sup> of The European Physical Journal, Special Topics.

Among these multiscale ingredients is a technique known as (bottom-up) *coarse-graining*:<sup>34,35,39</sup> complex systems are represented, where possible, by fictitious particles (called *beads*) with reduced degrees of freedom, and the fine-grained details are only reinserted into the overall system when necessary physically, cf. Praprotnik, Delle Site, and Kremer.<sup>38</sup> The construction of such a coarse-grained system requires (i) a sophisticated selection of the molecular pieces to be represented by a single bead (and its corresponding position, mass, and shape) and (ii) the design of interacting forces between individual beads. To be thermodynamically consistent, the latter should be defined via the so-called multibody potential of mean force (compare, e.g., Noid<sup>34</sup>) which, alas, is computationally intractable. A common alternative is to ignore three- or more-particle interactions and to calculate effective pair potentials by fitting structural properties of the coarse-grained system to real measurements (neutron scattering experiments) or data from fine-grained numerical case studies.

For the simplest conceivable case, let us assume that the coarse-grained system consists of a translation and rotation invariant ensemble of a single type of beads and that data are given for the so-called *radial distribution function*  $g$  of these beads when the fine-grained system is in thermodynamical equilibrium. Roughly speaking, this function assigns to each  $r > 0$  the expected number  $g(r)$  of beads on a sphere of radius  $r$  around any given bead, normalized by the surface area of the sphere and the square of the density  $\rho^{(1)}$  of the system. Then, the aforementioned problem amounts to finding a pair potential for the interaction of classical pointlike particles such that the statistical distribution of the corresponding canonical or grand canonical ensemble matches the given data. This is a typical instance of an inverse problem,<sup>18</sup> where the cause for a given observation is sought.

In an often cited paper, which may be considered the theoretical basis of modern bottom-up coarse-graining techniques, Henderson<sup>20</sup> has claimed that under given conditions of temperature and density, there is indeed at most one pair potential depending only on the distance

between the interacting particles, for which the statistics of the corresponding ensemble obeys a given radial distribution function in the *thermodynamical limit*, i.e., as volume and particle count go to infinity. To give credit to Henderson's contribution, we will refer to this inverse problem of statistical mechanics as the inverse Henderson problem.

Since there exist only approximate identities connecting a given pair potential with the associated radial distribution function (see Hansen and McDonald<sup>41</sup>) and those fail to provide sufficiently accurate potentials in general (Schommers<sup>42</sup>), physicists have started in the 1970s to design iterative algorithms for the numerical solution of the inverse Henderson problem,<sup>28,29,39,43,45,46</sup> and this still is a very active area of research.<sup>9,19,31,48</sup>

Today, one can say that the determination of effective pair potentials for a given radial distribution function is the working horse in state-of-the-art coarse-graining technology, and it has been applied successfully to a wide range of challenging complex physical, chemical, and biochemical case studies, see Refs. 1, 3, 30, and 36 and the references therein.

In his fundamental paper, Henderson employs a technique suggested by Hohenberg and Kohn,<sup>21</sup> Mermin,<sup>32</sup> and others for studying a similar inverse problem for *external potentials*; see also Refs. 5 and 22. The key idea is to apply a *Gibbs variational principle*, which states that in a system with given thermodynamic conditions, the associated thermodynamic potential becomes minimal, if and only if the distribution of the particles is given by the probability measure associated with this system. The particular version of this principle used by Henderson is based on the free energy functional in a canonical ensemble, where the finite volume pair correlation function defined in (2) below, and not the radial distribution function, is the relevant stochastic quantity. (The pair correlation function is also known as pair density function in the pertinent literature, e.g., Ref. 11). To extend the uniqueness result to the radial distribution function, Henderson subsequently turns to the thermodynamical limit, ignoring the possibility that the strict inequality of the variational principle for finite volumes may turn into an equality when the volume tends to infinity. Accordingly, there is a gap in the argument provided in Ref. 20—aside of the fact that no mention is made concerning the necessary requirements for the pair potential, e.g., its behavior for particle pairs with diminishing distances.

Given the importance of Henderson's statement, the purpose of this note is to fix this gap and to provide a rigorous proof of his result by using a version of the Gibbs variational principle due to Georgii.<sup>13</sup> We show that Henderson's result holds true for a suitably rich class of pair potentials including the most relevant ones for physical applications, namely, hard core potentials and the so-called Lennard-Jones type potentials. This class of potentials, however, is a strict subclass of all superstable potentials, for which the thermodynamical limit is well-defined, see the work of Ruelle<sup>41</sup> for this and further terminology. Strictly speaking, this means that our result does not answer the question whether the radial distribution function associated with, say, the classical Lennard-Jones potential can also occur for a much more exotic type of pair potential and the same values of temperature and density.

The thermodynamical limit may either be reached from a canonical or a grand canonical ensemble. We therefore also state a variant of Henderson's result which is more natural from the grand canonical perspective. It will be shown below that the pair potential is uniquely determined when given the temperature, the chemical potential, and the infinite volume pair correlation function; it is unknown whether in this second version of Henderson's statement, the pair correlation function can be replaced by the radial distribution function in the isotropic case.

For completeness, we mention that the uniqueness result for the inverse Henderson problem on the lattice is contained in the work by Griffiths and Ruelle;<sup>17</sup> see also the work of Caglioti, Kuna, Lebowitz, and Speer.<sup>2</sup>

The outline of this note is as follows: In Sec. II, we review the rigorous mathematical setting of the thermodynamical limit of a grand canonical ensemble when the system is translation invariant and its potential energy is given by pairwise interactions only. Then, we formulate in Sec. III the particular version of the Gibbs variational principle that is valid in this setting. Section IV is devoted to the proof of the uniqueness results, and eventually we close with a few comments and open problems in Sec. V.

## II. THE THERMODYNAMICAL LIMIT OF THE GRAND CANONICAL ENSEMBLE

We start from a grand canonical ensemble of pointlike classical particles in a bounded box  $\Lambda_\ell = [-\ell, \ell]^d$ , with specified inverse temperature  $\beta > 0$  and chemical potential  $\mu \in \mathbb{R}$ . We restrict our attention to the case that the interaction of the particles is given by a pair potential  $u : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ , which is an even function, i.e.,  $u(x) = u(-x)$ , satisfying the following two assumptions:

1. There exists  $r_0 > 0$  and a decreasing function  $\varphi : (0, r_0] \rightarrow \mathbb{R}_0^+$  with

$$\int_0^{r_0} r^{d-1} \varphi(r) dr = +\infty$$

and

$$u(x) \geq \varphi(|x|), \quad \text{for } |x| \leq r_0.$$

2. There exists a decreasing function  $\psi : [r_0, \infty) \rightarrow \mathbb{R}_0^+$  with

$$\int_{r_0}^{\infty} r^{d-1} \psi(r) dr < \infty$$

and

$$|u(x)| \leq \psi(|x|), \quad \text{for } |x| \geq r_0. \quad (1)$$

For this class  $\mathcal{U}$  of potentials, the associated configurational Hamiltonian of  $m \in \mathbb{N}_0$  particles at positions  $x_i \in \mathbb{R}^d$ ,  $i = 1, \dots, m$ , given by

$$H_u(\mathbf{x}_m) = \sum_{1 \leq i < j \leq m} u(x_i - x_j), \quad \mathbf{x}_m = (x_1, \dots, x_m),$$

is stable (cf. Dobrushin<sup>8</sup>), i.e., for every  $u \in \mathcal{U}$ , there exists  $B > 0$  such that

$$H_u(\mathbf{x}_m) \geq -Bm,$$

independent of the number  $m$  of particles. The statistical distribution of the particles of such a grand canonical ensemble is determined by the corresponding  $m$ -particle correlation functions

$$\rho_{\Lambda_\ell}^{(m)}(\mathbf{x}_m) = \frac{e^{\beta\mu m}}{\Xi(\Lambda_\ell, \beta, \mu, u)} \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{N!} \int_{\Lambda_\ell^N} e^{-\beta H_u(\mathbf{x}_m, \mathbf{y}_N)} d\mathbf{y}_N, \quad (2)$$

where  $\mathbf{x}_m = (x_1, \dots, x_m) \in \Lambda_\ell^m$ ,  $\mathbf{y}_N = (y_1, \dots, y_N) \in \Lambda_\ell^N$ , the integral  $\int_{\Delta} c dx_0$  with bounded domain  $\Delta \subset \mathbb{R}^d$  is always taken to be equal to  $c$ , and the normalizing constant

$$\Xi(\Lambda_\ell, \beta, \mu, u) = \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{N!} \int_{\Lambda_\ell^N} e^{-\beta H_u(\mathbf{x}_N)} d\mathbf{x}_N$$

is the associated grand canonical partition function. In (2),  $m$  varies in  $\mathbb{N}_0$ , with  $\rho_{\Lambda_\ell}^{(0)}$  being set to one.

We assume that for some sequence  $(\ell_k)_k$  going to infinity as  $k \rightarrow \infty$ , these correlation functions converge uniformly on compact subsets to translation-invariant functions  $\rho^{(m)} : (\mathbb{R}^d)^m \rightarrow \mathbb{R}_0^+$ ,  $m \in \mathbb{N}_0$ , defined on the entire space; in particular, this implies that  $\rho^{(1)}$  is a constant. It is known that these limiting correlation functions satisfy a so-called *Ruelle bound*, i.e.,

$$\sup_{\mathbf{x}_m \in (\mathbb{R}^d)^m} \rho^{(m)}(\mathbf{x}_m) \leq \xi^m, \quad m \in \mathbb{N}_0, \quad (3)$$

for some  $\xi > 0$ , depending only on  $\mu$ ,  $\beta$ , and  $u$ , and that they define a translation invariant probability measure  $\mathbb{P}$  on the configuration space

$$\Gamma = \{ \gamma \subset \mathbb{R}^d \mid \Delta \subset \mathbb{R}^d \text{ bounded} \Rightarrow \#(\gamma \cap \Delta) < \infty \},$$

i.e., the set of all locally finite subsets of  $\mathbb{R}^d$  representing the positions of the (at most countably many) individual particles in space, equipped with an appropriate  $\sigma$ -algebra, cf. the work of Ruelle.<sup>42</sup> This means that if  $m \in \mathbb{N}_0$  is fixed and an observable  $F$  depends on all possible  $m$ -tuples of particles in a given configuration, i.e.,

$$F(\gamma) = \sum_{\substack{x_1, \dots, x_m \in \gamma \\ x_i \neq x_j}} f(\mathbf{x}_m),$$

for some  $f = f_1 + f_2$  with  $f_1 \in L^1((\mathbb{R}^d)^m)$  and  $f_2 \geq 0$ , then

$$\int_{\Gamma} F(\gamma) d\mathbb{P}(\gamma) = \int_{(\mathbb{R}^d)^m} f(\mathbf{x}_m) \rho^{(m)}(\mathbf{x}_m) d\mathbf{x}_m \quad (4)$$

is the expected value of the corresponding observable. In particular, if  $|\Delta|$  denotes the volume of any bounded domain  $\Delta \subset \mathbb{R}^d$ , then

$$\rho(\mathbb{P}) = \frac{1}{|\Delta|} \int_{\Gamma} \#(\gamma \cap \Delta) d\mathbb{P}(\gamma) = \rho^{(1)}$$

is the limiting particle counting density.

According to Ref. 42,  $\mathbb{P}$  is a translation invariant tempered  $(\beta, \mu, u)$ -Gibbs measure, denoted  $\mathbb{P} \in \mathcal{G}(\beta, \mu, u)$ . This means that  $\mathbb{P}$  is supported by the set of tempered configurations (defined in Ref. 42) and that for every  $F \in L^1(\mathbb{P})$  and every bounded domain  $\Delta \subset \mathbb{R}^d$ , there holds

$$\int_{\Gamma} F(\gamma) d\mathbb{P}(\gamma) = \sum_{N=0}^{\infty} \frac{z^N}{N!} \int_{\Delta^N} \int_{\Gamma(\Delta^c)} F(\gamma') e^{-\beta W_u(\mathbf{x}_N; \gamma')} d\mathbb{P}(\gamma) e^{-\beta H_u(\mathbf{x}_N)} d\mathbf{x}_N, \quad (5)$$

where  $\gamma' = \gamma \cup \{x_1, \dots, x_N\}$ ,  $\Gamma(\Delta^c) = \{\gamma \in \Gamma : \gamma \subset \mathbb{R}^d \setminus \Delta\}$ , and the interaction  $W_u$  between particles at  $x_i$ ,  $i = 1, \dots, N$  and those of  $\gamma \in \Gamma$  is defined as

$$W_u(\mathbf{x}_N; \gamma) = \sum_{i=1}^N \sum_{y \in \gamma} u(x_i - y), \tag{6}$$

if the series converges absolutely, and as  $+\infty$  otherwise.

Given the limiting correlation functions, one can define Janossy densities  $j_{\Lambda_\ell}^{(m)} : \Lambda_\ell^m \rightarrow \mathbb{R}$  for every  $m \in \mathbb{N}_0$  and  $\ell > 0$  via

$$j_{\Lambda_\ell}^{(m)}(\mathbf{x}_m) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{\Lambda_\ell^k} \rho^{(m+k)}(\mathbf{x}_m, \mathbf{y}_k) d\mathbf{y}_k. \tag{7}$$

These Janossy densities provide the induced probability density on  $\Lambda_\ell$ , for which

$$\int_{\Gamma} F(\gamma) d\mathbb{P}(\gamma) = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\Lambda_\ell^m} f_m(\mathbf{x}_m) j_{\Lambda_\ell}^{(m)}(\mathbf{x}_m) d\mathbf{x}_m, \tag{8}$$

for every  $F \in L^1(\mathbb{P})$ , which satisfies  $F(\gamma) = F(\gamma \cap \Lambda_\ell)$ , and which is given by functions  $f_m : \Lambda_\ell^m \rightarrow \mathbb{R}$ ,  $m \in \mathbb{N}_0$ , such that  $F(\gamma_m) = f_m(\mathbf{x}_m)$ , when  $\gamma_m = \{x_1, \dots, x_m\} \subset \Lambda$ . Such observables  $F$  are thus called *local observables*.

Varying  $\beta > 0$ ,  $\mu \in \mathbb{R}$ , and  $u \in \mathcal{U}$ , we denote by

$$\mathcal{G} = \bigcup_{\beta, \mu, u} \mathcal{G}(\beta, \mu, u)$$

the union of all translation invariant tempered Gibbs measures, some of which may not be obtained as limits of finite-volume Gibbs measures with empty boundary conditions (see the work of Georgii<sup>12</sup>). We mention for later use that for almost every  $x \in \mathbb{R}^d$  and  $\mathbb{P}$  almost surely for every  $\mathbb{P} \in \mathcal{G}$ , the interaction defined in (6) is finite, and there holds

$$\lim_{\ell \rightarrow \infty} W_u(x; \gamma \cap \Lambda_\ell) = W_u(x; \gamma); \tag{9}$$

see Sec. V in Kondratiev and Kuna.<sup>23</sup> We mention further that there exists some  $\mu_0 \in \mathbb{R}$  depending on  $\beta$  and  $u \in \mathcal{U}$ , such that for  $\mu < \mu_0$ —the so-called *gas phase*—the set  $\mathcal{G}(\beta, \mu, u)$  consists of a single Gibbs measure  $\mathbb{P}$  only<sup>42</sup> and that in this case, the correlation functions  $\rho_{\Lambda_\ell}^{(m)}$  converge to  $\rho^{(m)}$  as  $\ell \rightarrow \infty$  for every  $m \in \mathbb{N}_0$ , cf. Ref. 41.

### III. THE GIBBS VARIATIONAL PRINCIPLE

The Gibbs variational principle goes back to Gibbs' work (cf., Ref. 16, p. 131) and appears in different variants in statistical mechanics and stochastic analysis; we refer to the books by Ruelle,<sup>41</sup> Gallavotti,<sup>10</sup> and Georgii<sup>14</sup> for rigorous mathematical treatments of this variational principle. Here, we apply a particular version established by Georgii and Zessin in a series of papers.<sup>12,13,15</sup>

For pair potentials  $u \in \mathcal{U}$  and Gibbs measures  $\mathbb{P} \in \mathcal{G}$ , we introduce the *specific energy*

$$E(u, \mathbb{P}) = \lim_{\ell \rightarrow \infty} \frac{1}{|\Lambda_\ell|} \int_{\Gamma} \frac{1}{2} \sum_{\substack{x, y \in \gamma \cap \Lambda_\ell \\ x \neq y}} u(x - y) d\mathbb{P}(\gamma) \tag{10}$$

and the *specific (relative) entropy*

$$S(\mathbb{P}) = \lim_{\ell \rightarrow \infty} \frac{1}{|\Lambda_\ell|} \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\Lambda_\ell^m} j_{\Lambda_\ell}^{(m)}(\mathbf{x}_m) \log(j_{\Lambda_\ell}^{(m)}(\mathbf{x}_m)) d\mathbf{x}_m,$$

where both limits are known to exist in  $\mathbb{R} \cup \{+\infty\}$ : concerning the specific energy, see Proposition 1; concerning the specific entropy, we refer to the work of Robinson and Ruelle<sup>40</sup>—in fact, using (7) and (3), it is not too difficult to see that  $S(\mathbb{P})$  is finite for every  $\mathbb{P} \in \mathcal{G}$ . The relative entropy differs from the standard entropy by a sign; to simplify language, we call  $S(\mathbb{P})$  nevertheless the entropy. We take similar liberties with the sign of the (specific) *grand potential*

$$\Omega_{\beta, \mu}(u, \mathbb{P}) = \mu \rho(\mathbb{P}) - E(u, \mathbb{P}) - \frac{1}{\beta} S(\mathbb{P}), \tag{11}$$

for which the following variational principle holds true; see Theorem 3.4 in the work of Georgii.<sup>13</sup>

**Theorem A.** For fixed  $\mu \in \mathbb{R}$ ,  $\beta > 0$ , and  $u \in \mathcal{U}$ , the grand potential  $\Omega_{\beta,\mu}(u, \cdot)$  has values in  $\mathbb{R} \cup \{-\infty\}$ . Its maximal value  $p$  on  $\mathcal{G}$  is attained for every  $\mathbb{P} \in \mathcal{G}(\beta, \mu, u)$ , and there holds

$$\Omega_{\beta,\mu}(u, \mathbb{P}) < p$$

for every other  $\mathbb{P} \in \mathcal{G}$ . Here,

$$p = \lim_{\ell \rightarrow \infty} \frac{1}{\beta|\Lambda_\ell|} \log \Xi(\Lambda_\ell, \beta, \mu, u)$$

is the pressure in the thermodynamical limit.

For the proof of the Henderson result, we will also need the following identity, for which we include a self-contained proof for the ease of the reader.

*Proposition 1.* For every  $u \in \mathcal{U}$  and every  $\mathbb{P} \in \mathcal{G}$ , the limit in (10) belongs to  $\mathbb{R} \cup \{+\infty\}$  and is given by

$$E(u, \mathbb{P}) = \frac{1}{2} \int_{\mathbb{R}^d} u(x) \rho^{(2)}(x, 0) dx, \quad (12)$$

where  $\rho^{(2)}$  is the pair correlation function associated with  $\mathbb{P}$ .

*Proof.* Let  $\rho^{(2)}$  be the translation invariant pair correlation function associated with  $\mathbb{P}$ . Then, we can apply (4) with  $m = 2$  and

$$f(x, y) = \begin{cases} u(x-y), & x, y \in \Lambda_\ell, \\ 0, & \text{else,} \end{cases}$$

to rewrite

$$\int_{\Gamma} \sum_{\substack{x, y \in \Lambda_\ell \\ x \neq y}} u(x-y) d\mathbb{P}(y) = \int_{\Lambda_\ell^2} u(x-y) \rho^{(2)}(x, y) d(x, y) = \int_{\Lambda_\ell} \left( \int_{\Delta_{x,\ell}} u(x-y) \rho^{(2)}(x, y) dy \right. \quad (13a)$$

$$\left. + \int_{\Lambda_\ell \setminus \Delta_{x,\ell}} u(x-y) \rho^{(2)}(x, y) dy \right) dx, \quad (13b)$$

where the set  $\Delta_{x,\ell} = \{y \in \Lambda_\ell \mid u(x-y) \geq 1\}$  is bounded and  $u(x-\cdot)$  is absolutely integrable over  $\mathbb{R}^d \setminus \Delta_{x,\ell}$  because of (1). Since  $\rho^{(2)}$  is bounded according to (3) the integral in (13b) is uniformly bounded, independent of  $\ell$  and  $x \in \Lambda_\ell$ . The integrand of the integral in (13a) is nonnegative. In case this integral diverges for some  $\ell \in \mathbb{N}$  and some  $x \in \Lambda_\ell$ , then the total right-hand side of (13) equals  $+\infty$ , and this remains true for all larger values of  $\ell$ , i.e.,  $E(u, \mathbb{P}) = +\infty$ . The same argument applied to the right-hand side of (12) shows that equality holds in (12) in this case because  $\rho^{(2)}$  is translation invariant.

On the other hand, if the integral over  $\Delta_{x,\ell}$  in (13a) is finite for every  $\ell \in \mathbb{N}$  and every  $x \in \Lambda_\ell$ , then the same argument as before, together with the translation invariance of  $\rho^{(2)}$  shows that

$$\int_{\mathbb{R}^d} u(x-y) \rho^{(2)}(x, y) dy = \int_{\mathbb{R}^d} u(y) \rho^{(2)}(y, 0) dy \quad (14)$$

is absolutely convergent. Now, we assume that  $\ell$  is greater than the parameter  $r_0$  which occurs in (1). Then, we choose some  $r$  between  $r_0$  and  $\ell$ , and we split the domain  $\Lambda_\ell^2$  of integration into

$$\Lambda_\ell^2 = \Delta_1 \cup \Delta_2 \cup \Delta_3,$$

where

$$\Delta_1 = \{(x, y) \in \Lambda_\ell^2 : x \in \Lambda_{\ell-r}, |y-x| \leq r\},$$

$$\Delta_2 = \{(x, y) \in \Lambda_\ell^2 : x \in \Lambda_\ell \setminus \Lambda_{\ell-r}, |y-x| \leq r\},$$

$$\Delta_3 = \{(x, y) \in \Lambda_\ell^2 : |x-y| > r\}.$$

Under these assumptions on  $r$  and  $\ell$ , it follows from the fact that (14) is absolutely convergent that

$$\left| \int_{\Delta_2} u(x-y)\rho^{(2)}(x,y)d(x,y) \right| \leq \int_{\Lambda_\ell \setminus \Lambda_{\ell-r}} \int_{|y-x| \leq r} |u(x-y)|\rho^{(2)}(x,y)dydx \leq (|\Lambda_\ell| - |\Lambda_{\ell-r}|) \int_{\mathbb{R}^3} |u(y)|\rho^{(2)}(y,0)dy$$

and

$$\left| \int_{\Delta_3} u(x-y)\rho^{(2)}(x,y)d(x,y) \right| \leq \int_{\Delta_3} |u(x-y)|\rho^{(2)}(x,y)d(x,y) \leq \int_{\Lambda_\ell} \int_{|y-x| > r} |u(x-y)|\rho^{(2)}(x,y)dydx \leq |\Lambda_\ell| \int_{|y| > r} |u(y)|\rho^{(2)}(y,0)dy.$$

Since (14) converges absolutely, we can thus choose  $r = r(\varepsilon)$  sufficiently large to make sure that

$$\limsup_{\ell \rightarrow \infty} \left| \frac{1}{|\Lambda_\ell|} \int_{\Delta_2 \cup \Delta_3} u(x-y)\rho^{(2)}(x,y)d(x,y) \right| \leq \varepsilon \tag{15}$$

for any given positive number  $\varepsilon$ . On the other hand, using the translation invariance again, we have

$$\int_{\Delta_1} u(x-y)\rho^{(2)}(x,y)d(x,y) = \int_{\Lambda_{\ell-r}} \int_{|y-x| \leq r} u(x-y)\rho^{(2)}(x,y)dydx = |\Lambda_{\ell-r}| \int_{|y| \leq r} u(y)\rho^{(2)}(y,0)dy,$$

and hence,

$$\limsup_{\ell \rightarrow \infty} \left| \frac{1}{|\Lambda_\ell|} \int_{\Delta_1} u(x-y)\rho^{(2)}(x,y)d(x,y) - \int_{\mathbb{R}^d} u(y)\rho^{(2)}(y,0)dy \right| \leq \int_{|y| > r} |u(y)|\rho^{(2)}(y,0)dy. \tag{16}$$

Combining (16) for  $r = r(\varepsilon)$  with (15), the assertion (12) follows by letting  $\varepsilon \rightarrow 0$ . □

We mention that the Kirkwood-Salsburg equations (e.g., Ref. 42) can be used to argue that the integrand of (12) is bounded near the origin when  $\mathbb{P}$  is a  $(\beta, \mu, u)$ -Gibbs measure so that for “matching”  $u$  and  $\mathbb{P}$ , the integral is absolutely convergent and finite by virtue of (1) and (3).

#### IV. UNIQUENESS RESULTS OF HENDERSON TYPE IN THE THERMODYNAMICAL LIMIT

We now formulate Henderson’s theorem in the spirit of his original paper<sup>20</sup> and provide a rigorous proof, based on arguments borrowed from Ref. 20 and from the Proof of Theorem 2.34 in Ref. 14.

**Theorem 2.** *Let  $u, v \in \mathcal{U}$ ,  $\beta > 0$ , and  $\mu, \mu' \in \mathbb{R}$  be given, and assume that  $\mathbb{P}_u \in \mathcal{G}(\beta, \mu, u)$  and  $\mathbb{P}_v \in \mathcal{G}(\beta, \mu', v)$  admit the same density  $\rho^{(1)}$  and the same pair correlation function  $\rho^{(2)}$ . Then,  $\mu = \mu'$  and  $u = v$  almost everywhere.*

*Proof.* By Theorem A, we have

$$\Omega_{\beta, \mu}(u, \mathbb{P}_v) \leq \Omega_{\beta, \mu}(u, \mathbb{P}_u)$$

and

$$\Omega_{\beta, \mu'}(v, \mathbb{P}_u) \leq \Omega_{\beta, \mu'}(v, \mathbb{P}_v).$$

Since  $\Omega_{\beta, \mu}(u, \mathbb{P}_u)$  and  $\Omega_{\beta, \mu'}(v, \mathbb{P}_v)$  are finite, we may write these inequalities as

$$\Omega_{\beta, \mu}(u, \mathbb{P}_v) - \Omega_{\beta, \mu}(u, \mathbb{P}_u) \leq 0 \tag{17}$$

and

$$\Omega_{\beta, \mu'}(v, \mathbb{P}_u) - \Omega_{\beta, \mu'}(v, \mathbb{P}_v) \leq 0, \tag{18}$$

and adding them, we get

$$\Omega_{\beta, \mu}(u, \mathbb{P}_v) - \Omega_{\beta, \mu}(u, \mathbb{P}_u) + \Omega_{\beta, \mu'}(v, \mathbb{P}_u) - \Omega_{\beta, \mu'}(v, \mathbb{P}_v) \leq 0. \tag{19}$$

Recalling the definition (11) of the grand potential, we have

$$\begin{aligned} \Omega_{\beta,\mu}(u, \mathbb{P}_v) - \Omega_{\beta,\mu}(u, \mathbb{P}_u) + \Omega_{\beta,\mu'}(v, \mathbb{P}_u) - \Omega_{\beta,\mu'}(v, \mathbb{P}_v) &= \mu\rho(\mathbb{P}_v) - E(u, \mathbb{P}_v) - \frac{1}{\beta}S(\mathbb{P}_v) - \mu\rho(\mathbb{P}_u) + E(u, \mathbb{P}_u) + \frac{1}{\beta}S(\mathbb{P}_u) \\ &\quad + \mu'\rho(\mathbb{P}_u) - E(v, \mathbb{P}_u) - \frac{1}{\beta}S(\mathbb{P}_u) - \mu'\rho(\mathbb{P}_v) + E(v, \mathbb{P}_v) + \frac{1}{\beta}S(\mathbb{P}_v) \\ &= -E(u, \mathbb{P}_v) + E(u, \mathbb{P}_u) - E(v, \mathbb{P}_u) + E(v, \mathbb{P}_v) \end{aligned} \quad (20)$$

because  $\rho(\mathbb{P}_u) = \rho(\mathbb{P}_v) = \rho^{(1)}$  by assumption. Furthermore, by virtue of Proposition 1 and the fact that the pair correlation functions of  $\mathbb{P}_u$  and  $\mathbb{P}_v$  coincide, there holds

$$E(u, \mathbb{P}_u) = \frac{1}{2} \int_{\mathbb{R}^d} u(x) \rho^{(2)}(x, 0) dx = E(u, \mathbb{P}_v)$$

and

$$E(v, \mathbb{P}_u) = \frac{1}{2} \int_{\mathbb{R}^d} v(x) \rho^{(2)}(x, 0) dx = E(v, \mathbb{P}_v).$$

Inserting this into (20), we conclude that

$$\Omega_{\beta,\mu}(u, \mathbb{P}_v) - \Omega_{\beta,\mu}(u, \mathbb{P}_u) + \Omega_{\beta,\mu'}(v, \mathbb{P}_u) - \Omega_{\beta,\mu'}(v, \mathbb{P}_v) = 0.$$

Accordingly, equality holds in (19), and thus necessarily in both (17) and (18). By the Gibbs variational principle (Theorem A), this implies that  $\mathbb{P}_v \in \mathcal{G}(\beta, \mu, u)$  and  $\mathbb{P}_u \in \mathcal{G}(\beta, \mu', v)$ .

It therefore follows from (5) that

$$\begin{aligned} \int_{\Gamma} F(\gamma) d\mathbb{P}_u(\gamma) &= \sum_{N=0}^{\infty} \frac{1}{N!} \int_{\Delta^N} \left( \int_{\Gamma(\Delta^c)} F(\gamma') e^{-\beta W_u(\mathbf{x}_N; \gamma)} d\mathbb{P}_u(\gamma) \right) e^{N\beta\mu - \beta H_u(\mathbf{x}_N)} d\mathbf{x}_N \\ &= \sum_{N=0}^{\infty} \frac{1}{N!} \int_{\Delta^N} \left( \int_{\Gamma(\Delta^c)} F(\gamma') e^{-\beta W_v(\mathbf{x}_N; \gamma)} d\mathbb{P}_u(\gamma) \right) e^{N\beta\mu' - \beta H_v(\mathbf{x}_N)} d\mathbf{x}_N \end{aligned}$$

for every  $F \in L^1(\mathbb{P}_u)$  and every bounded domain  $\Delta \subset \mathbb{R}^d$ . Therefore,

$$H_u(\mathbf{x}_N) + W_u(\mathbf{x}_N; \gamma) - N\mu = H_v(\mathbf{x}_N) + W_v(\mathbf{x}_N; \gamma) - N\mu' \quad (21)$$

for every  $N \in \mathbb{N}$ , almost every  $\mathbf{x}_N \in \Delta^N$  and  $\mathbb{P}_u$  almost surely for  $\gamma \in \Gamma(\Delta^c)$ . For  $N = 1$ , this means that

$$W_u(x; \gamma) - \mu = W_v(x; \gamma) - \mu' \quad (22)$$

for almost every  $x \in \Delta$  and  $\mathbb{P}_u$  and  $\mathbb{P}_v$  almost surely for  $\gamma \in \Gamma(\Delta^c)$ . Using the additivity of  $W_u$  and  $W_v$  in the first argument, we thus conclude from the case  $N = 2$  of (21) that

$$u(x - y) = v(x - y) \quad (23)$$

for almost every  $x, y \in \Delta$ , and since  $\Delta$  was arbitrarily chosen, we have  $u = v$  almost everywhere.

Inserting (23) into (6), it follows that

$$W_u(x; \gamma_0) = W_v(x; \gamma_0)$$

for almost every  $x \in \Delta$  and almost every finite subset  $\gamma_0 \subset \Delta^c \cap \Lambda_\ell$ . Together with (8), this implies that for every  $\mathbb{P} \in \mathcal{G}$ , there holds

$$W_u(x; \gamma \cap \Lambda_\ell) = W_v(x; \gamma \cap \Lambda_\ell) \quad (24)$$

$\mathbb{P}$  almost surely for  $\gamma \in \Gamma(\Delta^c)$ . By virtue of (9) and (24), we therefore have

$$W_u(x; \gamma) = W_v(x; \gamma)$$

for almost every  $x \in \Delta$  and  $\mathbb{P}_u$  almost surely for  $\gamma \in \Gamma(\Delta^c)$ , and hence we conclude from (22) that  $\mu = \mu'$ , which remained to be shown.  $\square$

Henderson stipulated the assumptions of Theorem 2 from the canonical ensemble point of view. Concerning the grand canonical perspective, an analogous uniqueness result is as follows.

**Theorem 3.** Let  $u, v \in \mathcal{U}$ ,  $\beta > 0$ , and  $\mu \in \mathbb{R}$  be given, and assume that  $\mathbb{P}_u \in \mathcal{G}(\beta, \mu, u)$  and  $\mathbb{P}_v \in \mathcal{G}(\beta, \mu, v)$  admit the same pair correlation function  $\rho^{(2)}$ . Then,  $u = v$  almost everywhere.

The proof is the same as for Theorem 2. This time (20) holds true because the chemical potentials are the same. We mention, however, that we do not know whether the counting densities of the two Gibbs measures are necessarily the same, unless it is assumed that the corresponding  $(\beta, \mu, u)$ -Gibbs measure is uniquely determined—as it is, e.g., in the gas phase.

## V. CONCLUDING REMARKS

We emphasize that for our results, we do not stipulate that the system is in the gas phase nor that the set  $\mathcal{G}(\beta, \mu, u)$  consists of a single Gibbs measure only.

In case it is known that  $u$  is also radially symmetric, i.e., if the interaction of two particles only depends on their distance, then one can show—using the Markov-Kakutani fixed point theorem as in the Proof of Theorem 5.8 in Ref. 41, compare Kuna<sup>25</sup>—that there exists at least one rotation and translation invariant Gibbs measure  $\mathbb{P}_u \in \mathcal{G}(\beta, \mu, u)$ , which can be used to define a *radial distribution function*  $g : \mathbb{R}^+ \rightarrow \mathbb{R}_0^+$  given by

$$g(r) = \frac{\rho^{(2)}(x_1, x_2)}{(\rho^{(1)})^2}, \quad r = |x_1 - x_2|, \quad (25)$$

provided that the density is nonzero. Assuming further that  $\mathbb{P}_v \in \mathcal{G}(\beta, \mu, v)$  is also rotation invariant, then one obviously can impose in Theorem 2—as did Henderson—that the radial distribution functions and the densities are the same for these two Gibbs measures, and the statement of the theorem remains valid. We do not know whether in the formulation of Theorem 3  $\rho^{(2)}$  can also be replaced by the radial distribution function in the isotropic case.

Finally, we remark that the representation (12) of the specific energy is not essential for the uniqueness argument. By its definition (10), the specific energy  $E(u, \mathbb{P})$  is the limit in  $\mathbb{R} \cup \{+\infty\}$  of the expression given in (13) normalized by two times the volume of  $\Lambda_\ell$ , and hence, its value only depends on  $u$  and on the pair correlation function  $\rho^{(2)}$  associated with  $\mathbb{P}$ . This suffices to conclude that the expression (20) sums up to zero.

## VI. SUMMARY

Henderson's theorem states that the pair potential of an isotropic system of identical classical particles in equilibrium without multibody interactions is uniquely determined from the finite-volume pair correlation function of the system and its temperature and density. In this note, we have provided sufficient conditions under which a rigorous proof of Henderson's theorem is possible in the thermodynamical limit. We have also shown that the same result is true if the chemical potential instead of the density is given.

We emphasize that when higher-order interactions are relevant, pairwise potentials are only crude approximations of the ideal multibody potential of mean force and will often fail to be thermodynamically useful. Chayes and Chayes<sup>4</sup> and Navrotskaya<sup>33</sup> have used variational principles to show that multibody interactions are also uniquely determined by the corresponding multibody statistics in continuous space for the canonical and grand canonical ensembles in finite volumes; see also the work of Tóth.<sup>47</sup> For translation invariant stable interactions with finite range, it is possible to extend these results to the thermodynamical limit of these systems in a similar manner as in the proof of Theorem 2 by using a version of the Gibbs variational principle established by Dereudre.<sup>6</sup> As particular examples for such interactions, Dereudre refers to Delaunay triangle interactions or the Widom-Rowlinson interaction. The corresponding problem on the lattice was again studied in Ref. 17.

Having settled the uniqueness problem, the natural follow-up question concerns the existence of solutions of the inverse Henderson problem, i.e., what are necessary and sufficient conditions on a given triplet  $\beta, \rho > 0$ ,  $\mu \in \mathbb{R}$ , and a nonnegative translation invariant function  $\rho^{(2)} : (\mathbb{R}^d)^2 \rightarrow \mathbb{R}$ , such that there exists a pair potential  $u \in \mathcal{U}$  for which  $\rho$  is the density and  $\rho^{(2)}$  is the pair correlation function of a Gibbs measure  $\mathbb{P} \in \mathcal{G}(\beta, \mu, u)$ . Partial results for this problem have been contributed, e.g., by Caglioti, Kuna, Lebowitz, and Speer,<sup>2,26,27</sup> and Koralov.<sup>24</sup> The general existence problem is widely open, though.

## ACKNOWLEDGMENTS

We are grateful to Friederike Schmid for helpful suggestions on the presentation of these results. This research has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project No. 233630050—TRR 146.

## REFERENCES

- <sup>1</sup>E. Brini, E. A. Algaer, P. Ganguly, C. Li, F. Rodríguez-Ropero, and N. F. A. van der Vegt, "Systematic coarse-graining methods for soft matter simulations – a review," *Soft Matter* **9**, 2108–2119 (2013).
- <sup>2</sup>E. Caglioti, T. Kuna, J. L. Lebowitz, and E. R. Speer, "Point processes with specified low order correlations," *Markov Process. Relat. Fields* **12**, 257–272 (2006).
- <sup>3</sup>S. P. Carmichael and M. S. Shell, "A new multiscale algorithm and its application to coarse-grained peptide models for self-assembly," *J. Phys. Chem. B* **116**, 8383–8393 (2012).

- <sup>4</sup>J. T. Chayes and L. Chayes, "On the validity of the inverse conjecture in classical density functional theory," *J. Stat. Phys.* **36**, 471–488 (1984).
- <sup>5</sup>J. T. Chayes, L. Chayes, and E. H. Lieb, "The inverse problem in classical statistical mechanics," *Commun. Math. Phys.* **93**, 57–121 (1984).
- <sup>6</sup>D. Dereudre, "Variational principle for Gibbs point processes with finite range interaction," *Electron. Commun. Probab.* **21**, 10–21 (2016).
- <sup>7</sup>L. Delle Site, M. Deserno, B. Dünweg, C. Holm, C. Peter, and H. Pleiner, "Modern simulation approaches in soft matter science: From fundamental understanding to industrial applications," *Eur. Phys. J. Spec. Top.* **225**, 1317–1321 (2016).
- <sup>8</sup>R. Dobrushin, "Investigation of conditions for the asymptotic existence of the configuration integral of Gibbs' distribution," *Theor. Probab. Appl.* **9**, 566–581 (1964).
- <sup>9</sup>F. Delbary, M. Hanke, and D. Ivanizki, "A generalized Newton iteration for computing the solution of the inverse Henderson problem," e-print [arXiv:1806.11135](https://arxiv.org/abs/1806.11135).
- <sup>10</sup>G. Gallavotti, *Statistical Mechanics: A Short Treatise* (Springer, Berlin, 1999).
- <sup>11</sup>J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 4th ed. (Academic Press, Oxford, 2013).
- <sup>12</sup>H.-O. Georgii, "Large deviations and the equivalence of ensembles for Gibbsian particle systems with superstable interaction," *Probab. Theory Relat. Fields* **99**, 171–195 (1994).
- <sup>13</sup>H.-O. Georgii, "The equivalence of ensembles for classical systems of particles," *J. Stat. Phys.* **80**, 1341–1378 (1995).
- <sup>14</sup>H.-O. Georgii, *Gibbs Measures and Phase Transitions*, 2nd ed. (De Gruyter, Berlin, 2011).
- <sup>15</sup>H.-O. Georgii and H. Zessin, "Large deviations and the maximum entropy principle for marked point random fields," *Probab. Theory Relat. Fields* **96**, 177–204 (1993).
- <sup>16</sup>J. Gibbs, *Elementary Principles in Statistical Mechanics* (Scribner's Sons, New York, 1902).
- <sup>17</sup>R. Griffiths and D. Ruelle, "Strict convexity ("continuity") of the pressure in lattice systems," *Commun. Math. Phys.* **23**, 169–175 (1971).
- <sup>18</sup>M. Hanke, *A Taste of Inverse Problems. Basic Theory and Examples* (SIAM, Philadelphia, 2017).
- <sup>19</sup>M. Heinen, "Calculating particle pair potentials from fluid-state pair correlations: Iterative Ornstein-Zernike inversion," *J. Comput. Chem.* **39**, 1531–1543 (2018).
- <sup>20</sup>R. Henderson, "A uniqueness theorem for fluid pair correlation functions," *Phys. Lett. A* **49**, 197–198 (1974).
- <sup>21</sup>P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.* **136**, B864–B871 (1964).
- <sup>22</sup>N. M. Hugenholtz, "On the inverse problem in statistical mechanics," *Commun. Math. Phys.* **85**, 27–38 (1982).
- <sup>23</sup>Y. G. Kondratiev and T. Kuna, "Correlation functionals for Gibbs measures and Ruelle bounds," *Methods Funct. Anal. Topol.* **9**, 9–58 (2003).
- <sup>24</sup>L. Korolov, "An inverse problem for Gibbs fields with hard core potential," *J. Math. Phys.* **48**, 053301 (2007).
- <sup>25</sup>T. Kuna, "Studies in configuration space analysis and applications," Ph.D. thesis, Rheinische Friedrich-Wilhelms-Universität, Bonn, 1999.
- <sup>26</sup>T. Kuna, J. L. Lebowitz, and E. R. Speer, "Realizability of point processes," *J. Stat. Phys.* **129**, 417–439 (2007).
- <sup>27</sup>T. Kuna, J. L. Lebowitz, and E. R. Speer, "Necessary and sufficient conditions for realizability of point processes," *Ann. Appl. Probab.* **21**, 1253–1281 (2011).
- <sup>28</sup>D. Levesque, J. J. Weis, and L. Reatto, "Pair interaction from structural data for dense classical liquids," *Phys. Rev. Lett.* **54**, 451–454 (1985).
- <sup>29</sup>A. P. Lyubartsev and A. Laaksonen, "Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach," *Phys. Rev. E* **52**, 3730–3737 (1995).
- <sup>30</sup>A. P. Lyubartsev, A. Naomé, D. P. Vercauteren, and A. Laaksonen, "Systematic hierarchical coarse-graining with the inverse Monte Carlo method," *J. Chem. Phys.* **143**, 243120 (2015).
- <sup>31</sup>S. Y. Mashayak, L. Miao, and N. R. Aluru, "Integral equation theory based direct and accelerated systematic coarse-graining approaches," *J. Chem. Phys.* **148**, 214105 (2018).
- <sup>32</sup>N. Mermin, "Thermal properties of the inhomogeneous electron gas," *Phys. Rev.* **137**, A1471–A1473 (1965).
- <sup>33</sup>I. Navrotskaya, "Inverse problem in classical statistical mechanics," Ph.D. thesis, University of Pittsburgh, 2016.
- <sup>34</sup>W. G. Noid, "Perspective: Coarse-grained models for biomolecular systems," *J. Chem. Phys.* **139**, 090901 (2013).
- <sup>35</sup>C. Peter and K. Kremer, "Multiscale simulation of soft matter systems," *Faraday Discuss.* **144**, 9–24 (2010).
- <sup>36</sup>B. L. Peters, K. M. Salerno, A. Agrawal, D. Perahia, and G. S. Grest, "Coarse-grained modeling of polyethylene melts: Effect on dynamics," *J. Chem. Theor. Comput.* **13**, 2890–2896 (2017).
- <sup>37</sup>R. Potestio, C. Peter, and K. Kremer, "Computer simulations of soft matter: Linking the scales," *Entropy* **16**, 4199–4245 (2014).
- <sup>38</sup>M. Praprotnik, L. Delle Site, and K. Kremer, "Adaptive resolution molecular-dynamics simulation: Changing the degrees of freedom on the fly," *J. Chem. Phys.* **123**, 224106 (2005).
- <sup>39</sup>D. Reith, M. Pütz, and F. Müller-Plathe, "Deriving effective mesoscale potentials from atomistic simulations," *J. Comput. Chem.* **24**, 1624–1636 (2003).
- <sup>40</sup>D. W. Robinson and D. Ruelle, "Mean entropy of states in classical statistical mechanics," *Commun. Math. Phys.* **5**, 288–300 (1967).
- <sup>41</sup>D. Ruelle, *Statistical Mechanics: Rigorous Results* (W. A. Benjamin Publisher, New York, 1969).
- <sup>42</sup>D. Ruelle, "Superstable interactions in classical statistical mechanics," *Commun. Math. Phys.* **18**, 127–159 (1970).
- <sup>43</sup>W. Schommers, "A pair potential for liquid rubidium from the pair correlation function," *Phys. Lett. A* **43**, 157–158 (1973).
- <sup>44</sup>W. Schommers, "Pair potentials in disordered many-particle systems: A study for liquid gallium," *Phys. Rev. A* **28**, 3599–3605 (1983).
- <sup>45</sup>M. S. Shell, "The relative entropy is fundamental to multiscale and inverse thermodynamic problems," *J. Chem. Phys.* **129**, 144108 (2008).
- <sup>46</sup>A. K. Soper, "Empirical potential Monte Carlo simulation of fluid structure," *Chem. Phys.* **202**, 295–306 (1996).
- <sup>47</sup>G. Tóth, "Interactions from diffraction data: Historical and comprehensive overview of simulation assisted methods," *J. Phys.: Condens. Matter* **19**, 335220 (2007).
- <sup>48</sup>A. Tsourtis, V. Harmandaris, and D. Tsagkarogiannis, "Parameterization of coarse-grained molecular interactions through potential of mean force calculations and cluster expansion techniques," *Entropy* **19**, 395 (2017).

# 4

## A variational framework for the inverse Henderson problem of statistical mechanics

### 4.1 Overview

In this chapter we investigate a constructive approach for the inverse Henderson problem. For this, we employ the specific relative entropy

$$S(P_* | \beta, \mu, u) = \frac{1}{\beta} S(P_*) + p(\beta, \mu, u) - \mu \rho_* + \frac{1}{2} \int_{\mathbb{R}^d} u(x) \rho_*^{(2)}(0, x) dx$$

of a fixed point process  $P_*$  (with density  $\rho_*$  and pair correlation function  $\rho_*^{(2)}$ ) with respect to  $(\beta, \mu, u)$ -Gibbs measures. Since we work at fixed  $\beta > 0$  we will write  $S(P_* | \mu, u)$  in place of  $S(P_* | \beta, \mu, u)$ . If there is some set of parameters  $(\beta, \mu_*, u_*) \in [0, \infty) \times \mathbb{R} \times \mathcal{U}_0$  such that  $P_* \in \mathcal{G}(\beta, \mu_*, u_*)$  then by Theorem 2.2.26 we have that  $S(P_* | \mu, u) \geq 0$  with equality if and only if  $\mu = \mu_*$  and  $u = u_*$ . This means solving the inverse Henderson problem is equivalent to minimizing the specific relative entropy for a fixed Gibbs measure.

In light of Proposition 2.2.11 the relative entropy is a convex function as the limit of the sum of a strictly convex and linear functions. However, it can be shown that the strict convexity of the pressure remains in the thermodynamic limit.

**Theorem 4.1.1** *The pressure  $p$  is a strictly convex function of  $(\mu, u)$ .*

As such the relative entropy functional is also strictly convex and we can use the framework of convex optimization for the minimization. However, since most physical or chemical simulations are done in a canonical ensemble, see Remark 2.2.8, it is natural to want to work at fixed density  $\rho_*$ . For this we need to investigate whether there is a chemical potential  $\mu = \mu(u, \rho_*)$  such that there is a  $(\beta, \mu, u)$ -Gibbs measure with this given density  $\rho_*$ .

**Theorem 4.1.2** *For any  $u \in \mathcal{U}_0$  and  $\beta, \rho_* > 0$  there is a unique  $\mu = \mu(u, \rho_*)$  and a  $(\beta, \mu, u)$ -Gibbs measure with density  $\rho_*$ .*

The mapping  $\mu: \mathcal{U}_0 \rightarrow \mathbb{R}, u \mapsto \mu(u, \rho_*)$  is highly nonlinear. In fact, the pressure at fixed density will no longer be strictly convex. The relative entropy though, retains strict convexity.

**Theorem 4.1.3** *The fixed density relative entropy functional  $\Phi: \mathcal{U}_0 \rightarrow \mathbb{R}$  defined by*

$$\Phi(u) = p(\beta, \mu(u, \rho_*), u) - \mu(u, \rho_*)\rho_* + \frac{1}{2} \int_{\mathbb{R}^d} u(x)\rho_*^{(2)}(0, x) dx \quad (4.1)$$

*is strictly convex in  $u$  and its value is minimal if and only if  $u = u_*$ .*

In light of this result a natural choice for a numerical algorithm for the minimization of  $\Phi$  would be Newton's method. For this we need to show differentiability of  $\Phi$  and its derivative and thus work on the space  $\mathcal{V}_{u_*}$ .

**Theorem 4.1.4** *The fixed density relative entropy functional  $\Phi$  is twice Fréchet differentiable in an open set  $\mathcal{B}(u_*) \subset \mathcal{V}_{u_*}$  around  $u_*$ . For every  $u \in \mathcal{B}(u_*)$  the first derivative  $\Phi' \in \mathcal{V}'_{u_*}$  is given by*

$$\Phi'(u)v = \int_{\mathbb{R}^d} v(x) \left( \rho_*^{(2)}(x) - \rho_u^{(2)}(x) \right) dx$$

*for  $v \in \mathcal{V}_{u_*}$  and the second derivative  $\Phi'': \mathcal{V}_{u_*} \times \mathcal{V}_{u_*} \rightarrow \mathbb{R}$  is a continuous symmetric and positive semidefinite bilinear form. Here  $\rho_u^{(2)}$  denotes the two-point correlation functions of the (unique)  $(\beta, \mu(u, \rho_*), u)$ -Gibbs measure.*

**Remark 4.1.5** *Another interpretation of  $\Phi''$  is as a covariance operator on the space of perturbations. There holds:*

$$\Phi''(u)(v, v) = \frac{\beta}{2} \lim_{\Lambda \nearrow \mathbb{R}^d} \frac{1}{\mathfrak{K}(\Lambda)} \left( \text{Var}_u(V(\cdot_\Lambda)) - \frac{\text{Cov}_u(V(\cdot_\Lambda), N_\Lambda)^2}{\text{Var}_u(N_\Lambda)} \right) \quad (4.2)$$

*where  $V$  is the Hamiltonian associated to  $v \in \mathcal{V}_{u_*}$  by (2.16) and the (co-)variances are taken with respect to  $\mathbb{P}_u$ .*

## 4.2 Reprint

The article is reprinted by permission from Springer Nature: Letters in Mathematical Physics: A variational framework for the inverse Henderson problem of statistical mechanics, Fabio Frommer, Martin Hanke, 2022 , 112, 17 July 2022 Springer Science+Business Media (doi.org/10.1007/s11005-022-01563-w).

Letters in Mathematical Physics (2022) 112:71  
https://doi.org/10.1007/s11005-022-01563-w



## A variational framework for the inverse Henderson problem of statistical mechanics

Fabio Frommer<sup>1</sup> · Martin Hanke<sup>1</sup>

Received: 20 April 2022 / Revised: 8 June 2022 / Accepted: 27 June 2022 /

Published online: 17 July 2022

© The Author(s) 2022, corrected publication 2022

### Abstract

The inverse Henderson problem refers to the determination of the pair potential which specifies the interactions in an ensemble of classical particles in continuous space, given the density and the equilibrium pair correlation function of these particles as data. For a canonical ensemble in a bounded domain, it has been observed that this pair potential minimizes a corresponding convex relative entropy functional, and that the Newton iteration for minimizing this functional coincides with the so-called inverse Monte Carlo (IMC) iterative scheme. In this paper, we show that in the thermodynamic limit analogous connections exist between the specific relative entropy introduced by Georgii and Zessin and a proper formulation of the IMC iteration in the full space. This provides a rigorous variational framework for the inverse Henderson problem, valid within a large class of pair potentials, including, for example, Lennard-Jones-type potentials. It is further shown that the pressure is strictly convex as a function of the pair potential and the chemical potential, and that the specific relative entropy at fixed density is a strictly convex function of the pair potential. At a given reference potential and a corresponding density in the gas phase, we determine the gradient and the Hessian of the specific relative entropy, and we prove that the Hessian extends to a symmetric positive semidefinite quadratic functional in the space of square integrable perturbations of this potential.

---

This article belongs to the themed collection: Mathematical Physics and Numerical Simulation of Many-Particle Systems; V. Bach and L. Delle Site (eds.). The research leading to this work has been done within the Collaborative Research Center TRR 146; corresponding funding by the DFG is gratefully acknowledged.

---

✉ Martin Hanke  
hanke@math.uni-mainz.de

Fabio Frommer  
fabiofrommer@uni-mainz.de

<sup>1</sup> Institut für Mathematik, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany

**Keywords** Coarse-graining · Relative entropy · Pair correlation function · Cluster function · Pressure · Fréchet derivative · Gibbs point processes

**Mathematics Subject Classification** 82B21 · 82B80 · 60G55

## 1 Introduction

Numerical simulation has established itself as an independent and indispensable branch of research in the natural sciences, on equal footing with theory and experiment. To be truly useful, numerical approaches have to face and master the multiscale nature which is ubiquitous in almost all relevant applications. This is particularly true for soft matter, where spatial scales may bridge from the electron scale up to the millimeter scale of biomaterials or polymers. Concerning examples we refer to the excellent survey by Noid [31], the collected volume edited by Monticelli and Salonen [28], or the recent special issue [40] of *Journal of Physics: Condensed Matter*.

An important technique to advance numerical algorithms to the particular needs of multiscale applications consists in *coarse-graining*, cf., e.g., Noid [32], or Peter and Kremer [33]: small-scale features are discarded on the coarser scale by replacing detailed descriptions of molecules or matter by artificial *beads* of a certain shape. The simulation then focusses only on these beads and their interaction with the other constituents of the system. When and where necessary, fine details can be reinserted back into the simulation for better accuracy as, e.g., in the AdResS scheme ([6, 34]).

Of course, to evaluate the equations of motion for the coarse-grained model, it is necessary to derive the prevailing effective forces on these beads. Concerning the transition from an atomic microscale to a molecular macroscale in thermodynamic equilibrium, for example, there are essentially two ways to settle this problem. On the one hand, one can employ an *ab initio* bottom-up approach and evaluate and assemble the resulting forces from the eliminated details of the fine-grained description, cf. Ercolessi and Adams [5], or run a fine-grained simulation, compute the corresponding forces, and somehow approximate them on the coarse level as suggested, e.g., by Izvekov and Voth [20], and Wang et al [46]. The other alternative is to follow a top-down strategy and use the given structural information about the location of the beads on the coarse scale to formulate an *inverse problem*: Which are the appropriate forces or interactions on the coarse level that define ensembles with the same structural properties?

In this work, we treat one of the simplest incarnations of the latter approach. Let us presume that the probabilities of the snapshots of the coarse-grained bead ensembles are in good agreement with a model which only uses additive pairwise translation invariant interactions of the beads. Such interactions can be formulated in terms of a scalar pair potential which only depends on the relative position of the respective pair of beads. The pair correlation function, which measures the empirical likelihood to observe two beads at a given relative position, appears to be an adequate piece of data to be used for finding the corresponding potential, because both the data and the unknown consist of a scalar function of the space variable in that case. In fact, in a celebrated paper, Henderson [18] argued that the pair potential is uniquely specified

this way, i.e., under given conditions of temperature and density, no two different pair potentials can give rise to the same piece of empirical data. Finding a pair potential that reproduces the given pair correlation function is therefore sometimes called the *inverse Henderson problem*.

The numerical solution of the inverse Henderson problem is demanding because of the lack of a mathematical formula for computing the pair correlation function for a given pair potential, or vice versa. In the old days people have developed approximate identities like the Percus–Yevick or hypernetted chain integral equations for this purpose, cf. Hansen and McDonald [17], or have used parametric ansatz functions like, e.g., Lennard-Jones potentials, and optimized the corresponding parameters numerically. Today, the state of the art is to use nonparametric potentials and employ iterative schemes which, in each iteration, simulate the equilibrium ensemble for the current guess of the pair potential, and use the associated data fit to somehow generate a new guess. Well-known examples are the iterative Boltzmann inversion (IBI), cf. Soper [42] and Reith, Pütz, and Müller-Plathe [35], and the inverse Monte Carlo method (IMC) by Lyubartsev and Laaksonen [26]. We recommend the valuable reviews by Toth [44] and Rühle et al [37] for a comparison of these and further methods.

We emphasize that the setting of the inverse Henderson problem is generally accepted to be far too simplistic to capture all the relevant features of a real system, cf., e.g., [25, 32, 45], mostly because multibody interactions are neglected. In particular, thermodynamic properties of the coarse-grained model may differ from the real system, especially at other temperatures or densities.

But its simplicity offers a great opportunity for a mathematical analysis, which in turn may lead to a better understanding of other, more flexible coarse-graining techniques that are routinely being employed in practice. Still, only few rigorous mathematical results have yet been obtained, e.g., in [2, 7, 16, 23, 24, 30], the reason being, again, the lack of explicit formulae to attack the problem.

The aim of this paper is to point out and advocate an alternative access point for theoretical investigations, which goes back to a nice observation by Shell [41] from within the chemical physics community: He argues that the Henderson potential minimizes the (information theoretic) relative entropy

$$S_{\text{rel}} = \sum_{\gamma} \mathcal{P}_*(\gamma) \log \frac{\mathcal{P}_*(\gamma)}{\mathcal{P}(\gamma)}, \quad (1.1)$$

where—in his words—the summation is over all possible (coarse grained) ensemble configurations  $\gamma$ ,  $\mathcal{P}_*(\gamma)$  denotes the target (or observed) probability of  $\gamma$ , and  $\mathcal{P}(\gamma)$  is the corresponding probability of a model with a given pair potential. (Compare, for example, Georgii [12] for background on the concept of relative entropy in stochastics.) Subsequently, Murtola, Karttunen, and Vattulainen [29] pointed out that the functional (1.1) is convex, and that the Newton iteration for minimizing the relative entropy coincides with the aforementioned IMC iteration (see also Rosenberger et al [36]).

Like Henderson’s paper [18], the results in [29, 36, 41] lack mathematical rigor, because their arguments are restricted to bounded domains, whereas an unambigu-

ous definition of a “translation invariant ensemble” is only possible in the full space. Concerning the Henderson theorem, this shortcoming has recently been fixed in [7], building on fundamental work by Ruelle and by Georgii. Here, we focus on the proposal by Shell and his colleagues, provide a rigorous justification of their results, and elaborate further on them.

To be specific, we outline in Sect. 2 that the (appropriately formulated) relative entropy (1.1) divided by the volume of the bounded domain converges in the thermodynamic limit to the specific relative entropy, first introduced for continuum systems by Gallavotti and Miracle-Sole [8] in the case of hard-core interactions, and further investigated by Georgii and Zessin in [9, 10, 13] for general additive pair interactions. Under mild assumptions on the model and target ensembles to be utilized, we prove that this specific relative entropy is a strictly convex functional on the corresponding set of pair potentials (which include Lennard-Jones type and hard-core potentials), and that its (unique) minimizer is the particular potential which solves the inverse Henderson problem (see Sects. 3 and 4).

From Sect. 5 onward, we restrict ourselves to low densities (the “gas phase”), where the specific relative entropy is a differentiable function of the pair potential. We calculate its Hessian, and in Sect. 6, we verify that the Newton iteration for minimizing the specific relative entropy does indeed coincide with the IMC iteration formulated in the thermodynamic limit. In Sect. 7, we investigate the Hessian in more detail and show that it can be represented by a self-adjoint positive semidefinite operator in  $L^2$ . The mapping properties of this operator can be analyzed somewhat further for the particular class of Lennard-Jones-type pair potentials; we conclude with a corresponding result in Sect. 8.

We hope that this variational framework for the inverse Henderson problem opens a possibility to discuss the convergence of the IMC iteration, or to come up with measures to stabilize or regularize this popular iterative scheme.

## 2 The relative entropy in the thermodynamic limit

To reformulate Shell’s approach within a rigorous mathematical framework we start with the assumption that the target ensemble is given by a translation invariant probability measure  $\mathbb{P}_*$  on the configuration space

$$\Gamma = \{ \gamma \subset \mathbb{R}^d : \Delta \subset \mathbb{R}^d \text{ bounded} \Rightarrow \#(\gamma \cap \Delta) < \infty \}$$

with density  $\rho_*$  and finite locally second moments, compare Georgii [9]. For the model ensemble, we restrict ourselves to ensembles of classical particles with additive pairwise interactions defined by a measurable even pair potential  $u : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ . Concerning the latter, we assume that there exists  $r_0 > 0$  and decreasing positive functions  $\varphi : (0, r_0) \rightarrow \mathbb{R}_0^+$  and  $\psi : [0, \infty) \rightarrow \mathbb{R}^+$  with

$$\int_0^{r_0} r^{d-1} \varphi(r) dr = +\infty \quad \text{and} \quad \int_0^\infty r^{d-1} \psi(r) dr < \infty, \quad (2.1)$$

such that

$$\begin{aligned} u(x) &\geq \varphi(|x|) && \text{for } 0 < |x| < r_0, \\ |u(x)| &\leq \psi(|x|) && \text{for } |x| \geq r_0, \end{aligned} \quad (2.2)$$

holds true almost everywhere; for our convenience, we take  $\psi$  to be a bounded and decreasing function defined for all  $r \geq 0$ . We denote by  $\mathcal{U}_0$  the subset of the above pair potentials, which also belong to  $L_{\text{loc}}^\infty(\mathbb{R}^3 \setminus \{0\})$ , and define  $\mathcal{U}$  to be the union of  $\mathcal{U}_0$  with the hard-core potentials, which satisfy (2.2) with  $\psi$  as above and  $\varphi$  replaced by  $+\infty$ ; accordingly,  $r_0$  is taken to be the hard-core radius in this case. Technically, we do not distinguish between potentials (and functions in general) which differ on sets of Lebesgue measure zero.

**Remark 2.1** The set  $\mathcal{U}$  is convex. To see this let  $u_i \in \mathcal{U}, i = 1, 2$ , be such that (2.1) and (2.2) hold for  $r_{0,i} > 0$  and decreasing functions  $\varphi_i$  and  $\psi_i$  satisfying (2.1), respectively. Without loss of generality, we may assume that  $r_{0,1} \leq r_{0,2}$ . Let  $u = tu_1 + (1-t)u_2$  for some fixed  $t \in (0, 1)$ . We distinguish two cases. If  $u_2$  is a hard-core potential, then  $u$  is also a hard-core potential with hard-core radius  $r_{0,2}$ . In this case, we can choose  $r_0 = r_{0,2}$ ,

$$\begin{aligned} \varphi(r) &= +\infty && \text{for } 0 < r < r_{0,2}, \\ \psi(r) &= t\psi_1(r) + (1-t)\psi_2(r) && \text{for } r \geq 0, \end{aligned}$$

to achieve the assumption (2.2) for  $u$ . On the other hand, if  $u_2$  is no hard-core potential, then  $u_2$  is bounded on  $[r_{0,1}, r_{0,2})$ , and there exists  $c \geq 1$ , such that

$$|u_2(x)| \leq c\psi_2(|x|) \quad \text{for } r_{0,1} \leq |x| < r_{0,2}.$$

It follows that  $u$  satisfies the inequalities (2.2) with  $r_0 = r_{0,1}$  and

$$\begin{aligned} \varphi(r) &= t\varphi_1(r) + (1-t)\varphi_2(r) && \text{for } 0 < r < r_{0,1}, \\ \psi(r) &= t\psi_1(r) + (1-t)c\psi_2(r) && \text{for } r \geq 0. \end{aligned}$$

In either case, we have verified that  $u \in \mathcal{U}$ , hence  $\mathcal{U}$  is convex.  $\diamond$

Let  $\Lambda = [-\ell, \ell]^d$  be a bounded box in  $\mathbb{R}^d$ . In analogy to Shell, who considered the relative entropy framework for canonical ensembles in  $\Lambda$ , we take the restriction  $\mathbb{P}_{*,\Lambda}$  of  $\mathbb{P}_*$  as target, and the grand canonical ensemble in  $\Lambda$  with chemical potential  $\mu \in \mathbb{R}$  and pair potential  $u \in \mathcal{U}$  as model. For a configuration  $\gamma \in \Gamma$  of  $N = N(\gamma)$  particles located at  $\{x_1, \dots, x_N\} \subset \Lambda$ , the probability density of the model is thus given by

$$\mathcal{P}(\gamma) = \frac{1}{\Xi_\Lambda} e^{\beta\mu N(\gamma)} e^{-\beta U(\gamma)},$$

where  $\beta > 0$  is the inverse temperature,

$$U(\gamma) = U(x_1, \dots, x_N) = \sum_{1 \leq i < j \leq N} u(x_i - x_j) \quad (2.3)$$

is the interaction energy, and

$$\Xi_\Lambda = \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{N!} \int_{\Lambda^N} e^{-\beta U(x_1, \dots, x_N)} dx_1 \cdots dx_N$$

is the corresponding grand canonical partition function.

Note that most of the quantities we deal with depend on  $\beta$ ; however, since we will keep the temperature fixed throughout this paper, we refrain from making this dependency explicit in our notation.

Let  $S(\mathbb{P}_{*,\Lambda})$  denote the entropy associated with  $\mathbb{P}_{*,\Lambda}$ . In the spirit of (1.1), we then specify the relative entropy

$$\begin{aligned} S_{\text{rel},\Lambda} &= S(\mathbb{P}_{*,\Lambda}) + \int_{\Gamma_\Lambda} (\log \Xi_\Lambda - \beta\mu N(\gamma) + \beta U(\gamma)) d\mathbb{P}_{*,\Lambda}(\gamma) \\ &= S(\mathbb{P}_{*,\Lambda}) + \log \Xi_\Lambda - \beta\mu \mathbb{E}_{*,\Lambda}[N] + \beta \mathbb{E}_{*,\Lambda}[U], \end{aligned}$$

where  $\mathbb{E}_{*,\Lambda}[\cdot]$  denotes expectation with respect to  $\mathbb{P}_{*,\Lambda}$ . Take note that the entropy and the expected interaction energy may be  $+\infty$ . Dividing by  $\beta|\Lambda|$ , we arrive at

$$\frac{1}{\beta|\Lambda|} S_{\text{rel},\Lambda} = \frac{1}{\beta|\Lambda|} S(\mathbb{P}_{*,\Lambda}) + \frac{1}{\beta|\Lambda|} \log \Xi_\Lambda - \mu\rho_* + \frac{1}{|\Lambda|} \mathbb{E}_{*,\Lambda}[U]. \quad (2.4)$$

Now, we want to derive the analog of this identity for the thermodynamic limit  $\ell \rightarrow \infty$ . Concerning this limit it is known that there is a sequence  $(\ell_k)_k$  with  $\ell_k \rightarrow \infty$ , such that the probability densities associated with the corresponding grand canonical ensembles converge in a local topology to a translation invariant tempered Gibbs measure on  $\Gamma$ , cf. Ruelle [39], for brevity called  $(\mu, u)$ -Gibbs measure in the sequel. In general, different sequences may lead to different  $(\mu, u)$ -Gibbs measures, e.g., when different phases coexist. For any such sequence, however, the limit

$$p(\mu, u) = \lim_{\ell \rightarrow \infty} \frac{1}{\beta|\Lambda|} \log \Xi_\Lambda \quad (2.5)$$

is always the same well-defined and nonnegative finite number, namely the pressure of the ensemble in the thermodynamic limit, cf. Ruelle [38]. Therefore, passing in (2.4) to the thermodynamic limit  $\ell \rightarrow \infty$ , we can (uniquely) define the relative entropy  $S_{\text{rel}}(\mu, u)$  of all these  $(\mu, u)$ -Gibbs measures with respect to the target model  $\mathbb{P}_*$  as

$$\frac{1}{\beta} S_{\text{rel}}(\mu, u) = \frac{1}{\beta} S_* + p(\mu, u) - \mu\rho_* + E(u, \mathbb{P}_*), \quad (2.6)$$

where the individual terms on the right-hand side of (2.6) are the corresponding limits of the respective terms in (2.4):  $S_* = S(\mathbb{P}_*)$  is the specific entropy and  $E(u, \mathbb{P}_*)$  is the specific interaction energy (with respect to the potential  $u$ ) of the target ensemble, both of which have been shown to be well-defined in  $\mathbb{R} \cup \{+\infty\}$ , cf. [11, 13]. If  $\mathbb{P}_*$  satisfies a Ruelle condition (compare (A.1) in the appendix), then the specific entropy

is finite; this is the case, e.g., when  $\mathbb{P}_*$  is a  $(\mu_*, u_*)$ -Gibbs measure for some  $u_* \in \mathcal{U}$  and  $\mu_* \in \mathbb{R}$ , compare [39, Corollary 5.3]. For this latter particular case, it has further been shown in [7] that

$$E(u, \mathbb{P}_*) = \frac{1}{2} \int_{\mathbb{R}^d} u(x) \rho_*^{(2)}(x) \, dx, \quad (2.7)$$

where  $\rho_*^{(2)}(x)$  is the pair correlation function associated with  $\mathbb{P}_*$ . Here, and throughout this paper, we deliberately use the short-hand notation  $\rho^{(2)}(x)$  instead of  $\rho^{(2)}(x, 0)$  for the pair correlation function of a translation invariant point process.

In [10], Georgii studied the relative entropy  $S_{\text{rel}}(\mu, u)$  in detail and established the following *Gibbs variational principle*:

**Theorem A** (Gibbs variational principle) *Let  $\mathbb{P}_*$  be a translation invariant probability measure on  $\Gamma$  with finite locally second moments. Then, the relative entropy (2.6) is nonnegative real or  $+\infty$  for every  $u \in \mathcal{U}$  and  $\mu \in \mathbb{R}$ . There holds  $S_{\text{rel}}(\mu, u) = 0$ , if and only if  $\mathbb{P}_*$  is a  $(\mu, u)$ -Gibbs measure.*

We have utilized this result in [7] to prove a rigorous version of the Henderson theorem:

**Theorem B** (Henderson theorem) *Let  $u_1, u_2 \in \mathcal{U}$  and  $\mu_1, \mu_2 \in \mathbb{R}$ . If  $\mathbb{P}_1$  and  $\mathbb{P}_2$  are  $(\mu_1, u_1)$ - and  $(\mu_2, u_2)$ -Gibbs measures, respectively, which share the same density and the same pair correlation function, then  $u_1 = u_2$  and  $\mu_1 = \mu_2$ .*

Combining Georgii's Gibbs variational principle and the Henderson theorem, we can formulate an alternative version of the Gibbs variational principle. This version is the one that we will mostly use below.

**Theorem 2.2** (Gibbs variational principle, alternative form) *If the target  $\mathbb{P}_*$  is a  $(\mu_*, u_*)$ -Gibbs measure for some  $u_* \in \mathcal{U}$  and  $\mu_* \in \mathbb{R}$ , then the relative entropy  $S_{\text{rel}}(\mu, u)$  becomes minimal, if and only if  $u = u_*$  and  $\mu = \mu_*$ .*

**Proof** According to Theorem A, it remains to investigate the case when  $S_{\text{rel}}(\mu, u) = 0$  for some  $u \in \mathcal{U}$  and  $\mu \in \mathbb{R}$ . The Gibbs variational principle states that  $\mathbb{P}_*$  is then a  $(\mu_*, u_*)$ - and  $(\mu, u)$ -Gibbs measure at the same time. In particular this means that these two Gibbs measures share the same density and pair correlation function. The assertion thus follows from the Henderson theorem.  $\square$

Georgii investigated the relative entropy for fixed interaction and chemical potentials  $u$  and  $\mu$ , and varied the target model  $\mathbb{P}_*$ . In connection with the inverse Henderson problem, our interest is sort of dual to this: we assume that  $\mathbb{P}_*$  is fixed, and consider the relative entropy as a function of  $\mu$  and  $u$ .

### 3 Strict convexity of the pressure and the relative entropy

It is well-known that the pressure is a convex function of the chemical potential, cf. [38, Theorem 3.4.6]. This convexity is strict, whenever a Gibbs variational principle is valid,

cf., e.g., Hughes [19, Sect. 4.3]. In the sequel, we show that under our assumptions the pressure is also strictly convex in  $\mu$  and  $u$ .

**Theorem 3.1** *The pressure  $p = p(\mu, u)$  of (2.5) is a strictly convex function of  $\mu \in \mathbb{R}$  and  $u \in \mathcal{U}$ . Moreover, for  $u \in \mathcal{U}_0$ , there holds*

$$\frac{p(\mu, u)}{\mu} \rightarrow +\infty \quad \text{as } \mu \rightarrow +\infty. \quad (3.1)$$

**Proof** To begin with, we first observe that  $\mathbb{R} \times \mathcal{U}$  is convex by virtue of Remark 2.1.

Now let  $\mu_1, \mu_2 \in \mathbb{R}$  and  $u_1, u_2 \in \mathcal{U}$  be arbitrarily chosen, not both being equal at the same time, and define

$$\mu = t\mu_1 + (1-t)\mu_2 \quad \text{and} \quad u = tu_1 + (1-t)u_2$$

for some fixed  $t \in (0, 1)$ . When choosing for  $\mathbb{P}_*$  a corresponding  $(\mu, u)$ -Gibbs measure, then it follows from (2.6) and the Gibbs variational principle of Theorem 2.2 that

$$p(\mu, u) = -\frac{1}{\beta} S_* + \mu\rho_* - E(u, \mathbb{P}_*). \quad (3.2)$$

In particular, the specific entropy  $S_*$  is finite because  $\mathbb{P}_*$  is a Gibbs measure, and hence, so is the specific interaction energy. Likewise, we obtain

$$p(\mu_1, u_1) > -\frac{1}{\beta} S_* + \mu_1\rho_* - E(u_1, \mathbb{P}_*)$$

and

$$p(\mu_2, u_2) > -\frac{1}{\beta} S_* + \mu_2\rho_* - E(u_2, \mathbb{P}_*),$$

which gives

$$t p(\mu_1, u_1) + (1-t) p(\mu_2, u_2) > -\frac{1}{\beta} S_* + \mu\rho_* - E(u, \mathbb{P}_*)$$

because of the linearity of the specific interaction energy. A comparison with (3.2) thus shows that the pressure is strictly convex.

Consider now a fixed pair potential  $u \in \mathcal{U}_0$ . It has been shown in the proof of [9, Lemma 7.1] that for any  $\rho > 0$  there exists a translation invariant probability measure  $\mathbb{P}_\rho$  on  $\Gamma$  with density  $\rho$ , such that  $S(\mathbb{P}_\rho) < \infty$ , and  $E(u, \mathbb{P}_\rho) < \infty$ . Choosing  $\mathbb{P}_* = \mathbb{P}_\rho$  in (2.6), Georgii's Gibbs variational principle (Theorem A) yields the inequality

$$p(\mu, u) - \mu\rho \geq -\frac{1}{\beta} S(\mathbb{P}_\rho) - E(u, \mathbb{P}_\rho) =: c_\rho > -\infty \quad (3.3)$$

for every  $\mu \in \mathbb{R}$ . In other words,

$$\frac{p(\mu, u)}{\mu} \geq \rho + \frac{c_\rho}{\mu},$$

and hence,

$$\liminf_{\mu \rightarrow +\infty} \frac{p(\mu, u)}{\mu} \geq \rho.$$

Since  $\rho > 0$  has been arbitrary, this implies (3.1).  $\square$

We thus have shown that the relative entropy (2.6) is the sum of a strictly convex functional and an affine function of  $(\mu, u)$ . Accordingly, the relative entropy is also a strictly convex functional of  $\mu$  and  $u$ , as long as it is finite.

Another immediate consequence of Theorem 2.2 is the following inequality for the pressure.

**Corollary 3.2** For  $\mu_1, \mu_2 \in \mathbb{R}$  and  $u_1, u_2 \in \mathcal{U}$ , there holds

$$\begin{aligned} (\mu_2 - \mu_1)\rho_1 - \frac{1}{2} \int_{\mathbb{R}^d} (u_2 - u_1)(x)\rho_1^{(2)}(x) dx \\ \leq p(\mu_2, u_2) - p(\mu_1, u_1) \leq (\mu_2 - \mu_1)\rho_2 - \frac{1}{2} \int_{\mathbb{R}^d} (u_2 - u_1)(x)\rho_2^{(2)}(x) dx. \end{aligned} \quad (3.4)$$

whenever  $\rho_i$  and  $\rho_i^{(2)}$  are the density and pair correlation function of a  $(\mu_i, u_i)$ -Gibbs measure, respectively. Both inequalities are strict, unless  $\mu_1 = \mu_2$  and  $u_1 = u_2$ . In particular, if  $u_1 = u_2$  and  $\mu_1 < \mu_2$ , then  $\rho_1 < \rho_2$ .

**Proof** The two inequalities (3.4) follow readily from Theorem 2.2 by choosing for  $\mathbb{P}_*$  the corresponding  $(\mu_i, u_i)$ -Gibbs measures, respectively. They are strict, unless  $\mu_1 = \mu_2$  and  $u_1 = u_2$ .  $\square$

## 4 The relative entropy functional for fixed density

Returning to the inverse Henderson problem formulated in the introduction, we now constrain our model ensembles to have the same density  $\rho_*$  as the target ensemble. We do so because numerical simulations are often done in the canonical ensemble, where the density is the crucial quantity, and because this is also the setting used by Henderson and by Shell, as well as in the definition of the IMC scheme mentioned before. Since the attainable densities for hard-core potentials are bounded, we need to distinguish the case whether  $u$  is a hard-core potential or not. We focus our analysis on the latter case and mention the necessary modifications for hard-core potentials in Remark 4.4 later in this section.

The first fundamental problem to settle concerns the question whether and how the prescribed density  $\rho_*$  can be attained by some  $(\mu, u)$ -Gibbs measure for a given

$u \in \mathcal{U}_0$ . When the chemical potential is sufficiently small, i.e., when the system is in the *gas phase* (see Sect. 5 for a specification of this term), then it is known that there is a one-to-one relation between the corresponding chemical potentials and the associated densities, and in this case, the inverse map  $\rho \mapsto \mu$  can even be computed by means of cluster expansions, cf., e.g., Jansen, Kuna, and Tsagkarogiannis [22]. Outside the gas phase, where phase transitions may occur, the problem becomes more difficult. Adopting a method from Chayes and Chayes [2], we can establish the following result, which holds for the full range of possible densities.

**Theorem 4.1** *Let  $u \in \mathcal{U}_0$  and  $\rho_* > 0$  be fixed. Then, there is a unique chemical potential  $\mu_* = \mu_*(u) \in \mathbb{R}$ , for which there exists a  $(\mu_*, u)$ -Gibbs measure with density  $\rho_*$ .*

**Proof** For every  $\mu \in \mathbb{R}$  and the given  $u \in \mathcal{U}_0$ , let  $\mathbb{P}_{\mu,u}$  be a  $(\mu, u)$ -Gibbs measure, and denote by  $p(\mu)$  and  $\rho(\mu)$  the associated pressure and density, respectively.

It is well-known, cf. [38, Theorem 4.3.1], that for small chemical potentials, the pressure is a differentiable function with

$$p'(\mu) = \rho(\mu). \quad (4.1)$$

In other words, if  $\rho_*$  is sufficiently small, then the corresponding chemical potential  $\mu_*$  from the formulation of the theorem is given as the unique minimum of the function

$$\Theta(\mu) = p(\mu) - \mu\rho_*, \quad (4.2)$$

the latter being strictly convex by virtue of Theorem 3.1. We will proceed by showing that the minimizer of  $\Theta$  is also the appropriate chemical potential to choose for larger values of  $\rho_*$ .

To see this, we first observe that

$$\Theta(\mu) \geq -\mu\rho_*,$$

because the pressure is nonnegative, whereas

$$\Theta(\mu) \geq c_{(1+\varepsilon)\rho_*} + \varepsilon\mu\rho_* \quad (4.3)$$

for any suitable  $\varepsilon > 0$  and corresponding constant  $c_{(1+\varepsilon)\rho_*}$  by virtue of (3.3) (with  $\rho = (1 + \varepsilon)\rho_*$ ). This shows that  $\Theta$  is bounded from below and that

$$\Theta(\mu) \rightarrow +\infty, \quad \text{whenever } |\mu| \rightarrow \infty.$$

Therefore,  $\Theta$  attains its minimum for a uniquely defined value  $\mu = \mu_*$ .

For any  $\mu \in \mathbb{R}$ ,  $\mu \neq \mu_*$ , we now conclude from (3.4) and (4.2) that

$$(\mu_* - \mu)\rho(\mu) < p(\mu_*) - p(\mu) = \Theta(\mu_*) + \mu_*\rho_* - \Theta(\mu) - \mu\rho_*,$$

i.e.,

$$(\mu_* - \mu)(\rho(\mu) - \rho_*) < \Theta(\mu_*) - \Theta(\mu) < 0.$$

This means that

$$\begin{aligned} \rho(\mu) &< \rho_* && \text{for } \mu < \mu_*, \\ \rho(\mu) &> \rho_* && \text{for } \mu > \mu_*. \end{aligned} \quad (4.4)$$

Now, let  $(\mu_k^+)_k$  be a strictly decreasing sequence and  $(\mu_k^-)_k$  a strictly increasing sequence of chemical potentials, both of which converge to  $\mu_*$ . By virtue of Lemma A.1 from the appendix, there exist  $(\mu_*, u)$ -Gibbs measures  $\mathbb{P}^-$  and  $\mathbb{P}^+$  with densities

$$\rho(\mathbb{P}^-) = \lim_{k \rightarrow \infty} \rho(\mu_k^-) \quad \text{and} \quad \rho(\mathbb{P}^+) = \lim_{k \rightarrow \infty} \rho(\mu_k^+).$$

The inequalities (4.4) imply that

$$\rho(\mathbb{P}^-) \leq \rho_* \leq \rho(\mathbb{P}^+),$$

and hence, there is some  $t \in [0, 1]$ , for which  $\mathbb{P} = t\mathbb{P}^- + (1-t)\mathbb{P}^+$  has density

$$\rho(\mathbb{P}) = t\rho(\mathbb{P}^-) + (1-t)\rho(\mathbb{P}^+) = \rho_*.$$

Since the set of  $(\mu_*, u)$ -Gibbs measures is convex,  $\mathbb{P}$  has all the desired properties from the statement of this theorem, and the proof is done.  $\square$

In the light of the above theorem, we can now restrict our attention to  $(\mu, u)$ -Gibbs measures with density  $\rho_*$ , i.e., with chemical potential  $\mu = \mu_*(u)$ , when looking at the relative entropy functional. Further, we drop the constant offset  $S_*$  in (2.6), as it is independent of  $u$ . This leads to the functional

$$\Phi(u) = p_*(u) - \mu_*(u)\rho_* + E(u, \mathbb{P}_*) \quad (4.5)$$

for  $u \in \mathcal{U}_0$ , where we have set

$$p_*(u) = p(\mu_*(u), u) \quad (4.6)$$

for brevity. By a slight abuse of wording we will keep calling  $\Phi$  the relative entropy functional. As we will see next, although  $p_*$  may fail to be convex, in general, the functional  $\Phi$  is strictly convex, again.

**Theorem 4.2** *Let  $\mathbb{P}_*$  be as in Theorem A. Then, the functional  $\Phi : \mathcal{U}_0 \rightarrow \mathbb{R}$  of (4.5) is strictly convex as long as the specific interaction energy  $E(u, \mathbb{P}_*)$  is finite.*

**Proof** Let  $u_1$  and  $u_2$  be two different pair potentials from  $\mathcal{U}_0$ . For any fixed  $0 < t < 1$ , define

$$u = tu_1 + (1-t)u_2,$$

and, as in Theorem 4.1, denote by  $\mu = \mu_*(u)$  the chemical potential associated with  $u$  and density  $\rho_*$ . Finally, let  $\rho^{(2)}$  be the pair correlation function of the associated  $(\mu, u)$ -Gibbs measure constructed in Theorem 4.1.

Then, we obtain from (4.6) and (3.4)—with  $\mu_i = \mu_*(u_i)$  for  $i = 1, 2$ —that

$$\begin{aligned} & p_*(u) - tp_*(u_1) - (1-t)p_*(u_2) \\ &= t(p_*(u) - p_*(u_1)) + (1-t)(p_*(u) - p_*(u_2)) \\ &< t\left((\mu - \mu_1)\rho_* - \frac{1-t}{2} \int_{\mathbb{R}^d} (u_2 - u_1)(x)\rho^{(2)}(x) dx\right) \\ &\quad + (1-t)\left((\mu - \mu_2)\rho_* - \frac{t}{2} \int_{\mathbb{R}^d} (u_1 - u_2)(x)\rho^{(2)}(x) dx\right) \\ &= \mu\rho_* - t\mu_1\rho_* - (1-t)\mu_2\rho_* . \end{aligned}$$

Note that this inequality is strict because  $u$  is different from  $u_1$  and  $u_2$  by construction. Reordering terms, we thus arrive at

$$p_*(u) - \mu\rho_* < t(p_*(u_1) - \mu_1\rho_*) + (1-t)(p_*(u_2) - \mu_2\rho_*). \quad (4.7)$$

It thus follows from (4.5) and (4.7) that if the specific interaction energy  $E(\cdot, \mathbb{P}_*)$  stays finite, then  $\Phi$  is strictly convex, because  $E(\cdot, \mathbb{P}_*)$  is linear in the first argument.  $\square$

Theorem 4.2 implies that the relative entropy functional has at most one local minimizer, which is then also a global one. Concerning this minimizer, we have the following result.

**Theorem 4.3** *Let  $\mathbb{P}_*$  be a  $(\mu, u_*)$ -Gibbs measure for some  $u_* \in \mathcal{U}_0$  and  $\mu \in \mathbb{R}$ , and let  $\rho_*$  be its density. Then, the relative entropy function (4.5) attains its minimum for  $u = u_*$ .*

**Proof** The given Gibbs measure  $\mathbb{P}_*$  has finite specific entropy  $S_*$ , and hence, (2.6) implies that

$$\Phi(u) = \frac{1}{\beta} S_{\text{rel}}(\mu_*(u), u) - \frac{1}{\beta} S_* \quad (4.8)$$

for every  $u \in \mathcal{U}_0$ . Since  $\mathbb{P}_*$  has density  $\rho_*$  the chemical potential  $\mu$  associated with  $\mathbb{P}_*$  must be given by  $\mu = \mu_*(u_*)$  according to Theorem 4.1. From (4.8) and the Gibbs variational principle (Theorems A and 2.2), it follows that

$$\Phi(u_*) = -\frac{1}{\beta} S_* < \frac{1}{\beta} S_{\text{rel}}(\mu_*(u), u) - \frac{1}{\beta} S_* = \Phi(u) \quad (4.9)$$

for every  $u \in \mathcal{U}_0 \setminus \{u_*\}$ , and this was to be shown.  $\square$

Note from (4.9) that the minimal value of  $\Phi$  depends on the specific entropy of the target Gibbs measure, and is therefore unknown in general.

**Remark 4.4** For a hard-core potential  $u \in \mathcal{U}$  the density  $\rho$  of any associated  $(\mu, u)$ -Gibbs measure is bounded from above by a finite *closest-packing density*  $\rho_{\text{cp}} = \rho_{\text{cp}}(u)$ : This bound only depends on the hard-core radius  $r_0$  of (2.2) and is a decreasing function of  $r_0$ , cf. [9, Sect. 7]. We formally set  $\rho_{\text{cp}}(u) = +\infty$  for  $u \in \mathcal{U}_0$ .

If the given density  $\rho_*$  of the target happens to be below  $\rho_{\text{cp}}(u)$  for a given  $u \in \mathcal{U}$ , then Theorem 4.1 is still valid for this pair potential; to establish (4.3) in its proof,  $\varepsilon > 0$  must be so small that  $(1 + \varepsilon)\rho_*$  is still below  $\rho_{\text{cp}}(u)$ , because this is needed for [9, Lemma 7.1], and hence, for (3.3). The relative entropy functional (4.5), however, is only well-defined on the domain

$$\mathcal{D}(\Phi) = \{u \in \mathcal{U} : \rho_{\text{cp}}(u) > \rho_*\},$$

but this is again a convex set, compare Remark 2.1, and  $\Phi$  happens to be strictly convex on  $\mathcal{D}(\Phi)$ . Theorem 4.3 extends literally to every  $u_* \in \mathcal{U}$  with this understanding of the domain of  $\Phi$ .  $\diamond$

Theorem 4.3 is the basis for a variational setting of the inverse Henderson problem: If the density and the pair correlation function of a  $(\mu_*, u_*)$ -Gibbs measure target  $\mathbb{P}_*$  are given, then the unique minimizer of the strictly convex relative entropy functional (4.5) with  $E(u, \mathbb{P}_*)$  of (2.7) yields the corresponding pair potential  $u_*$ , i.e., the solution of the inverse Henderson problem. Nevertheless, this approach also has some pitfalls as discussed in the following remark.

**Remark 4.5** If the target  $\mathbb{P}_*$  fails to be some  $(\mu, u)$ -Gibbs measure, then it is not clear to us whether the relative entropy functional  $\Phi$  will still be bounded from below on  $\mathcal{U}$ , and even if it may, its infimum need not be attained on  $\mathcal{U}$ .

Vice versa, if  $u_* \in \mathcal{U}$  happens to be the minimizer of  $\Phi$ , then this does not imply that  $\mathbb{P}_*$  is a  $(\mu_*(u_*), u_*)$ -Gibbs measure. To see the latter, consider the following example: Let  $u_*$  be any hard-core potential in  $\mathcal{U}$  and  $\rho_* < \rho_{\text{cp}}(u_*)$ ; compare Remark 4.4. Then, provided  $\rho_*$  is sufficiently small, a result by Kuna, Lebowitz, and Speer [24, Corollary 4.3] states that there exist uncountably many distinct translation invariant probability measures with finite specific entropy and density  $\rho_*$ , which share the pair correlation function  $\rho_*^{(2)}$  with the  $(\mu_*(u_*), u_*)$ -Gibbs measure  $\mathbb{P}$ , but  $\mathbb{P}$  is the only  $(\mu, u)$ -Gibbs measure among them according to the Henderson theorem. Because the specific interaction energy is given by (2.7) for all of these probability measures, the relative entropy functional does not differ, and hence,  $u_*$  is its unique minimizer, regardless which of them has been the target  $\mathbb{P}_*$ .

Finally, we mention that even if the target  $\mathbb{P}_*$  is a  $(\mu, u_*)$ -Gibbs measure, and hence, the functional (4.5) is minimized by  $u_*$  according to Theorem 4.3, then this does not seem to imply that the model and the target have the same pair correlation function. To illustrate this, imagine that a fluid corresponding to a pair potential  $u_* \in \mathcal{U}$  exhibits a so-called *triple point* (compare, e.g., [17]), where three different phases coexist at the same thermodynamical state point, i.e., for the same values of pressure (or chemical potential  $\mu_*$ , say) and temperature (or inverse temperature  $\beta$ ). It can be expected that the different phases have linearly independent pair correlation functions; taking convex combinations of the corresponding Gibbs measures, one can thus determine

two  $(\mu_*, u_*)$ -Gibbs measures which have the same density  $\rho_*$ , but have different pair correlation functions. One of them could be the target and the other one the minimizing model. We will see in Sect. 5 that if the density  $\rho_*$  belongs to the gas phase of the minimizer of  $\Phi$ , then the pair correlation function of the minimizing Gibbs measure always coincides with the given  $\rho_*^{(2)}$ ; see Remark 5.5.  $\diamond$

## 5 Differentiability of the relative entropy functional

In the remainder of this paper, we further analyze the case when  $\Phi$  is a differentiable function of  $u$ . For this, we build upon our earlier work [14, 15].<sup>1</sup> A conceptual difficulty in this context is the fact that  $\mathcal{U}$  lacks a universal topology. Therefore, following [14], we consider a tailored neighborhood for any given  $u_0 \in \mathcal{U}$  by introducing a corresponding Banach space  $\mathcal{V}$  of *perturbations*, consisting of all even measurable functions  $v: \mathbb{R}^d \rightarrow \mathbb{R}$ , for which the associated norm

$$\|v\|_{\mathcal{V}} = \sup_{x \in \mathbb{R}^d} \frac{|v(x)|}{\psi_0(|x|)} \quad (5.1)$$

is finite. Here,  $\psi_0$  is the majorant  $\psi$  of (2.1) associated with  $u_0$ . Note that the norm (5.1) is somewhat stronger than the one employed in [14] because it is more restrictive in the core region  $0 < r < r_0$ , and hence, the resulting space of perturbations is smaller. We believe that this restriction allows a more natural formulation of our results.

**Remark 5.1** It can easily be seen that for any  $v \in \mathcal{V}$  the sum  $u_0 + v$  belongs to  $\mathcal{U}$ : If  $u_0$  is a hard-core potential, then  $r_0$  is its hard-core radius and the corresponding function  $\varphi$  from (2.1) equals  $+\infty$ , and we have

$$\begin{aligned} u_0(x) + v(x) &\geq \varphi(|x|) && \text{for } 0 < |x| < r_0, \\ |u_0(x) + v(x)| &\leq (1 + \|v\|_{\mathcal{V}})\psi_0(|x|) && \text{for } |x| \geq r_0. \end{aligned}$$

Otherwise, there exists  $r_1 \in (0, r_0]$  such that  $\varphi(r) \geq \|v\|_{\mathcal{V}}\psi_0(0)$  for  $0 < r \leq r_1$ , and therefore

$$\begin{aligned} u_0(x) + v(x) &\geq \varphi(|x|) - \|v\|_{\mathcal{V}}\psi_0(0) && \text{for } 0 < |x| < r_1, \\ |u_0(x) + v(x)| &\leq \left( \sup_{|x'| \geq r_1} \frac{|u(x')|}{\psi_0(|x'|)} + \|v\|_{\mathcal{V}} \right) \psi_0(|x|) && \text{for } |x| \geq r_1. \end{aligned}$$

In either case, this shows that  $u_0 + v \in \mathcal{U}$ .  $\diamond$

We now consider the ball

$$\mathcal{B}_{\mathcal{V}}(u_0) = \{u = u_0 + v : \|v\|_{\mathcal{V}} < \delta_0\} \subset \mathcal{U} \quad (5.2)$$

<sup>1</sup> The results in [14, 15] are formulated in three-dimensional space (i.e., for  $d = 3$ ) and for pair potentials, which are rotation invariant; also, hard-core potentials had not been considered. All results extend verbatim to the setting considered here.

around  $u_0$  for a suitable radius  $\delta_0$  with  $0 < \delta_0 < 1$ ; we further specify  $\delta_0$  in the context of (5.7) below. As has been verified in [14, Proposition 2.1], there are positive constants  $c_\beta$  and  $B$ , such that every  $u \in \mathcal{B}_\gamma(u_0)$  satisfies

$$\int_{\mathbb{R}^d} |e^{-\beta u(x)} - 1| dx \leq c_\beta, \quad (5.3)$$

and the interaction energy of every configuration  $\gamma \subset \mathbb{R}^d$  with  $N(\gamma)$  particles is bounded from below by

$$U(\gamma) \geq -BN(\gamma). \quad (5.4)$$

Moreover, for every  $u \in \mathcal{B}_\gamma(u_0)$  and  $\mu \in \mathbb{R}$  and every associated  $(\mu, u)$ -Gibbs measure, the sequence

$$\boldsymbol{\rho}(\mu, u) = [\rho, \rho^{(2)}, \dots]^T$$

of its  $m$ -particle correlation functions is a solution of the so-called *Kirkwood-Salsburg equations* (cf., e.g., [39, Corollary 5.3], or compare (A.3) in the appendix), which can be written in the form

$$(I - e^{\beta\mu} A(u))\boldsymbol{\rho}(\mu, u) = e^{\beta\mu} \mathbf{e}_1, \quad (5.5)$$

where  $\mathbf{e}_1 = [1, 0, \dots]^T$ ,  $I$  is the identity operator, and  $A(u)$  is a bounded operator in an associated Banach space  $\mathcal{X}$  of sequences of  $L^\infty$  functions. This operator  $A(u)$  is given by a certain infinite-dimensional matrix of integral operators, compare [14, 38], and is Fréchet differentiable with respect to  $u \in \mathcal{B}_\gamma(u_0)$ .

We fix

$$\mu_0 < \frac{1}{\beta} \log \frac{1}{c_\beta e^{2\beta B+1}}, \quad (5.6)$$

and call the range  $\mu \leq \mu_0$  of chemical potentials the *gas phase* associated with the pair potentials in  $\mathcal{B}_\gamma(u_0)$ . In this gas phase, the Kirkwood-Salsburg equations (5.5) have a unique solution  $\boldsymbol{\rho}(\mu, u)$ , which can be developed in a converging Neumann series in  $\mathcal{X}$ ; in particular, this means that for  $u \in \mathcal{B}_\gamma(u_0)$ , there exists only one  $(\mu, u)$ -Gibbs measure for each chemical potential  $\mu$  within the gas phase. The individual components  $\rho^{(m)}(\mu, u)$  of  $\boldsymbol{\rho}(\mu, u)$  are Fréchet differentiable (in  $L^\infty$ ) with respect to  $u$ , analytic with respect to  $\mu < \mu_0$  and continuous in  $\mu \leq \mu_0$ . It is also easy to see that  $\boldsymbol{\rho}$  is a  $C^1$  function of  $\mu \leq \mu_0$  and  $u \in \mathcal{B}_\gamma(u_0)$ .

Further, for fixed  $u \in \mathcal{B}_\gamma(u_0)$ , the density  $\rho = \rho(\mu, u)$  of the associated Gibbs measure, i.e., the first entry of  $\boldsymbol{\rho}(\mu, u)$ , is differentiable and strictly increasing in  $\mu$  up to  $\mu = \mu_0$  by virtue of Corollary 3.2. Since it is also continuous in  $u \in \mathcal{B}_\gamma(u_0)$ , we can choose  $\delta_0$  in (5.2) so small that

$$\rho_0 = \inf \{ \rho(\mu_0, u) : u \in \mathcal{B}_\gamma(u_0) \} \geq (1 - \varepsilon) \rho(\mu_0, u_0) \quad (5.7)$$

for any given  $\varepsilon > 0$ , and then every density  $\rho \in (0, \rho_0)$  belongs to the gas phase of all pair potentials in  $\mathcal{B}_\gamma(u_0)$ . Finally, for fixed  $u \in \mathcal{B}_\gamma(u_0)$ , the pressure  $p(\mu, u)$  is also

differentiable and strictly increasing in  $\mu$  up to  $\mu = \mu_0$ , and its derivative is given by the density, cf. (4.1).

**Proposition 5.2** *Let  $0 < \rho_* < \rho_0$  with  $\rho_0$  as in (5.7). Then, the chemical potential  $\mu_* = \mu_*(u)$  defined in Theorem 4.1 is differentiable with respect to  $u \in \mathcal{B}_{\mathcal{V}}(u_0)$  with derivative  $\mu'_*(u) \in \mathcal{V}'$  given by*

$$\mu'_*(u) v = -\frac{\partial_u \rho(\mu_*(u), u) v}{\partial_\mu \rho(\mu_*(u), u)} \quad \text{for } v \in \mathcal{V}, \quad (5.8)$$

where  $\partial_\mu \rho$  and  $\partial_u \rho$  denote the partial derivatives of  $\rho(\mu, u)$ , and  $\mathcal{V}'$  is the dual space of  $\mathcal{V}$ .

**Proof** Since the pressure and the density of a fixed pair potential  $u \in \mathcal{B}_{\mathcal{V}}(u_0)$  are strictly increasing and differentiable functions of the chemical potential in the gas phase, we can rewrite each of these functions as a strictly increasing function of any of the other ones. The pressure, for example, can be written as a function of density, which we call  $\pi$  to avoid any confusion, i.e.,

$$p(\mu, u) = \pi(\rho(\mu, u), u).$$

Then, the chain rule can be applied to obtain

$$\partial_\mu p(\mu, u) = \partial_\rho \pi(\rho(\mu, u), u) \partial_\mu \rho(\mu, u), \quad (5.9)$$

because it has been shown in [39, Theorem 4.3] that  $\pi$  is a differentiable function of the density in the gas phase. For  $\mu = \mu_*(u)$ , we thus conclude from (4.1) and (5.9) that

$$\rho_* = \rho(\mu_*(u), u) = \partial_\mu p(\mu_*(u), u) = \alpha \partial_\mu \rho(\mu_*(u), u) \quad (5.10)$$

with

$$\alpha = \partial_\rho \pi(\rho_*, u) > 0.$$

It thus follows from (5.10) that  $\partial_\mu \rho(\mu_*(u), u) > 0$ , so that the implicit function theorem is applicable to the equation

$$\rho(\mu_*(u+v), u+v) = \rho_*$$

near  $v = 0$ . From this, we readily obtain (5.8); moreover,  $\mu'_*(u)$  belongs to  $\mathcal{V}'$  because  $\partial_u \rho \in \mathcal{V}'$ .  $\square$

Now, we return to our analysis of the relative entropy functional (4.5).

**Theorem 5.3** *Let  $u_0 \in \mathcal{U}$  be a fixed pair potential and  $\mathcal{B}_{\mathcal{V}}(u_0) \subset \mathcal{U}$  be defined as in (5.2) with  $\delta_0$  so small that (5.7) holds. Furthermore, let the target  $\mathbb{P}_*$  of the relative entropy functional have density  $\rho_* < \rho_0$ , and assume that the associated specific*

interaction energy  $E(\cdot, \mathbb{P}_*)$  is bounded on  $\mathcal{B}_{\mathcal{V}}(u_0)$ . Then,  $\Phi$  is Fréchet-differentiable in every  $u \in \mathcal{B}_{\mathcal{V}}(u_0)$  with derivative  $\Phi'(u) \in \mathcal{V}'$ , given by

$$\Phi'(u)v = E(v, \mathbb{P}_*) - \frac{1}{2} \int_{\mathbb{R}^d} v(x) \rho^{(2)}(x) \, dx \quad (5.11)$$

for  $v \in \mathcal{V}$ , where  $\rho^{(2)}$  is the pair correlation function of the associated  $(\mu_*(u), u)$ -Gibbs measure.

**Proof** Let  $u \in \mathcal{B}_{\mathcal{V}}(u_0)$  and  $v \in \mathcal{V}$ . We apply Corollary 3.2 to  $u_1 = u$  and  $u_2 = u + v$  with  $\mu_1 = \mu_*(u)$  and  $\mu_2 = \mu_*(u + v)$ . This yields

$$\begin{aligned} -\frac{1}{2} \int_{\mathbb{R}^d} v(x) \rho^{(2)}(x) \, dx &\leq p_*(u + v) - \mu_*(u + v) \rho_* - (p_*(u) - \mu_*(u) \rho_*) \\ &\leq -\frac{1}{2} \int_{\mathbb{R}^d} v(x) \tilde{\rho}^{(2)}(x) \, dx, \end{aligned}$$

where  $\tilde{\rho}^{(2)}$  denotes the pair correlation function of the corresponding  $(\mu_*(u + v), u + v)$ -Gibbs measure. It follows that

$$\begin{aligned} E(v, \mathbb{P}_*) - \frac{1}{2} \int_{\mathbb{R}^d} v(x) \rho^{(2)}(x) \, dx &\leq \Phi(u + v) - \Phi(u) \\ &\leq E(v, \mathbb{P}_*) - \frac{1}{2} \int_{\mathbb{R}^d} v(x) \tilde{\rho}^{(2)}(x) \, dx, \end{aligned}$$

i.e., that

$$\begin{aligned} 0 &\leq \Phi(u + v) - \Phi(u) - \left( E(v, \mathbb{P}_*) - \frac{1}{2} \int_{\mathbb{R}^d} v(x) \rho^{(2)}(x) \, dx \right) \\ &\leq \frac{1}{2} \int_{\mathbb{R}^d} v(x) (\rho^{(2)} - \tilde{\rho}^{(2)})(x) \, dx. \end{aligned}$$

This shows that

$$\begin{aligned} &\left| \Phi(u + v) - \Phi(u) - \left( E(v, \mathbb{P}_*) - \frac{1}{2} \int_{\mathbb{R}^d} v(x) \rho^{(2)}(x) \, dx \right) \right| \\ &\leq \frac{C_{\psi_0}}{2} \|v\|_{\mathcal{V}} \|\tilde{\rho}^{(2)} - \rho^{(2)}\|_{L^\infty(\mathbb{R}^d)}, \end{aligned}$$

where

$$C_{\psi_0} = |\mathbb{S}^{d-1}| \int_0^\infty r^{d-1} \psi_0(r) \, dr < \infty$$

by virtue of (2.1). We finally observe that

$$\tilde{\rho}^{(2)} = \rho^{(2)}(\mu_*(u + v), u + v)$$

converges to  $\rho^{(2)} = \rho^{(2)}(\mu_*(u), u)$  in  $L^\infty(\mathbb{R}^d)$  as  $\|v\|_{\mathcal{Y}} \rightarrow 0$  because of Proposition 5.2 and the smoothness of  $\rho^{(2)}$  as a function of  $\mu$  and  $u$ . We have therefore established that

$$\left| \Phi(u+v) - \Phi(u) - \left( E(v, \mathbb{P}_*) - \frac{1}{2} \int_{\mathbb{R}^d} v(x) \rho^{(2)}(x) dx \right) \right| = o(\|v\|_{\mathcal{Y}})$$

as  $\|v\|_{\mathcal{Y}} \rightarrow 0$ , which yields (5.11). Since  $E(\cdot, \mathbb{P}_*)$  is bounded on  $\mathcal{B}_{\mathcal{Y}}(u_0)$ , it is easy to see that  $E(\cdot, \mathbb{P}_*) \in \mathcal{Y}'$ , and hence, it follows in the same way as above that  $\Phi'(u) \in \mathcal{Y}'$  with

$$\|\Phi'(u)\|_{\mathcal{Y}'} \leq \|E(\cdot, \mathbb{P}_*)\|_{\mathcal{Y}'} + \frac{C\psi_0}{2} \|\rho^{(2)}\|_{L^\infty(\mathbb{R}^d)}.$$

This concludes the proof.  $\square$

**Remark 5.4** We emphasize that the dual space  $\mathcal{Y}'$  of  $\mathcal{Y}$  is a complicated Banach space. However, the representation (5.11), in combination with (2.7), reveals that if  $\mathbb{P}_*$  has a bounded and measurable pair correlation function  $\rho_*^{(2)}$ , then  $\Phi'(u)$  can be identified with

$$\nabla \Phi(u) = \frac{1}{2} (\rho_*^{(2)} - \rho^{(2)}(\mu_*(u), u)) \in L^\infty(\mathbb{R}^d), \quad (5.12)$$

when using the natural dual pairing

$$\Phi'(u)v = \langle v, \nabla \Phi(u) \rangle = \int_{\mathbb{R}^d} v(x) (\nabla \Phi(u))(x) dx \quad (5.13)$$

of  $L^2(\mathbb{R}^d)$ .

This is the case, for example, if  $\mathbb{P}_*$  is a  $(\mu, u_*)$ -Gibbs measure for some  $u_* \in \mathcal{U}$  (with  $\mu = \mu_*(u_*)$ ), and then (5.12) can further be rewritten as

$$\nabla \Phi(u) = \frac{1}{2} (\omega^{(2)}(\mu, u_*) - \omega^{(2)}(\mu_*(u), u)),$$

where

$$\omega^{(2)}(\mu, u) = \rho^{(2)}(\mu, u) - \rho(\mu, u)^2. \quad (5.14)$$

It follows from (5.17) below that  $\nabla \Phi(u)$  then also belongs to  $L^1(\mathbb{R}^d)$ , and hence, to  $L^2(\mathbb{R}^d)$  as well. Note that  $\mathcal{Y}$  is embedded in  $L^2(\mathbb{R})$  for the same reason. We refer to Sect. 7 for further arguments which support the choice of the  $L^2$ -pairing.  $\diamond$

**Remark 5.5** Assume that the target  $\mathbb{P}_*$  has a bounded and measurable pair correlation function  $\rho_*^{(2)}$ , and that the relative entropy functional  $\Phi$  is minimized by some  $u_* \in \mathcal{U}$ . Assume further that the given density  $\rho_*$  belongs to the gas phase of  $u_*$ . Then, it follows from Theorem 5.3 that  $\Phi$  is differentiable at  $u_*$ , and that  $\nabla \Phi(u_*) = 0$ . From (5.12), we thus conclude that the target and the model share the same pair correlation function in this case; compare Remark 4.5.  $\diamond$

**Remark 5.6** For a  $(\mu, u)$ -Gibbs measure the *cluster functions*, sometimes also called Ursell functions (e.g., Stell [43], and [15]), or truncated correlation functions (e.g., Jansen [21]), are defined recursively by the constant function  $\omega^{(1)} = \rho$  and, for  $m \geq 2$ , by

$$\omega^{(m)}(x_1, \dots, x_m) = \rho^{(m)}(x_1, \dots, x_m) - \sum_{k=2}^m \sum_{\pi \in \Pi_k^m} \prod_{i=1}^k \omega^{(|\pi_i|)}(\mathbf{x}_{\pi_i}), \quad (5.15)$$

cf. [38, p. 87], where  $\Pi_k^m$  denotes the set of partitions  $\pi$  of  $x_1, \dots, x_m$  into  $k$  subsets  $\mathbf{x}_{\pi_1}, \dots, \mathbf{x}_{\pi_k}$ . Note that it immediately follows from (5.15) that each cluster function inherits the translation and permutation invariance of the correlation functions.

For  $m = 2$ , we recover from (5.15) the definition (5.14), and for fixed  $\mu$  and  $u$ , we adopt the short-hand notation

$$\omega^{(2)}(x) = \rho^{(2)}(x) - \rho^2$$

for  $\omega^{(2)}(x, 0)$  from the pair correlation function. Note that

$$\omega^{(2)}(x) = \omega^{(2)}(-x) \quad \text{for every } x \in \mathbb{R}^d \quad (5.16)$$

because of the translation invariance.

We will make repeatedly use of [38, Theorem 4.4.8] which states that

$$\int_{(\mathbb{R}^d)^{(m-2)}} |\omega^{(m)}(\cdot, 0, x_3, \dots, x_m)| dx_3 \cdots dx_m \in L^1(\mathbb{R}^d), \quad (5.17)$$

provided that  $u \in \mathcal{U}$  and  $\mu \leq \mu_0$ ; see also Lemma A.4 in the appendix. In particular, (5.17) shows that  $\omega^{(2)} \in L^1(\mathbb{R}^d)$ .  $\diamond$

Based on Theorem 5.3, we now compute the Hessian of  $\Phi$ .

**Theorem 5.7** *Under the assumptions of Theorem 5.3, the relative entropy functional (4.5) is twice differentiable, and its Hessian at  $u \in \mathcal{B}_\psi(u_0)$  is given by*

$$\Phi''(u)(v, \tilde{v}) = -\frac{1}{\partial_\mu \rho} (\partial_u \rho \tilde{v})(\partial_u \rho v) - \frac{1}{2} \langle v, \partial_u \rho^{(2)} \tilde{v} \rangle \quad (5.18)$$

in terms of the dual pairing (5.13) of  $L^2(\mathbb{R}^d)$ , where  $v, \tilde{v} \in \mathcal{V}$ , and the partial derivatives  $\partial_u \rho$  and  $\partial_u \rho^{(2)}$  are evaluated at the pair potential  $u$  and the chemical potential  $\mu_*(u)$ .

**Proof** We differentiate (5.11) with respect to  $u$  in the direction  $\tilde{v} \in \mathcal{V}$ . Keeping in mind that  $\rho^{(2)}$  in (5.11) stands for  $\rho^{(2)}(\mu_*(u), u)$ , the chain rule gives

$$\begin{aligned} \Phi''(u)(v, \tilde{v}) &= -\frac{1}{2} \int_{\mathbb{R}^d} v(x) \left( \partial_\mu \rho^{(2)} \mu'_*(u) \tilde{v} + \partial_u \rho^{(2)} \tilde{v} \right)(x) dx \\ &= -\frac{\mu'_*(u) \tilde{v}}{2} \langle v, \partial_\mu \rho^{(2)} \rangle - \frac{1}{2} \langle v, \partial_u \rho^{(2)} \tilde{v} \rangle, \end{aligned} \quad (5.19)$$

where we have adopted the notation (5.13), and the partial derivatives of  $\rho^{(2)}$ —which belong to  $L^\infty(\mathbb{R}^d)$ —are evaluated at the pair potential  $u$  and the chemical potential  $\mu_*(u)$ .

To determine  $\partial_\mu \rho^{(2)}(\mu, u)$  for an arbitrary chemical potential  $\mu < \mu_0$  we apply Corollary 3.2 with  $u_1 = u$  and  $u_2 = u + w$  for any  $w \in \mathcal{V}$ , and with  $\mu_1 = \mu_2 = \mu$ : As in the proof of Theorem 5.3, we conclude that the pressure is differentiable with respect to  $u$  with derivative  $\partial_u p(\mu, u) \in \mathcal{V}'$  given by

$$\partial_u p(\mu, u) w = -\frac{1}{2} \int_{\mathbb{R}^d} w(x) \rho^{(2)}(x) dx = -\frac{1}{2} \langle w, \rho^{(2)} \rangle,$$

where  $\rho^{(2)}$  is the pair correlation function of the  $(\mu, u)$ -Gibbs measure. Schwarz's theorem and (4.1) therefore yield

$$\langle w, \partial_\mu \rho^{(2)} \rangle = \partial_\mu \langle w, \rho^{(2)} \rangle = -2 \partial_\mu ((\partial_u p) w) = -2 \partial_u (\partial_\mu p) w = -2 \partial_u \rho w. \quad (5.20)$$

Inserting this into (5.19), we obtain

$$\Phi''(u)(v, \tilde{v}) = (\mu'_*(u) \tilde{v})(\partial_u \rho v) - \frac{1}{2} \langle v, \partial_u \rho^{(2)} \tilde{v} \rangle,$$

and assertion (5.18) thus follows from Proposition 5.2.  $\square$

Note that  $\Phi''(u)$  is a continuous bilinear form: This follows readily from the result in [14] that  $\partial_u \rho \in \mathcal{V}'$  and  $\partial_u \rho^{(2)} \in \mathcal{L}(\mathcal{V}, L^\infty(\mathbb{R}^d))$ . We further observe that  $\Phi''$  is continuous in  $u$  because  $\rho$  and  $\rho^{(2)}$  are  $C^1$  functions of  $\mu$  and  $u$ . This implies that  $\Phi''(u)$  is a symmetric bilinear form. Finally, this bilinear form is positive semidefinite, because  $\Phi$  is strictly convex.

## 6 The connection to the IMC iterative scheme

In the preceding section, we have seen that the relative entropy functional is twice differentiable at any  $u_0 \in \mathcal{U}$ , provided that the target density  $\rho_*$  belongs to its gas phase. The convex quadratic approximation

$$\Phi_2(u_0 + v) := \Phi(u_0) + \Phi'(u_0)v + \frac{1}{2} \Phi''(u_0)(v, v) \approx \Phi(u_0 + v), \quad (6.1)$$

valid for small enough  $v \in \mathcal{V}$ , attains its minimum for every solution  $v_0 \in \mathcal{V}$  of the variational problem

$$\Phi''(u_0)(v_0, w) = -\Phi'(u_0)w \quad \text{for all } w \in \mathcal{V}, \quad (6.2)$$

and the well-known Newton scheme for minimizing  $\Phi$  consists in updating  $u_0$  via

$$u_1 = u_0 + v_0 \quad (6.3)$$

to obtain a (hopefully) better approximation of the global unique minimizer of  $\Phi$ . Unfortunately, however, we only know that  $\Phi''(u_0)$  is semidefinite; it is therefore not clear whether (6.2) has a solution, and if it does, whether it is unique.

The bilinear form  $\Phi''(u_0)$  defines a linear operator  $A : \mathcal{V} \rightarrow \mathcal{V}'$  via the identity

$$\Phi''(u_0)(v, \tilde{v}) = \langle \tilde{v}, Av \rangle_{\mathcal{V}' \times \mathcal{V}} \quad \text{for all } v, \tilde{v} \in \mathcal{V}.$$

As in Remark 5.4 the dual form on the right-hand side can be replaced by the dual pairing induced by  $L^2(\mathbb{R}^d)$ , when  $A$  is identified with the (bounded) operator

$$A : \mathcal{V} \rightarrow L^\infty(\mathbb{R}^d), \quad A : v \mapsto -\frac{1}{2} \partial_u \rho^{(2)} v - \frac{1}{\partial_\mu \rho} (\partial_u \rho v) \nabla_u \rho, \quad (6.4)$$

where  $\nabla_u \rho \in L^\infty(\mathbb{R}^d)$  denotes the representative of the functional  $\partial_u \rho \in \mathcal{V}'$  for this pairing. This is an immediate consequence of Theorem 5.7. According to [15, Sect. 6] there holds

$$\nabla_u \rho(x) = -\beta \rho^{(2)}(x) - \frac{\beta}{2} \int_{\mathbb{R}^d} \chi^{(3)}(0, x + x', x') dx', \quad (6.5)$$

where

$$\chi^{(3)}(0, x + x', x') = \omega^{(3)}(0, x + x', x') + \rho \omega^{(2)}(x') + \rho \omega^{(2)}(x + x')$$

is given in terms of the second and third cluster functions associated with  $\mu$  and  $u$ ; compare Remark 5.6. Using the translation and permutation invariance of  $\omega^{(3)}$ , this can be rewritten as

$$\chi^{(3)}(0, x + x', x') = \omega^{(3)}(x, 0, -x') + \rho \omega^{(2)}(x') + \rho \omega^{(2)}(x + x'). \quad (6.6)$$

Note that it has been shown in [15, Proposition 4.2] that the integral in (6.5) is absolutely convergent and that it defines a bounded function of  $x \in \mathbb{R}^d$ . Since  $\omega^{(2)} \in L^1(\mathbb{R}^d)$ , it thus follows from (6.6) that

$$\int_{\mathbb{R}^d} |\omega^{(3)}(\cdot, 0, -x')| dx' \in L^\infty(\mathbb{R}^d), \quad (6.7)$$

and inserting (6.6) into (6.5), we finally arrive at the representation

$$\nabla_u \rho(x) = -\beta \rho^{(2)}(x) - \frac{\beta}{2} \int_{\mathbb{R}^d} \omega^{(3)}(x, 0, x') dx' - \beta \rho \int_{\mathbb{R}^d} \omega^{(2)}(x') dx', \quad (6.8)$$

which will be used later on.

Now, we introduce the (nonlinear) ‘‘Henderson operator’’

$$F : u \mapsto \rho^{(2)}(\mu_*(u), u), \quad (6.9)$$

which maps  $u \in \mathcal{B}_{\mathcal{Y}}(u_0)$  to its associated pair correlation function at the given density  $\rho_*$ , and which is an injective operator according to the Henderson theorem. Since  $F(u) + 2\nabla\Phi(u)$  is independent of  $u$  according to (5.11), see also (5.12), it follows that  $F : \mathcal{B}_{\mathcal{Y}}(u_0) \rightarrow L^\infty(\mathbb{R}^d)$  is differentiable for  $\rho_* < \rho_0$  with

$$\langle \tilde{v}, F'(u_0)v \rangle = -2 \Phi''(u_0)(v, \tilde{v}). \quad (6.10)$$

This shows that if the target  $\mathbb{P}_*$  has a pair correlation function  $\rho_*^{(2)} \in L^\infty(\mathbb{R}^d)$ , then the variational problem (6.2) is equivalent to the linear operator equation

$$F'(u_0)v_0 = \rho_*^{(2)} - F(u_0). \quad (6.11)$$

The step (6.3) for minimizing  $\Phi$  with Newton's method is thus equivalent to one iteration of the IMC method, where the update  $v_0$  is determined by solving (6.11). This is the thermodynamic limit analog of the observation in [29, 36], referred to in the introduction.

**Remark 6.1** As follows from Theorem 5.3 and (2.7), the Henderson operator satisfies

$$\langle v, F(u_0) \rangle = 2 E(v, \mathbb{P}_0),$$

where  $\mathbb{P}_0$  denotes the corresponding  $(\mu_*(u_0), u_0)$ -Gibbs measure. In the IMC scheme, as introduced in [26], this interaction energy is approximated by the corresponding expectation value

$$E(v, \mathbb{P}_0) \approx \frac{1}{|\Lambda|} \langle V \rangle_{\rho_*, \Lambda}$$

of the *canonical ensemble* corresponding to the pair potential  $u_0$  at density  $\rho_*$  in the bounded box  $\Lambda \subset \mathbb{R}^d$ , where  $V$  is the observable

$$V(\gamma) = \sum_{1 \leq i < j \leq N} v(x_i - x_j)$$

associated with the perturbation  $v \in \mathcal{V}$ , compare (2.3). Likewise, the derivative  $F'(u_0)$  is approximated in the IMC scheme via the cross-correlation matrix

$$-\frac{2\beta}{|\Lambda|} \left( \langle V \tilde{V} \rangle_{\rho_*, \Lambda} - \langle V \rangle_{\rho_*, \Lambda} \langle \tilde{V} \rangle_{\rho_*, \Lambda} \right) \approx \langle \tilde{v}, F'(u_0)v \rangle.$$

It has to be emphasized that the density constraint  $\rho(u) = \rho_*$  is inherently built into the canonical ensemble. Accordingly, this approximation of  $F'(u_0)$  does respect this constraint. When working in the grand canonical ensemble instead (or in the thermodynamic limit), this constraint necessitates a correction term to project the unconstrained derivative into the tangent plane of the manifold of pair correlation functions corresponding to particle systems with the correct density. This is the reason

for the first term on the right-hand side of (5.18) which does not (and need not) occur in the original IMC scheme.  $\diamond$

## 7 Extension of the Hessian to a semidefinite bilinear form on $L^2(\mathbb{R}^d)$

In the remainder of this paper, we study the Hessian of the relative entropy functional somewhat further. More precisely, we examine the mapping properties of the Jacobian of the Henderson map, which is connected to the Hessian  $\Phi''$  via (6.10).

To begin with, recall from Remark 5.4 that the right-hand side of (6.11) belongs to  $L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ , when the target  $\mathbb{P}_*$  is a  $(\mu, u_*)$ -Gibbs measure. As our first result of this section, we show that  $F'(u_0)$  has matching mapping properties.

**Theorem 7.1** *Under the assumptions of Theorem 5.3, the Jacobian  $F'(u_0)$  of the Henderson operator (6.9) belongs to  $\mathcal{L}(\mathcal{V}, L^\infty(\mathbb{R}^d)) \cap \mathcal{L}(\mathcal{V}, L^1(\mathbb{R}^d))$ .*

**Proof** We already know from Sect. 6 that  $F'(u_0) = -2A \in \mathcal{L}(\mathcal{V}, L^\infty(\mathbb{R}^d))$ . It thus remains to establish that  $F'(u_0) \in \mathcal{L}(\mathcal{V}, L^1(\mathbb{R}^d))$ . To this end, we rewrite the Henderson operator in terms of the cluster function (5.14), i.e.,

$$F(u_0) = \omega^{(2)}(\mu_*(u_0), u_0) + \rho_*^2,$$

which yields

$$F'(u_0)v = \partial_\mu \omega^{(2)}(\mu_*, u_0) \mu'_*(u_0) v + \partial_u \omega^{(2)}(\mu_*, u_0) v, \quad (7.1)$$

where  $\mu_* = \mu_*(u_0)$ . The partial derivative of the cluster function with respect to the chemical potential is given by

$$\begin{aligned} \partial_\mu \omega^{(2)}(\mu, u) &= \partial_\mu (\rho^{(2)}(\mu, u) - \rho(\mu, u)^2) \\ &= \partial_\mu \rho^{(2)}(\mu, u) - 2\rho(\mu, u) \partial_\mu \rho(\mu, u). \end{aligned} \quad (7.2)$$

From (5.20) and (6.8), we conclude that

$$\begin{aligned} \partial_\mu \rho^{(2)}(x) &= -2 \nabla_u \rho(x) \\ &= 2\beta \rho^{(2)}(x) + \beta \int_{\mathbb{R}^d} \omega^{(3)}(x, 0, x') dx' + 2\beta \rho \int_{\mathbb{R}^d} \omega^{(2)}(x') dx', \end{aligned}$$

and hence, together with Lemma A.2, it follows that

$$\partial_\mu \omega^{(2)}(x) = 2\beta \omega^{(2)}(x) + \beta \int_{\mathbb{R}^d} \omega^{(3)}(x, 0, x') dx'. \quad (7.3)$$

In particular, this shows that  $\partial_\mu \omega^{(2)} \in L^1(\mathbb{R}^d)$  by virtue of (5.17). Together with Proposition 5.2 and (7.1), we thus obtain

$$\begin{aligned} & \|F'(u_0)v\|_{L^1(\mathbb{R}^d)} \\ & \leq \left( \|\partial_\mu \omega^{(2)}(\mu_*, u_0)\|_{L^1(\mathbb{R}^d)} \|\mu'_*(u_0)\|_{\mathcal{V}'} + \|\partial_u \omega^{(2)}(\mu_*, u_0)\|_{\mathcal{L}(\mathcal{V}, L^1(\mathbb{R}^d))} \right) \|v\|_{\mathcal{V}}, \end{aligned}$$

because it has been proved in [15] that  $\partial_u \omega^{(2)} \in \mathcal{L}(\mathcal{V}, L^1(\mathbb{R}^d))$ . This shows that  $F'(u_0) \in \mathcal{L}(\mathcal{V}, L^1(\mathbb{R}^d))$ .  $\square$

Since  $L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$ , Theorem 7.1 shows that  $F'(u_0)$  is a bounded operator from  $\mathcal{V}$  to  $L^2(\mathbb{R}^d)$ . As  $\mathcal{V}$  is a dense subspace of  $L^2(\mathbb{R}^d)$ , and because we have repeatedly identified the dual pairing  $\mathcal{V} \times \mathcal{V}'$  with the dual pairing of  $L^2(\mathbb{R}^d)$ , this raises the question whether  $F'(u_0)$  extends to a bounded operator in  $L^2(\mathbb{R}^d)$ , or, in terms of the relative entropy functional, whether  $\Phi''(u_0)$  extends to a quadratic form on  $L^2(\mathbb{R}^d)$ . This will be answered to the affirmative in the remainder of this section.

**Lemma 7.2** *Under the assumptions of Theorem 5.3, the Jacobian of the Henderson map is given by*

$$(F'(u_0)v)(x) = -\beta \rho^{(2)}(x)v(x) \quad (7.4a)$$

$$- 2\beta \rho(\omega^{(2)} * v)(x) - \beta(\omega^{(2)} * \omega^{(2)} * v)(x) \quad (7.4b)$$

$$+ \frac{1}{2\partial_\mu \rho} \partial_\mu \omega^{(2)}(x) \int_{\mathbb{R}^d} \partial_\mu \omega^{(2)}(x') v(x') dx' \quad (7.4c)$$

$$- 2\beta \int_{\mathbb{R}^d} \omega^{(3)}(x, 0, x') v(x') dx' \quad (7.4d)$$

$$- \frac{\beta}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \omega^{(4)}(x, 0, x'', x' + x'') dx'' v(x') dx' \quad (7.4e)$$

for  $v \in \mathcal{V}$  and  $x \in \mathbb{R}^d$ .

**Proof** According to (6.4), we have

$$F'(u_0)v = -2Av = \partial_u \rho^{(2)}v + \frac{2}{\partial_\mu \rho} (\partial_u \rho v) \nabla_u \rho,$$

where

$$\begin{aligned} (\partial_u \rho^{(2)}v)(x) &= -\beta \rho^{(2)}(x)v(x) - 2\beta \int_{\mathbb{R}^d} \rho^{(3)}(x, 0, x') v(x') dx' \\ &\quad - \frac{\beta}{2} \int_{\mathbb{R}^d} v(x') \int_{\mathbb{R}^d} (\rho^{(4)}(x, 0, x'', x'' + x') - \rho^{(2)}(x)\rho^{(2)}(x')) dx'' dx' \end{aligned} \quad (7.5)$$

by virtue of [15, Eq. (6.9)]. Concerning the nested integral in the second line of (7.5), it has further been shown in [15, Proposition 4.3] that the inner integral converges absolutely and defines an  $L^\infty$  function of  $x$  and  $x'$ . It follows that

$$(F'(u_0)v)(x) = -\beta \rho^{(2)}(x)v(x) + \int_{\mathbb{R}^d} k(x, x')v(x') dx' \quad (7.6)$$

with integral kernel

$$\begin{aligned} k(x, x') &= \frac{2}{\partial_{\mu}\rho} \nabla_u \rho(x) \nabla_u \rho(x') - 2\beta \rho^{(3)}(x, 0, x') \\ &\quad - \frac{\beta}{2} \int_{\mathbb{R}^d} (\rho^{(4)}(x, 0, x'', x'' + x') - \rho^{(2)}(x) \rho^{(2)}(x')) dx''. \end{aligned} \quad (7.7)$$

We now rewrite the individual terms of this kernel function. To begin with, we recall from (5.20) and (7.2) that

$$\nabla_u \rho = -\frac{1}{2} \partial_{\mu} \rho^{(2)} = -\frac{1}{2} \partial_{\mu} \omega^{(2)} - \rho \partial_{\mu} \rho,$$

so that

$$\begin{aligned} \frac{2}{\partial_{\mu}\rho} \nabla_u \rho(x) \nabla_u \rho(x') &= \frac{1}{2\partial_{\mu}\rho} \partial_{\mu} \omega^{(2)}(x) \partial_{\mu} \omega^{(2)}(x') \\ &\quad + \rho \partial_{\mu} \omega^{(2)}(x) + \rho \partial_{\mu} \omega^{(2)}(x') + 2\rho^2 \partial_{\mu} \rho. \end{aligned}$$

The partial derivatives with respect to  $\mu$  in the final three terms of the right-hand side can be eliminated with the help of (7.3) and Lemma A.2: This yields

$$\begin{aligned} \frac{2}{\partial_{\mu}\rho} \nabla_u \rho(x) \nabla_u \rho(x') &= \frac{1}{2\partial_{\mu}\rho} \partial_{\mu} \omega^{(2)}(x) \partial_{\mu} \omega^{(2)}(x') \\ &\quad + 2\beta \rho \omega^{(2)}(x) + \beta \rho \int_{\mathbb{R}^d} \omega^{(3)}(x, 0, x'') dx'' + 2\beta \rho \omega^{(2)}(x') \\ &\quad + \beta \rho \int_{\mathbb{R}^d} \omega^{(3)}(x', 0, x'') dx'' + 2\beta \rho^3 + 2\beta \rho^2 \int_{\mathbb{R}^d} \omega^{(2)}(x'') dx''. \end{aligned} \quad (7.8)$$

Concerning the second term of the integral kernel (7.7), we apply definition (5.15) of the cluster functions to rewrite

$$\rho^{(3)}(x, 0, x') = \omega^{(3)}(x, 0, x') + \rho \omega^{(2)}(x) + \rho \omega^{(2)}(x - x') + \rho \omega^{(2)}(x') + \rho^3, \quad (7.9)$$

where we have also used (5.16).

In the same way, we can rewrite the integrand of the integral in (7.7) in terms of the cluster functions:

$$\begin{aligned} &\rho^{(4)}(x, 0, x'', x'' + x') - \rho^{(2)}(x) \rho^{(2)}(x') \\ &= \omega^{(4)}(x, 0, x'', x'' + x') + \omega^{(3)}(x, 0, x'') \rho + \omega^{(3)}(x, 0, x'' + x') \rho \\ &\quad + \omega^{(3)}(x, x'', x'' + x') \rho + \omega^{(3)}(0, x'', x'' + x') \rho \\ &\quad + \omega^{(2)}(x - x'') \omega^{(2)}(x'' + x') + \omega^{(2)}(x - x' - x'') \omega^{(2)}(x'') \\ &\quad + \omega^{(2)}(x'' - x) \rho^2 + \omega^{(2)}(x'' + x' - x) \rho^2 \\ &\quad + \omega^{(2)}(x'') \rho^2 + \omega^{(2)}(x'' + x') \rho^2 \end{aligned}$$

$$\begin{aligned}
&= \omega^{(4)}(x, 0, x'', x'' + x') + \omega^{(3)}(x, 0, x'') \rho + \omega^{(3)}(x, 0, x'' + x') \rho \\
&\quad + \omega^{(3)}(x', 0, x - x'') \rho + \omega^{(3)}(x', 0, -x'') \rho \\
&\quad + \omega^{(2)}(x - x'') \omega^{(2)}(x'' + x') + \omega^{(2)}(x - x' - x'') \omega^{(2)}(x'') \\
&\quad + \omega^{(2)}(x'' - x) \rho^2 + \omega^{(2)}(x'' + x' - x) \rho^2 \\
&\quad + \omega^{(2)}(x'') \rho^2 + \omega^{(2)}(x'' + x') \rho^2,
\end{aligned}$$

where we have used in the second step that the cluster functions are translation and permutation invariant. Integrating over  $x'' \in \mathbb{R}^d$ , we thus obtain

$$\begin{aligned}
&\int_{\mathbb{R}^d} (\rho^{(4)}(x, 0, x'', x'' + x') - \rho^{(2)}(x) \rho^{(2)}(x')) dx'' \\
&= \int_{\mathbb{R}^d} \omega^{(4)}(x, 0, x'', x'' + x') dx'' + 4\rho^2 \int_{\mathbb{R}^d} \omega^{(2)}(x'') dx'' \\
&\quad + 2\rho \int_{\mathbb{R}^d} \omega^{(3)}(x, 0, x'') dx'' + 2\rho \int_{\mathbb{R}^d} \omega^{(3)}(x', 0, x'') dx'' \\
&\quad + (\omega^{(2)} * \omega^{(2)})(x + x') + (\omega^{(2)} * \omega^{(2)})(x - x'),
\end{aligned} \tag{7.10}$$

where all the terms on the right-hand side are bounded functions of  $x$  and  $x'$ : This follows from (6.7) and the fact that  $\omega^{(2)} \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ .

Inserting (7.8), (7.9), and (7.10) into (7.7), we conclude that

$$\begin{aligned}
k(x, x') &= \frac{1}{2\partial_{\mu\rho}} \partial_{\mu} \omega^{(2)}(x) \partial_{\mu} \omega^{(2)}(x') - 2\beta \omega^{(3)}(x, 0, x') - 2\beta \rho \omega^{(2)}(x - x') \\
&\quad - \frac{\beta}{2} \int_{\mathbb{R}^d} \omega^{(4)}(x, 0, x'', x'' + x') dx'' \\
&\quad - \frac{\beta}{2} (\omega^{(2)} * \omega^{(2)})(x + x') - \frac{\beta}{2} (\omega^{(2)} * \omega^{(2)})(x - x').
\end{aligned}$$

Note that the final two terms of this integral kernel representation define the same convolution operator

$$v \mapsto -\frac{\beta}{2} \int_{\mathbb{R}^d} (\omega^{(2)} * \omega^{(2)})(\cdot - x') v(x') dx',$$

because every  $v \in \mathcal{V}$  is an even function. The assertion (7.4) thus follows readily from (7.6).  $\square$

Now, we can prove the main result of this section.

**Theorem 7.3** *Under the assumptions of Theorem 5.3, the operator  $F'(u_0)$  extends to a self-adjoint negative semidefinite operator on  $L^2(\mathbb{R}^d)$ .*

**Proof** We already know from Theorem 7.1 that  $F'(u_0) \in \mathcal{L}(\mathcal{V}, L^2(\mathbb{R}^d))$ . Since  $\mathcal{V}$  is a dense subset of  $L^2(\mathbb{R}^d)$ , it remains to discuss the continuity from  $L^2(\mathbb{R}^d)$  to  $L^2(\mathbb{R}^d)$  of each of the terms defined in lines (7.4a)–(7.4e) of Lemma 7.2.

For the multiplication operator in (7.4a), this is true because the pair correlation function  $\rho^{(2)}$  is bounded. The autoconvolution  $\omega^{(2)} * \omega^{(2)}$  of the cluster function  $\omega^{(2)} \in L^1(\mathbb{R}^d)$  belongs to  $L^1(\mathbb{R}^d)$  again, hence the two convolution operators in (7.4b) are bounded operators in  $L^2(\mathbb{R}^d)$ . From (7.3), it follows that  $\partial_\mu \omega^{(2)}$  belongs to  $L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ : the  $L^1$  property has been verified in the argument following (7.3); the  $L^\infty$  property of the second term in (7.3) has been established in (6.7). This shows that  $\partial_\mu \omega^{(2)} \in L^2(\mathbb{R}^d)$ , and the continuity of the operator in (7.4c) is therefore a consequence of the Cauchy–Schwarz inequality. By virtue of (5.17), the cluster function  $\omega^{(3)}(\cdot, 0, \cdot)$  belongs to  $L^1(\mathbb{R}^d \times \mathbb{R}^d)$ . Since it is bounded, it also belongs to  $L^2(\mathbb{R}^d \times \mathbb{R}^d)$ . Therefore, (7.4d) defines a compact operator in  $\mathcal{L}(L^2(\mathbb{R}^d))$ . Concerning (7.4e), we have already pointed out after (7.10) that the inner integral of  $\omega^{(4)}$  is a bounded function of  $x$  and  $x'$ . It also belongs to  $L^1(\mathbb{R}^d \times \mathbb{R}^d)$  by virtue of (5.17), again. The continuity of the operator defined in (7.4e) therefore follows in the same way as in the previous case.

In summary, this shows that  $F'(u_0)$  extends to an operator in  $\mathcal{L}(L^2(\mathbb{R}^d))$ . By virtue of (6.10), this extension is self-adjoint and negative semidefinite.  $\square$

Concerning the Hessian of the relative entropy functional, we readily conclude from (6.10) the following corollary.

**Corollary 7.4** *Under the assumptions of Theorem 5.3 the Hessian  $\Phi''(u_0)$  of the relative entropy functional (4.5) defines a symmetric and positive semidefinite bilinear form on  $L^2(\mathbb{R}^d)$ .*

As indicated before, we leave it as an open problem whether  $F'(u_0)$  is injective, i.e., whether the local quadratic approximation (6.1) of the relative entropy functional is strictly convex.

## 8 Lennard-Jones-type pair potentials

In this final section, we establish a stronger regularity result for the derivative of the Henderson map, respectively, the Hessian of the relative entropy functional, which is valid when the potential  $u_0 \in \mathcal{U}$  satisfies (2.2) with a majorant of the form

$$\psi_0(r) = C(1 + r^2)^{-\alpha/2} \quad (8.1)$$

for some  $C > 0$  and  $\alpha > d$ . This class of potentials includes the so-called potentials of Lennard-Jones type, cf. [38].

It is known that in the gas phase, the cluster function  $\omega^{(2)}$  corresponding to such a pair potential satisfies the same rate of decay near infinity as  $\psi_0$  does. This entails the following result.

**Theorem 8.1** *Let  $u_0$  satisfy (2.2) for some  $\varphi$  as in (2.1) and  $\psi = \psi_0$  of (8.1). Let the space  $\mathcal{V}$  of perturbations of  $u_0$  be defined as before, cf. (5.1), and let  $\rho_*$  belong to the gas phase of  $u_0$ . Then,  $F'(u_0) \in \mathcal{L}(\mathcal{V})$ .*

**Proof** The aforementioned result about the cluster function states that  $\omega^{(2)} \in \mathcal{V}$  under the given assumptions; compare Lemma A.4 in the appendix.

Now, we discuss the continuity of each of the operators corresponding to the individual terms in (7.4). Continuity obviously holds for the multiplication operator in (7.4a), because  $\rho^{(2)}$  is bounded. Concerning the convolutions in (7.4b), we refer to [16, Proposition 4.1] for the fact that the convolution is a continuous bilinear mapping from  $\mathcal{V} \times \mathcal{V}$  to  $\mathcal{V}$ . Accordingly, the two terms in (7.4b) also define operators in  $\mathcal{L}(\mathcal{V})$ .

For the remaining terms, we make use of Lemma A.4 again. For  $m = 3$ , this auxiliary result gives

$$\int_{\mathbb{R}^d} |\omega^{(3)}(\cdot, 0, x')| dx' \in \mathcal{V}, \quad (8.2)$$

and since

$$\left| \int_{\mathbb{R}^d} \omega^{(3)}(x, 0, x') v(x') dx' \right| \leq \|v\|_{L^\infty(\mathbb{R}^d)} \int_{\mathbb{R}^d} |\omega^{(3)}(x, 0, x')| dx'$$

for every  $x \in \mathbb{R}^d$ , it follows that

$$\left\| \int_{\mathbb{R}^d} \omega^{(3)}(\cdot, 0, x') v(x') dx' \right\|_{\mathcal{V}} \leq \|v\|_{L^\infty(\mathbb{R}^d)} \left\| \int_{\mathbb{R}^d} |\omega^{(3)}(\cdot, 0, x')| dx' \right\|_{\mathcal{V}}.$$

Much the same argument applies to the term in (7.4e):

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \omega^{(4)}(x, 0, x'', x' + x'') dx'' v(x') dx' \right| \\ & \leq \|v\|_{L^\infty(\mathbb{R}^d)} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |\omega^{(4)}(x, 0, x'', x')| dx' dx'' \end{aligned}$$

for every  $x \in \mathbb{R}^d$ , and the right-hand side again belongs to  $\mathcal{V}$  as a function of  $x$  by virtue of Lemma A.4 for  $m = 4$ . Since  $\mathcal{V}$  is continuously embedded in  $L^\infty(\mathbb{R}^d)$  this shows that the two terms in (7.4d) and (7.4e) correspond to operators in  $\mathcal{L}(\mathcal{V})$ .

Finally, concerning the term in (7.4c), we conclude from (7.3), (8.2), and the fact that  $\omega^{(2)} \in \mathcal{V}$  that  $\partial_\mu \omega^{(2)}$  belongs to  $\mathcal{V}$  as well. Accordingly, this term also represents a bounded operator from  $\mathcal{V}$  to  $\mathcal{V}$ , and the proof is done.  $\square$

Note that if the target  $\mathbb{P}_*$  and the model  $\mathbb{P}_0$  are Gibbs measures corresponding to pair potentials  $u_*$  and  $u_0$  in  $\mathcal{U}$  and their associated chemical potentials, where  $u_*$  and  $u$  satisfy (2.2) for the same majorant  $\psi = \psi_0$  given by (8.1), and if the density  $\rho_*$  belongs to the gas phase of both pair potentials, then the cluster functions  $\omega_*^{(2)}$  and  $\omega_0^{(2)}$  of  $\mathbb{P}_*$  and  $\mathbb{P}_0$ , respectively, both belong to  $\mathcal{V}$ . It thus follows from Remark 5.4 that the gradient  $\nabla \Phi(u_0)$  of the relative entropy functional belongs to  $\mathcal{V}$  as well, and therefore the Newton equation (6.11) of the IMC iterative scheme is an operator equation in  $\mathcal{V}$  by virtue of Theorem 8.1. As shown in [16, Remark 6.5], a valid choice  $u_0 \in \mathcal{U}$ , which satisfies (2.2) for the same majorant (8.1), is given by the potential of mean

force,

$$u_0 = -\frac{1}{\beta} \log(\rho_*^{(2)}/\rho_*^2) = -\frac{1}{\beta} \log(1 + \omega_*^{(2)}/\rho_*^2),$$

provided the density  $\rho_*$  is sufficiently small. Moreover,  $u_* - u_0$  belongs to  $\mathcal{V}$  for this initial guess.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Data Availability Statement** Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

## Declarations

**Conflict of interest** On behalf of both authors, the corresponding author states that there is no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix

In this appendix, we provide some auxiliary results. In doing so, we will repeatedly make use of the short-hand notation

$$\mathbf{x}_n = (x_1, \dots, x_n) \in (\mathbb{R}^d)^n.$$

The first result is needed for the proof of Theorem 4.1, where we construct a  $(\mu, u)$ -Gibbs measure for a pair potential  $u \in \mathcal{U}$  with prescribed density  $\rho_* < \rho_{\text{cp}}(u)$ .

**Lemma A.1** *Let  $u \in \mathcal{U}$  and assume that the sequence  $(\mu_k)_k \subset \mathbb{R}$  satisfies  $\mu_k \rightarrow \mu \in \mathbb{R}$  as  $k \rightarrow \infty$ . Furthermore, for every  $k \in \mathbb{N}$  let  $\mathbb{P}_k$  be a  $(\mu_k, u)$ -Gibbs measure. Then,  $(\mathbb{P}_k)_k$  has an accumulation point with respect to the local topology, and every such accumulation point is a  $(\mu, u)$ -Gibbs measure.*

**Proof** Let  $(\rho_k^{(m)})_m$  be the sequence of correlation functions of  $\mathbb{P}_k$ . Since the sequence  $(\mu_k)_k$  is bounded, the measures  $\mathbb{P}_k$  satisfy a uniform Ruelle condition by virtue of [39, Corollary 5.3], i.e., there exists  $\xi > 0$ , such that

$$\|\rho_k^{(m)}\|_{L^\infty((\mathbb{R}^d)^m)} \leq \xi^m \quad \text{for every } k, m \in \mathbb{N}. \quad (\text{A.1})$$

This implies that for fixed  $m \in \mathbb{N}$ , the sequence  $(\rho_k^{(m)})_k$  of correlation functions has a weak\* convergent subsequence in  $L^\infty((\mathbb{R}^d)^m)$  with limit  $\rho^{(m)}$ , say. With a diagonalization argument, we can thus find a subsequence of indices  $k \in \mathbb{N}$ , for which

$$\rho_k^{(m)} \xrightarrow{*} \rho^{(m)}, \quad k \rightarrow \infty, \quad (\text{A.2})$$

holds for every  $m \in \mathbb{N}$ . This is equivalent to saying that  $(\mathbb{P}_k)_k$  has a subsequence which converges to some probability measure  $\mathbb{P}$  in the local topology. To simplify the notation, we assume in the sequel that the entire sequence  $(\mathbb{P}_k)_k$  is convergent.

Since every  $\mathbb{P}_k$  is a Gibbs measure, the respective correlation functions satisfy the Kirkwood-Salsburg equations, cf., e.g., [39, Corollary 5.3]: For each  $k \in \mathbb{N}$  and every  $m \in \mathbb{N}_0$ , there holds

$$\begin{aligned} \rho_k^{(m+1)}(x_0, \mathbf{x}_m) &= e^{\beta\mu_k} e^{-\beta W(x_0; \mathbf{x}_m)} \cdot \\ &\left( \rho_k^{(m)}(\mathbf{x}_m) + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{(\mathbb{R}^d)^n} \prod_{j=1}^n f(x_0 - y_j) \rho_k^{(m+n)}(\mathbf{x}_m, \mathbf{y}_n) \, d\mathbf{y}_n \right), \end{aligned} \quad (\text{A.3})$$

where  $\rho_k^{(0)} = 1$ ,  $\mathbf{y}_n = (y_1, \dots, y_n) \in (\mathbb{R}^d)^n$ ,

$$f(x) = e^{-\beta u(x)} - 1 \quad (\text{A.4})$$

is the well-known Mayer function, and

$$W(x_0; \mathbf{x}_m) = \sum_{i=1}^m u(x_0 - x_i),$$

the latter being taken to be zero for  $m = 0$ . Note that the right-hand side of (A.3) converges in  $L^\infty((\mathbb{R}^d)^{m+1})$  by virtue of the Ruelle condition (A.1) and the fact that the Mayer function (A.4) belongs to  $L^1(\mathbb{R}^d)$ ; compare (5.3). Moreover, since the Ruelle condition is uniform in  $k$ , this convergence is also uniform with respect to  $k$ , and hence, the weak\* convergence (A.2) and the convergence  $\mu_k \rightarrow \mu$  imply that the Kirkwood-Salsburg equations (A.3) also hold true for the limiting correlation functions  $(\rho^{(m)})_m$ . It thus follows from [39, Corollary 5.3] again that  $\mathbb{P}$  is a tempered  $(\mu, u)$ -Gibbs measure. Finally, since  $\rho_k^{(m)}$  is translation invariant for every  $k \in \mathbb{N}$  and every  $m \in \mathbb{N}$ , its weak\* limits  $\rho^{(m)}$  must also be translation invariant, and hence  $\mathbb{P}$  is translation invariant.  $\square$

The remaining two results of the appendix concern properties of the cluster functions (see Remark 5.6) within the gas phase. The proofs make use of their *cluster expansions* (compare [38, Sect. 4.4] and [43])

$$\omega^{(m)}(\mathbf{x}_m) = \sum_{n=m}^{\infty} \frac{e^{n\beta\mu}}{(n-m)!} \int_{(\mathbb{R}^d)^{n-m}} \phi^{(n)}(\mathbf{x}_n) \, dx_{m+1} \cdots dx_n \quad (\text{A.5})$$

for  $\mathbf{x}_m \in (\mathbb{R}^d)^m$  with

$$\phi^{(n)}(\mathbf{x}_n) = \sum_{\mathcal{C} \in \mathfrak{C}_n} \prod_{(i,j) \in \mathcal{C}} f(x_i - x_j), \quad (\text{A.6})$$

where  $f$  is, again, the Mayer function (A.4), and  $\mathfrak{C}_n$  denotes the set of undirected connected graphs with vertices  $\{1, \dots, n\}$  and edges  $(i, j) \in \mathcal{C}$ , which connect vertices  $i$  and  $j$ ; for  $n = m$  the integral in (A.5) has to be replaced by the value  $\phi^{(m)}(\mathbf{x}_m)$  of its integrand. These cluster expansions are absolutely convergent for  $u_0 \in \mathcal{U}$  and  $\mu \leq \mu_0$  with  $\mu_0$  of (5.6).

**Lemma A.2** *Let  $u_0 \in \mathcal{U}$  and  $\mu < \mu_0$  with  $\mu_0$  as in (5.6). Then, the density  $\rho$  of the corresponding Gibbs measure is differentiable with respect to  $\mu$  and the derivative is given by*

$$\partial_\mu \rho = \beta \rho + \beta \int_{\mathbb{R}^d} \omega^{(2)}(x) dx. \quad (\text{A.7})$$

**Proof** By (A.5) and (5.15), there holds (with  $x_1 = 0$ )

$$\rho = \omega^{(1)}(0) = \sum_{n=1}^{\infty} \frac{e^{n\beta\mu}}{(n-1)!} \sum_{\mathcal{C} \in \mathfrak{C}_n} \int_{(\mathbb{R}^d)^{n-1}} \prod_{(i,j) \in \mathcal{C}} f(x_i - x_j) dx_2 \cdots dx_n.$$

Differentiating with respect to  $\mu$ , we obtain

$$\begin{aligned} \partial_\mu \rho &= \sum_{n=1}^{\infty} \frac{n\beta e^{n\beta\mu}}{(n-1)!} \sum_{\mathcal{C} \in \mathfrak{C}_n} \int_{(\mathbb{R}^d)^{n-1}} \prod_{(i,j) \in \mathcal{C}} f(x_i - x_j) dx_2 \cdots dx_n \\ &= \beta \sum_{n=1}^{\infty} \frac{e^{n\beta\mu}}{(n-1)!} \sum_{\mathcal{C} \in \mathfrak{C}_n} \int_{(\mathbb{R}^d)^{n-1}} \prod_{(i,j) \in \mathcal{C}} f(x_i - x_j) dx_2 \cdots dx_n \\ &\quad + \beta \int_{\mathbb{R}^d} \left( \sum_{n=2}^{\infty} \frac{e^{n\beta\mu}}{(n-2)!} \sum_{\mathcal{C} \in \mathfrak{C}_n} \int_{(\mathbb{R}^d)^{n-2}} \prod_{(i,j) \in \mathcal{C}} f(x_i - x_j) dx_3 \cdots dx_n \right) dx_2 \\ &= \beta \rho + \beta \int_{\mathbb{R}^d} \omega^{(2)}(0, x_2) dx_2. \end{aligned}$$

Using the short-hand notation for  $\omega^{(2)}$  and (5.16), the assertion follows.  $\square$

**Remark A.3** Lemma A.2 is general folklore, at least its “grand canonical version”

$$\partial_\mu \rho_\Lambda(0) = \beta \rho_\Lambda(0) + \beta \int_{\Lambda} \left( \rho_\Lambda^{(2)}(0, x) - \rho_\Lambda(0) \rho_\Lambda(x) \right) dx \quad (\text{A.8})$$

for a bounded domain  $\Lambda \subset \mathbb{R}^d$  is well-known, where  $\rho_\Lambda$  and  $\rho_\Lambda^{(2)}$  denote the corresponding grand canonical correlation functions; compare, e.g., [43, Eq. (2-9)], or Ben-Naim [1, Eq. (1.53)]. An alternative proof of (A.7) is thus obtained by passing in (A.8) to the thermodynamic limit.  $\diamond$

Our final lemma strenghtens Ruelle's result (5.17) for Lennard-Jones-type pair potentials, and extends [16, Corollary 5.2] to higher-order cluster functions.<sup>2</sup>

**Lemma A.4** *Let  $u \in \mathcal{U}$  be a pair potential which satisfies (2.1), (2.2) with majorant  $\psi = \psi_0$  as in (8.1), and let  $\omega^{(m)}$ ,  $m \in \mathbb{N}$ , be the associated cluster functions. Further, let the chemical potential satisfy  $\mu \leq \mu_0$ , cf. (5.6). Then,  $\omega^{(2)} \in \mathcal{V}$ , and for  $m \geq 3$ , there holds*

$$\int_{(\mathbb{R}^d)^{m-2}} |\omega^{(m)}(\cdot, 0, x_3, \dots, x_m)| dx_2 \cdots dx_m \in \mathcal{V}.$$

**Proof** For  $\mu < \mu_0$ , it follows from [16, Lemma 5.1] that  $\phi^{(n)}$  of (A.6) satisfies

$$\int_{(\mathbb{R}^d)^{n-2}} |\phi^{(n)}(\mathbf{x}_n)| dx_3 \cdots dx_n \leq C_\Sigma (c_\beta e^{\beta B})^n n^{n-2} \psi_0(|x_1 - x_2|)$$

for some constant  $C_\Sigma > 0$ , every  $x_1, x_2 \in \mathbb{R}^d$  and every  $n \geq 2$ , where  $c_\beta$  and  $B$  are the constants defined in (5.3), (5.4); for  $n = 2$  the left-hand side of this inequality has to be replaced by

$$|\phi^{(2)}(x_1, x_2)| = |f(x_1 - x_2)|.$$

Plugging this into the cluster expansion (A.5) for  $m \geq 3$ , it follows that

$$\begin{aligned} \int_{(\mathbb{R}^d)^{m-2}} |\omega^{(m)}(\mathbf{x}_m)| dx_3 \cdots dx_m &\leq C_\Sigma \psi_0(|x_1 - x_2|) \sum_{n=m}^{\infty} \frac{n^{n-2}}{(n-m)!} (c_\beta e^{\beta(B+\mu)})^n \\ &\leq C_\Sigma \psi_0(|x_1 - x_2|) \sum_{n=m}^{\infty} n^{m-2} (c_\beta e^{\beta(B+\mu)+1})^n, \end{aligned}$$

where we have used the inequality  $n^{n-m}/(n-m)! \leq e^n$  in the final step. Since  $\mu \leq \mu_0$ , the positive number  $q = c_\beta e^{\beta(B+\mu)+1}$  is below one, and hence the series on the right-hand side converges, its value being  $C_{\mu,m}$ , say. This estimate extends to  $m = 2$  when the left-hand side is replaced by  $|\omega^{(2)}(\mathbf{x}_2)|$ . The assertion therefore follows from (5.1).  $\square$

We refer to [3, 4, 27] for further results on the decay of the higher-order cluster functions.

## References

1. Ben-Naim, A.: Molecular Theory of Solutions. Oxford University Press, New York (2006)
2. Chayes, J.T., Chayes, L.: On the validity of the inverse conjecture in classical density functional theory. J. Stat. Phys. **36**, 471–488 (1984)

<sup>2</sup> The results in [16], like those in [14, 15], extend to the present setting.

3. Dorlas, T.C., Rebenko, A.L., Savoie, B.: Correlation of clusters: partially truncated correlation functions and their decay. *J. Math. Phys.* **61**, 033303 (2020)
4. Duneau, M., Iagolnitzer, D., Souillard, B.: Decay of correlations for infinite-range interactions. *J. Math. Phys.* **16**, 1662–1666 (1975)
5. Ercolessi, F., Adams, J.B.: Interatomic potentials from first-principles calculations: the force-matching method. *Europhys. Lett.* **26**, 583–588 (1994)
6. Fritsch, S., Junghans, C., Kremer, K.: Structure formation of Toluene around C60: implementation of the adaptive resolution scheme (AdResS) into GROMACS. *J. Chem. Theor. Comput.* **8**, 398–403 (2012)
7. Frommer, F., Hanke, M., Jansen, S.: A note on the uniqueness result for the inverse Henderson problem. *J. Math. Phys.* **60**, 093303 (2019)
8. Gallavotti, G., Miracle-Sole, S.: A variational principle for the equilibrium of hard sphere systems. *Ann. IHP Phys. théor.* **8**, 287–299 (1968)
9. Georgii, H.-O.: Large deviations and the equivalence of ensembles for Gibbsian particle systems with superstable interaction. *Probab. Theory Related Fields* **99**, 171–195 (1994)
10. Georgii, H.-O.: The equivalence of ensembles for classical systems of particles. *J. Stat. Phys.* **80**, 1341–1378 (1995)
11. Georgii, H.-O.: *Gibbs Measures and Phase Transitions*. de Gruyter, Berlin (2011)
12. Georgii, H.-O.: *Stochastics: Introduction to Probability and Statistics*, 2nd edn. de Gruyter, Berlin (2013)
13. Georgii, H.-O., Zessin, H.: Large deviations and the maximum entropy principle for marked point random fields. *Probab. Theory Related Fields* **96**, 177–204 (1993)
14. Hanke, M.: Fréchet differentiability of molecular distribution functions I:  $L^\infty$  analysis. *Lett. Math. Phys.* **108**, 285–306 (2018)
15. Hanke, M.: Fréchet differentiability of molecular distribution functions II: the Ursell function. *Lett. Math. Phys.* **108**, 307–329 (2018)
16. Hanke, M.: Well-Posedness of the iterative Boltzmann inversion. *J. Stat. Phys.* **170**, 536–553 (2018)
17. Hansen, J.-P., McDonald, I.R.: *Theory of Simple Liquids: With Applications to Soft Matter*, 4th edn. Academic Press, Oxford (2013)
18. Henderson, R.L.: A uniqueness theorem for fluid pair correlation functions. *Phys. Lett. A* **49**, 197–198 (1974)
19. Hughes, A.M.: *Entropy minimization, convergence and Gibbs ensembles (local and global)*. University of Missouri, PhD-Thesis (2021)
20. Izvekov, S., Voth, G.A.: Multiscale coarse graining of liquid-state systems. *J. Chem. Phys.* **123**, 134105 (2005)
21. Jansen, S.: Cluster expansions for Gibbs point processes. *Adv. Appl. Probab.* **51**, 1129–1178 (2019)
22. Jansen, S., Kuna, T., Tsagkarogiannis, D.: Virial inversion and density functionals, [arXiv:1906.02322](https://arxiv.org/abs/1906.02322) [math-ph] (2019)
23. Korolov, L.: An inverse problem for Gibbs fields with hard core potential. *J. Math. Phys.* **48**, 053301 (2007)
24. Kuna, T., Lebowitz, J., Speer, E.: Realizability of point processes. *J. Stat. Phys.* **129**, 417–439 (2007)
25. Larini, L., Lu, L., Voth, G.A.: The multiscale coarse-graining method. VI. Implementation of three-body coarse-grained potentials. *J. Chem. Phys.* **132**, 164107 (2010)
26. Lyubartsev, A.P., Laaksonen, A.: Calculation of effective interaction potentials from radial distribution functions: a reverse Monte Carlo approach. *Phys. Rev. E* **52**, 3730–3737 (1995)
27. Minlos, R.A., Poghosyan, S.: Estimates of Ursell functions, group functions and their derivatives. *Theor. Math. Phys.* **31**, 408–418 (1977)
28. Monticelli, L., Salonen, E. (eds.): *Biomolecular Simulations. Methods and Protocols*. Springer, New York (2013)
29. Murtola, T., Karttunen, M., Vattulainen, I.: Systematic coarse graining from structure using internal states: application to phospholipid/cholesterol bilayer. *J. Chem. Phys.* **131**, 055101 (2009)
30. Navrotskaya, I.: *Inverse problem in classical statistical mechanics*. PhD thesis, University of Pittsburgh (2016)
31. Noid, W.G.: Perspective: coarse-grained models for biomolecular systems. *J. Chem. Phys.* **139**, 090901 (2013)
32. Noid, W.G.: Systematic methods for structurally consistent coarse-grained models. In [28], pp. 487–531 (2013)

33. Peter, C., Kremer, K.: Multiscale simulation of soft matter systems - from the atomistic to the coarse-grained level and back. *Soft Matter* **5**, 4357–4366 (2009)
34. Praprotnik, M., Delle Site, L., Kremer, K.: Adaptive resolution molecular-dynamics simulation: changing the degrees of freedom on the fly. *J. Chem. Phys.* **123**, 224106 (2005)
35. Reith, D., Pütz, M., Müller-Plathe, F.: Deriving effective mesoscale potentials from atomistic simulations. *J. Comput. Chem.* **24**, 1624–1636 (2003)
36. Rosenberger, D., Sanyal, T., Shell, M.S., van der Vegt, N.F.A.: Transferability of local density-assisted implicit solvation models for homogeneous fluid mixtures. *J. Chem. Theor. Comput.* **15**, 2881–2895 (2019)
37. Rühle, V., Junghans, C., Lukyanov, A., Kremer, K., Andrienko, D.: Versatile object-oriented toolkit for coarse-graining applications. *J. Chem. Theor. Comput.* **5**, 3211–3223 (2009)
38. Ruelle, D.: *Statistical Mechanics: Rigorous Results*. W.A. Benjamin Publ, New York (1969)
39. Ruelle, D.: Superstable interactions in classical statistical mechanics. *Commun. Math. Phys.* **18**, 127–159 (1970)
40. Schmid, F.: Editorial: Multiscale simulation methods for soft matter systems. *J. Phys.: Condens. Matter* **34**, 160401 (2022)
41. Shell, M.S.: The relative entropy is fundamental to multiscale and inverse thermodynamic problems. *J. Chem. Phys.* **129**, 144108 (2008)
42. Soper, A.K.: Empirical potential Monte Carlo simulation of fluid structure. *Chem. Phys.* **202**, 295–306 (1996)
43. Stell, G.: Cluster expansions for classical systems in equilibrium. In: Fritsch, H.L., Lebowitz, J.L. (eds.) *The Equilibrium Theory of Classical Fluids*, pp. II-171–II-266. W.A. Benjamin Publ., New York (1964)
44. Tóth, G.: Interactions from diffraction data: historical and comprehensive overview of simulation assisted methods. *J. Phys.: Condens. Matter* **19**, 335220 (2007)
45. Tsourtis, A., Harmandaris, V., Tsagkarogiannis, D.: Parameterization of coarse-grained molecular interactions through potential of mean force calculations and cluster expansion techniques. *Entropy* **19**, 395 (2017)
46. Wang, Y., Noid, W.G., Liu, P., Voth, G.A.: Effective force coarse-graining. *Phys. Chem. Chem. Phys.* **11**, 2002–2015 (2009)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

# 5

## An Inverse Cluster Expansion for the Chemical Potential

### 5.1 Overview

The goal of this chapter is to find an explicit expansion of the chemical potential in terms of functions depending on the correlation functions. The ansatz we use was first developed by Nettleton and Green in [45] for a grand-canonical ensemble. Their idea is to make use of (2.8) and develop the correlation functions in a cluster expansion similar to (2.43), using certain *n-body potentials of mean-force*. However, they were not able to show convergence of their expansion. We instead rely on an approach using Theorem 2.1.21 and a representation

$$j_{\Lambda}^{(1)}(x) = \rho \exp \left( \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \int_{(\mathbb{R}^d)^n} F_1^{(n)}(x, \mathbf{y}_n) d\mathbf{y}_n \right).$$

By Remark 2.1.23 we know that any  $(\beta, \mu, u)$ -Gibbs measure satisfies Assumption A, from which we can find bounds of the type (2.10) for  $(F_1^{(n)})_{n \geq 1}$ . We get the following result.

**Theorem 5.1.1** *Let  $\beta > 0$ ,  $\mu \in \mathbb{R}$ ,  $u \in \mathcal{U}$  and  $\mathbf{P}$  be a  $(\beta, \mu, u)$ -Gibbs measure. If  $D_{\rho}$  from Assumption A is small enough there holds*

$$\mu = \frac{1}{\beta} \left( \log \rho + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \int_{(\mathbb{R}^d)^n} \tilde{\rho}_T^{(1+n)}(0, \mathbf{y}_n) d\mathbf{y}_n \right) \quad (5.1)$$

where the family  $(\tilde{\rho}_T^{(1+n)})_{n \geq 1}$  is recursively defined by  $\tilde{\rho}_T^{(2)}(x, y) = \rho_T^{(2)}(x, y)/\rho$  and for  $n \geq 2$  by

$$\tilde{\rho}_T^{(1+n)}(x, \mathbf{y}_n) = \frac{\rho_T^{(1+n)}(x, \mathbf{y}_n)}{\rho} - \sum_{l=2}^n \sum_{\pi \in \Pi_l(\{\mathbf{y}_n\})} \prod_{i=1}^l \tilde{\rho}_T^{(1+\kappa_i)}(x, \pi_i). \quad (5.2)$$

**Remark 5.1.2** *It can be shown, see Section 4.3-4.5 in [55], that  $D_{\rho} = D_{\rho}(\mu)$  is decreasing in  $\mu$  and thus  $D_{\rho}$  is sufficiently small for  $\mu \ll 1$ .*

## 5.2 Reprint

The article is reprinted by permission from Springer Nature: *Journal of Statistical Physics: An Inverse Cluster Expansion for the Chemical Potential*, Fabio Frommer, 2024, 191, 112, 3 September 2024 Springer Science+Business Media ([doi.org/10.1007/s10955-024-03319-3](https://doi.org/10.1007/s10955-024-03319-3)).



## An Inverse Cluster Expansion for the Chemical Potential

Fabio Frommer<sup>1</sup>

Received: 20 December 2023 / Accepted: 5 August 2024 / Published online: 3 September 2024  
© The Author(s) 2024

### Abstract

Interacting particle systems in a finite-volume in equilibrium are often described by a grand-canonical ensemble induced by the corresponding Hamiltonian, i.e. a finite-volume Gibbs measure. However, in practice, directly measuring this Hamiltonian is not possible, as such, methods need to be developed to calculate the Hamiltonian potentials from measurable data. In this work, we give an expansion of the chemical potential in terms of the correlation functions of such a system in the thermodynamic limit. This is a justification of a formal approach of Nettleton and Green from the 50's, that can be seen as an inverse cluster expansion.

**Keywords** Cluster expansion · Pressure · Gibbs point processes · Chemical potential · Statistical mechanics

### 1 Introduction

In classical equilibrium statistical physics, interacting particle systems are often described via so-called Gibbs measures, either in a finite or infinite-volume. A possible arrangement of particles (also called configuration) has a higher probability when it is energetically favorable. This energy gets calculated via a Hamiltonian that depends on the chemical potential and the interaction potentials associated to the particles. However, when working with actual data, these potentials are impossible to measure. What instead can be measured or extrapolated from simulation data are the so-called correlation functions of the underlying Gibbs measure. This gives rise to the inverse problem: Given the correlation functions, determine the chemical potential and the interaction potentials.

For the chemical potential that induces a given density this problem was first discussed by Chayes and Chayes [1] and Chayes et al. [2]. The case of higher-order interactions is also briefly discussed.

For the pair-interaction case the inverse problem was investigated in [3] in the case of the finite-volume and in [4, 5] these results were extended upon and made rigorous in the thermodynamic limit, i.e. in infinite-volume. However, all these results do not give a constructive way of recovering the interaction. Estimation methods for the pair-interaction were developed by Takacs [6] and Fiksel [7] and have since been expanded upon for more general

Communicated by Aernout van Enter.

✉ Fabio Frommer  
fabiofrommer@uni-mainz.de

<sup>1</sup> Institut für Mathematik, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany

Hamiltonians in e.g. [8], another approach based on variational methods was developed in [9], however, it cannot be used to estimate the chemical potential. A first result of an inverse expansion of the chemical potential is found in [10] where the expansion is given in terms of the one-point correlation function (the density) and the Mayer function, see also [11] for an extension of this result to the non-homogeneous case. For this formula however prior knowledge of the pair-interaction is required. The goal of this paper is to find an expansion of the chemical potential, that solely depends on functions derived from the correlation functions of the Gibbs measure.

The idea we use goes back to [12], where an expansion of this type is used in finite-volume. However, the ansatz used in these works constructs multi-body interactions and thus showing convergence of these inversion formulas cannot be done using the classical tools of cluster expansions.

In this work, we will use a different approach to prove convergence for the explicit inversion formula for the chemical potential by using an exponential representation formula. We then use Bell-polynomials to find bounds on the appearing coefficients, which have been used for a similar scope independently in [13] to get bounds on the virial coefficients. In fact, we can show that this formula converges in both finite and infinite-volume. In particular, as we only make very mild assumptions on the Hamiltonian, which include many classic models, and some integrability conditions on the correlation functions this expansion holds not only for pair-interactions, but also in some cases of multi-body interactions of arbitrary order.

The outline is as follows: In Sect. 2 we introduce the general setting of Gibbs measures we work with and introduce the main idea of the expansion. Then, we formulate in Sect. 3 our main assumptions as well as the main result and give a few examples in which these assumptions hold in Sect. 4. The last two Sections are devoted to the proof of the main result, in Sect. 5 we start by proving a result about the zero-point density, which then gets generalized in Sect. 6 to the one-point density to obtain the expansion for the chemical potential.

## 2 Setting

A point process  $\mathbb{P}$  is a probability measure on the set of configurations

$$\Gamma = \{\eta \subset \mathbb{R}^d \mid N(\eta \cap \Delta) < \infty \text{ for all } \Delta \subset \mathbb{R}^d \text{ bounded}\}$$

equipped with the  $\sigma$ -algebra  $\mathcal{F} := \sigma(\{N(\cdot \cap \Delta) = m \mid m \in \mathbb{N}_0, \Delta \subset \mathbb{R}^d \text{ bounded}\})$ . Here  $N(\eta) = \#\eta$  is the number of elements of  $\eta \subset \mathbb{R}^d$ . We denote by  $\Gamma_0 = \{\gamma \in \Gamma \mid N(\gamma) < \infty\}$  the set of finite configurations. Such a point process is a Gibbs *measure* for some chemical potential  $\mu \in \mathbb{R}$  and a Hamiltonian  $H: \Gamma \rightarrow \mathbb{R} \cup \{\infty\}$ , we write  $\mathbb{P}$  is a  $(\mu, H)$ -Gibbs measure, if it is supported on a set of *tempered* configurations, see e.g. [14] for a discussion, and satisfies the *Ruelle-equation*, see [15] (herein called *system of equilibrium equations*), namely if, for every non-negative function  $G: \Gamma \rightarrow [0, \infty]$  and every bounded set  $\Lambda \subset \mathbb{R}^d$ , there holds

$$\int_{\Gamma} G(\eta) \, d\mathbb{P}(\eta) = \int_{\Gamma_{\Lambda^c}} \sum_{n=0}^{\infty} \frac{e^{n\mu}}{n!} \int_{\Lambda^n} G(\{\mathbf{x}_n\} \cup \eta) e^{-H(\{\mathbf{x}_n\}) - W(\{\mathbf{x}_n\}|\eta)} \, d\mathbf{x}_n \, d\mathbb{P}(\eta)$$

where  $\Lambda^c = \mathbb{R}^d \setminus \Lambda$ ,  $\Gamma_{\Lambda^c} = \{\eta \in \Gamma \mid \eta \subset \Lambda^c\}$  and  $W(\{\mathbf{x}_n\} | \eta)$  is the interaction between the part of the particles inside  $\Lambda$  and those on the outside. If  $H$  has finite-range  $W$  can be

defined as

$$W(\{\mathbf{x}_n \mid \eta\}) := H(\{\mathbf{x}_n\} \cup \eta) - H(\{\mathbf{x}_n\}) - H(\eta), \quad \mathbf{x}_n \in \Lambda^n, \eta \in \Gamma_{\Lambda^c}$$

with the convention that we set  $W(\{\mathbf{x}_n \mid \eta\}) = +\infty$  if  $H(\{\mathbf{x}_n\} \cup \eta) = +\infty$ . For general Hamiltonians it is possible to extend this definition, cf. [14]. However, the specific definition of  $W$  is not important for this work. We also assume that the Hamiltonian  $H$  is *translation-invariant*, i.e. for every  $x \in \mathbb{R}^d$  we have that  $H(\eta) = H(\tau_x(\eta))$  where  $\tau_x(\eta) := \{y - x \mid y \in \eta\}$ , that  $H$  is *hereditary*, i.e. that  $H(\eta \cup \{x\}) = +\infty$  for all  $x \in \mathbb{R}^d$  whenever  $H(\eta) = +\infty$  and that  $H$  does not contain any contribution of a self-interaction of particles, i.e.  $H(\{x\}) = 0$  for all  $x \in \mathbb{R}^d$ . The last assumption we make on  $H$  is, that  $H$  is *stable*, i.e. there is a constant  $B > 0$  such that

$$H(\{\mathbf{x}_n\}) \geq -Bn.$$

In this case, under some additional technical assumptions, e.g. finite-range of the interaction or the Hamiltonian consists of a regular pair-interaction, there is at least one translation-invariant  $(\mu, H)$ -Gibbs measure  $\mathbb{P}$  associated to these parameters, cf. [14]. In the sequel, we will assume that  $\mathbb{P}$  is such a translation-invariant Gibbs measure. A family of functions  $(j_{\Lambda}^{(n)})_{n \geq 0, \Lambda \subset \mathbb{R}^d \text{ bounded}}$  are called the *finite-volume densities* of  $\mathbb{P}$ , or *Janossy densities*, if for every bounded set  $\Lambda \subset \mathbb{R}^d$  and every non-negative function  $G: \Gamma \rightarrow [0, \infty]$  there holds

$$\int_{\Gamma} G(\eta \cap \Lambda) \, d\mathbb{P}(\eta) = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\Lambda^n} G(\{\mathbf{x}_n\}) j_{\Lambda}^{(n)}(\mathbf{x}_n) \, dx_k.$$

The Ruelle-equation characterizes the Janossy densities, they exist for every  $(\mu, H)$ -Gibbs measure  $\mathbb{P}$  and for fixed  $\Lambda \subset \mathbb{R}^d$  they are given by

$$j_{\Lambda}^{(n)}(\mathbf{x}_n) = \int_{\Gamma_{\Lambda^c}} \exp\left(n\mu - H(\{\mathbf{x}_n\}) - W(\{\mathbf{x}_n\} \mid \eta)\right) \, d\mathbb{P}(\eta), \quad n \in \mathbb{N}_0. \quad (2.1)$$

Note that by definition the Janossy densities are *symmetric* functions, i.e. for every permutation  $\sigma \in \mathcal{S}_n$  there holds  $j_{\Lambda}^{(n)}(x_1, \dots, x_n) = j_{\Lambda}^{(n)}(x_{\sigma(1)}, \dots, x_{\sigma(n)})$ . Another family of symmetric functions we are going to need are the *correlation functions*  $(\rho^{(n)})_{n \geq 0}$ , where  $\rho^{(0)} := 1$ ; when a point process has Janossy densities they are known to exist and for  $\mathbf{x}_n \in (\mathbb{R}^d)^n$  the  $n$ -point correlation function is given by

$$\rho^{(n)}(\mathbf{x}_n) = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\Lambda^n} j_{\Lambda}^{(n+k)}(\mathbf{x}_n, \mathbf{y}_k) \, d\mathbf{y}_k,$$

for any  $\Lambda$  such that  $\mathbf{x}_n \in \Lambda^n$ . Note that we write  $j_{\Lambda}^{(n+k)}(\mathbf{x}_n, \mathbf{y}_k) = j_{\Lambda}^{(n+k)}(x_1, \dots, x_n, y_1, \dots, y_k)$  for simplicity and that the Janossy densities contain the averaged contributions of the whole space by Eq. (2.1) and thus the right-hand side above is independent of  $\Lambda$ , if  $\mathbf{x}_n \in \Lambda^n$ . In particular, if  $\mathbb{P}$  is translation-invariant, then  $\rho^{(1)} = \rho$  is constant. Under the additional assumption that  $\mathbb{P}$  satisfies a so-called *Ruelle-condition*, i.e. there is a  $\xi > 0$  such that

$$0 \leq \rho^{(n)}(\mathbf{x}_n) \leq \xi^n \quad \text{for all } \mathbf{x}_n \in (\mathbb{R}^d)^n, \quad n \geq 1$$

there also holds the well-known inverse formula

$$j_{\Lambda}^{(n)}(\mathbf{x}_n) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{\Lambda^k} \rho^{(n+k)}(\mathbf{x}_n, \mathbf{y}_k) \, d\mathbf{y}_k, \quad \mathbf{x}_n \in \Lambda^n, \quad (2.2)$$

see e.g. [15].

**Remark 2.1** In experiments, when only having samples of a Gibbs measure, it is straightforward to estimate the correlation functions of the process without having to make prior assumptions on the Hamiltonian. Thus, the correlation functions are the experimentally easier to obtain starting object in contrast to the Janossy densities.

The idea of this work is to find an explicit expansion of  $\log j_\Lambda^{(1)}$ , using the identity Eq. (2.2), in terms that only depend on the correlation functions  $(\rho^{(n)})_{n \geq 0}$ . From this, we can then obtain an explicit formula for the chemical potential  $\mu$ . The reason is as follows:

By Eq. (2.1) there holds

$$\log j_\Lambda^{(1)}(x) = \mu + \log \int_{\Gamma_{\Lambda^c}} \exp(-W(\{x\} | \eta)) \, d\mathbb{P}(\eta).$$

Now it is well-known that both sides of the above equation diverge to  $-\infty$  for  $\Lambda \nearrow \mathbb{R}^d$ . In the sequel we will make the following technical assumption, that can be interpreted as an implicit decay condition on the Hamiltonian and in particular is obvious for finite-range potentials,

$$\log \int_{\Gamma_{\Lambda^c}} \exp(-W(\{x\} | \eta)) \, d\mathbb{P}(\eta) = \log \int_{\Gamma_{\Lambda^c}} d\mathbb{P}(\eta) + \varepsilon(\Lambda),$$

where  $\varepsilon(\Lambda) \rightarrow 0$  as  $\Lambda \nearrow \mathbb{R}^d$ . The first part of the right-hand side is equal to  $\log j_\Lambda^{(0)}$  and thus  $\log j_\Lambda^{(1)}(x) - \log j_\Lambda^{(0)} \rightarrow \mu$  as  $\Lambda \nearrow \mathbb{R}^d$ .

Our main tool is a result about exponential representation that has been used by Ruelle in [16, Sect. 4.4] for the correlation functions, see also [17] for a more general version, which we will state here for ease of reading.

**Theorem A** (Exponential representation [17]) *Let  $(F_n)_{n \geq 0}$  be a family of symmetric functions with  $F_n: (\mathbb{R}^d)^n \rightarrow \mathbb{R}$ ,  $F_0 \equiv 1$  and there exist  $0 < c < 1/2$  and  $D > 0$  such that for every bounded  $\Lambda \subset \mathbb{R}^d$  there holds*

$$\int_{\Lambda^n} |F_n(\mathbf{x}_n)| \, d\mathbf{x}_n \leq |\Lambda| n! D c^n \quad (2.3)$$

for every  $n \in \mathbb{N}$ , here  $|\Lambda|$  denotes the Lebesgue measure of the set  $\Lambda$ . Then the function  $\Phi: \Gamma_0 \rightarrow \mathbb{R}$  defined by

$$\Phi(\eta) = \sum_{k=1}^{N(\eta)} \sum_{\pi \in \Pi_k(\eta)} \prod_{i=1}^k F_{\kappa_i}(\pi_i), \quad (2.4)$$

where  $\Pi_k(\eta)$  denotes the set of partitions of  $\eta$  into  $k$  non-empty sets  $\pi_i$   $i = 1, \dots, k$  where  $\pi_i$  consists of  $\kappa_i$  elements, satisfies

$$\sum_{n=0}^{\infty} \frac{1}{n!} \int_{\Lambda^n} \Phi(\mathbf{x}_n) \, d\mathbf{x}_n = \exp \left( \sum_{n=1}^{\infty} \frac{1}{n!} \int_{\Lambda^n} F_n(\mathbf{x}_n) \, d\mathbf{x}_n \right). \quad (2.5)$$

The last tool we will need are the so-called *truncated correlation functions*.

**Definition 2.2** For a point process with correlation functions  $(\rho^{(n)})_{n \geq 0}$ , the *truncated correlation functions*  $(\rho_T^{(n)})_{n \geq 1}$  are defined recursively. For  $n = 1$  we define  $\rho_T^{(1)} = \rho^{(1)} = \rho$  and for  $n \geq 2$

$$\rho_T^{(n)}(\mathbf{x}_n) = \rho^{(n)}(\mathbf{x}_n) - \sum_{k=2}^n \sum_{\pi \in \Pi_k(\{\mathbf{x}_n\})} \prod_{i=1}^k \rho_T^{(\kappa_i)}(\pi_i). \quad (2.6)$$

### 3 Main Result

We will now state our main result, which will be proved in Sect. 6. To formulate our result we introduce the following two assumptions:

**Assumption A** Let  $W$  be the interaction associated to the Hamiltonian  $H$ . We assume for any bounded set  $\Delta \subset \mathbb{R}^d$  and every  $(\mu, H)$ -Gibbs measure  $\mathbb{P}$  there holds

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \sup_{x \in \Delta} \left| \log \int_{\Gamma_{\Lambda^c}} \exp(-W(\{x\} | \eta)) d\mathbb{P}(\eta) - \log \int_{\Gamma_{\Lambda^c}} d\mathbb{P}(\eta) \right| = 0.$$

**Assumption B** There are constants  $D > 0$  and  $q > 0$  such that for the truncated correlation functions of  $\mathbb{P}$  there holds

$$\sup_{x \in \mathbb{R}^d} \int_{\Lambda^n} \frac{1}{\rho} \left| \rho_T^{(1+n)}(x, \mathbf{y}_n) \right| d\mathbf{y}_n \leq n! D q^n, \quad (3.1)$$

for every bounded set  $\Lambda \subset \mathbb{R}^d$  and every  $n \in \mathbb{N}$ .

**Theorem 3.1** Let  $\mathbb{P}$  be a  $(\mu, H)$ -Gibbs measure that satisfies a Ruelle-condition and Assumptions A and B. If  $q < q_0$ , where

$$q_0 = \frac{1}{2(2 + \zeta D)} \quad \text{with} \quad \zeta = \frac{1}{2 \log 2 - 1}, \quad (3.2)$$

then there holds

$$\mu = \log \rho + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{(\mathbb{R}^d)^k} \tilde{\rho}_T^{(1+k)}(0, \mathbf{y}_k) d\mathbf{y}_k \quad (3.3)$$

where the family  $(\tilde{\rho}_T^{(1+k)})_{k \geq 1}$  is recursively defined by  $\tilde{\rho}_T^{(2)}(x, y) = \rho_T^{(2)}(x, y)/\rho$  and for  $k \geq 2$  by

$$\tilde{\rho}_T^{(1+k)}(x, \mathbf{y}_k) = \frac{\rho_T^{(1+k)}(x, \mathbf{y}_k)}{\rho} - \sum_{l=2}^k \sum_{\pi \in \Pi_l(\{\mathbf{y}_k\})} \prod_{i=1}^l \tilde{\rho}_T^{(1+\kappa_i)}(x, \pi_i). \quad (3.4)$$

**Remark 3.2** The functions  $(\tilde{\rho}_T^{(1+k)})_{k \geq 1}$  take a similar role as the so-called *Ursell functions* in classical cluster expansions and thus the expansion above can be seen as a type of multi-body cluster expansion.

### 4 Examples

We will now take a look at three cases in which the assumptions of Theorem 3.1 are satisfied. First, we will look at a classic superstable pair-interaction, then at a Hamiltonian consisting of multi-body potentials of arbitrary order, which are non-negative and of finite-range, in both settings there always exists at least one translation-invariant  $(\mu, H)$ -Gibbs measure, cf. [14]. Finally, we will look at the Kirkwood-closure process from [18] which can also be shown to be a Gibbs point process for a Hamiltonian consisting of multi-body potentials of every order.

#### 4.1 Superstable Pair-Interaction

Let us consider the case that  $H$  consists only of a pair-interaction, i.e. a measurable even function  $u: \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  such that  $H(\{\mathbf{x}_n\}) = \sum_{i < j} u(x_i - x_j)$ . If we further assume that  $u$  is *superstable*, i.e. that there exists  $r_0 > 0$  and decreasing positive functions  $\varphi: (0, r_0) \rightarrow \mathbb{R}_0^+$  and  $\psi: [0, \infty) \rightarrow \mathbb{R}^+$  with

$$\int_0^{r_0} r^{d-1} \varphi(r) dr = +\infty \quad \text{and} \quad \int_0^\infty r^{d-1} \psi(r) dr < \infty,$$

and

$$u(x) \geq \varphi(|x|) \quad \text{for } 0 < |x| < r_0,$$

$$|u(x)| \leq \psi(|x|) \quad \text{for } |x| \geq r_0.$$

It is well-known, cf. [16, Chap. 4] that  $\mathbb{P}$  satisfies Assumption B if

$$e^\mu = \frac{\bar{q}}{e^{2B+1} C(u)} \quad (4.1)$$

for some  $\bar{q} < 1/2$ , where  $C(u) := \int_{\mathbb{R}^d} |e^{-u(x)} - 1| dx$  and  $B = B(u) > 0$  is the stability constant of  $H$ , with

$$D = \frac{e^\mu e^{-2B}}{(1 - \bar{q})} \frac{1 - \bar{q}}{1 - 2\bar{q}} \quad \text{and} \quad q = \frac{\bar{q}}{1 - \bar{q}}.$$

Let  $\mathbb{P}$  be such a  $(\mu, u)$ -Gibbs measure. From [15, Theorem 5.5], it follows immediately that  $\mathbb{P}$  satisfies Assumption A. Therein, it is also shown that  $\mathbb{P}$  satisfies a Ruelle-condition. We thus get the easy result:

**Corollary 4.1** *Let  $\mathbb{P}$  be a  $(\mu, u)$ -Gibbs measure with a superstable pair-interaction  $u$ , if*

$$e^\mu < \min \left\{ \frac{1}{3}, \frac{1}{1 + \frac{2}{2+\zeta D}} \right\} \left( e^{2B+1} C(u) \right)^{-1},$$

*then the expansion Eq. (3.3) holds for the true chemical potential of  $\mathbb{P}$ .*

#### 4.2 Non-negative Multi-body Interaction of Finite-Range

We say a Hamiltonian has *finite-range*, if there exists an  $R > 0$  such that, whenever  $|x_0 - y| > R$  there holds

$$W(\{x_0\} | \{\mathbf{x}_n\} \cup \{y\}) = W(\{x_0\} | \{\mathbf{x}_n\}) \quad (4.2)$$

for every set of points  $\{\mathbf{x}_n\} \subset \mathbb{R}^d$ . In particular, this means that  $W(\{x_0\} | \eta) = 0$ , for any  $\eta \in \Gamma$  with  $\text{dist}(\{x_0\}, \eta) > R$ . As mentioned, in this case there always exists at least one  $(\mu, H)$ -Gibbs measure, see e.g. [14].

**Corollary 4.2** *Let  $\mathbb{P}$  be a  $(\mu, H)$ -Gibbs measure where  $H$  satisfies Eq. (4.2) as above, then  $\mathbb{P}$  satisfies Assumption A.*

**Proof** Fix some  $\Delta \subset \mathbb{R}^d$  compact and let  $\Lambda \subset \mathbb{R}^d$  be such that  $\text{dist}(\Delta, \partial\Lambda) > R$ , then Eqs. (4.2) and (2.1) for  $n = 1$  imply

$$j_\Lambda^{(1)}(x) = e^\mu \int_{\Gamma_{\Lambda^c}} d\mathbb{P}(\eta).$$

This proves the claim.  $\square$

It was shown by Moraal in [19] that if  $H$  is additionally non-negative, then the multi-body version of the Kirkwood-Salsburg operator is bounded and there exists a solution to the multi-body version of the Kirkwood-Salsburg equations. Using the same techniques as in the pair-interaction case, see [16, Sect. 4.4.6], it then follows that  $\mathbb{P}$  satisfies Assumption B. Furthermore,  $\mathbb{P}$  satisfies a Ruelle-condition with  $\xi = e^\mu$ .

**Corollary 4.3** *Let  $H$  be a non-negative Hamiltonian of finite range,  $\mu \in \mathbb{R}$  and  $\mathbb{P}$  be a  $(\mu, H)$ -Gibbs measure. If  $\mu$  is sufficiently small, then the expansion Eq. (3.3) holds.*

**Remark 4.4** In [20] Skrypnik extended the results of Moraal to the case that  $H$  consists of the sum of a very particular non-negative multi-body interaction of finite-range and a superstable pair-interaction. In this case, it can also be shown that Assumption A is fulfilled.

### 4.3 The Kirkwood-Closure Process

A point process  $\mathbb{P}$  is called the *Kirkwood-closure process* for a given density  $\rho > 0$  and a non-negative even function  $g$  on  $\mathbb{R}^d$  if the correlation functions of  $\mathbb{P}$  satisfy

$$\rho^{(n)}(\mathbf{x}_n) = \rho^n \prod_{1 \leq i < j \leq n} g(x_i - x_j). \quad (4.3)$$

It was shown in [18] by Kuna et al. that, if  $C(g) := \int_{\mathbb{R}^d} |g(x) - 1| dx < \infty$  and there is a  $b \geq 1$  such that  $\prod_{i=1}^n g(x_i - x_0) \leq b$  for any  $x_0, x_1, \dots, x_n \in \mathbb{R}^d$  with  $\prod_{i < j} g(x_i - x_j) > 0$ , then the Kirkwood-closure process exists for  $\rho < (ebC(g))^{-1}$ . This was an extension of the results in [21], where only the case  $g \leq 1$  was discussed. The Kirkwood-closure, which in computational physics is often called Kirkwood-superposition was first suggested in [22] and has been widely used to approximate higher-order correlation functions and thus just having to calculate  $\rho$  and  $\rho^{(2)}$  from simulation data. In this case, it also known that  $\mathbb{P}$  satisfies a Ruelle condition with  $\xi = \rho b^{1/2}$  and for every bounded  $\Lambda \subset \mathbb{R}^d$

$$\sup_{x \in \mathbb{R}^d} \int_{\Lambda^n} \left| \rho_T^{(1+n)}(x, \mathbf{y}_n) \right| d\mathbf{y}_n \leq n! D (\rho e b C(g))^{n+1},$$

i.e. the Kirkwood-closure process satisfies Assumption B, cf. [18, Sect. 3.1]. In [23] Glötzl gave sufficient conditions to guarantee the existence of a chemical potential  $\mu$  and a Hamiltonian  $H$  for which the point process is Gibbs. These conditions are for the *Campbell-measure*  $C_{\mathbb{P}}$  of a point process  $\mathbb{P}$ , i.e. the measure on  $(\mathbb{R}^d \times \Gamma, \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{F})$  (here  $\mathcal{B}(\mathbb{R}^d)$  are the Borel-sets of  $\mathbb{R}^d$ ) defined by

$$C_{\mathbb{P}}(B \times F) := \int_{\Gamma} \sum_{x \in \eta} \mathbb{1}_F(\eta \setminus \{x\}) \mathbb{1}_B(x) d\mathbb{P}(\eta).$$

Namely, if the Campbell-measure  $C_{\mathbb{P}}$  is absolutely continuous with respect to  $dx \times d\mathbb{P}$  and the density  $k$  satisfies some additional assumptions, such a Hamiltonian exists, see Satz 2 in [23]. Due to the easy structure Eq. (4.3) of the correlation functions, a straightforward computation, using Eq. (2.2), shows that  $\mathbb{P}$  satisfies these assumptions and the Radon-Nikodym derivative

of the Campbell-measure is given by

$$k(x, \eta) = \rho \prod_{w \in \eta} g(x - w) \exp \left( \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{(\mathbb{R}^d)^k} \left( \prod_{m=1}^k g(x - y_m) - 1 \right) \prod_{w \in \eta} \prod_{m=1}^k g(w - y_m) \rho_T^{(k)}(\mathbf{y}_k) \, d\mathbf{y}_k \right),$$

for  $x \in \mathbb{R}^d$  and  $\eta \in \Gamma$ . Note that, due to the conditions on  $g$  the function  $k$  is well-defined on  $\mathbb{R}^d \times \Gamma$  under mild decay conditions on  $g$ . Since  $k(x, \eta) = \exp(\mu - W(\{x\} | \eta))$ , i.e.  $k(x, \emptyset) = e^\mu$ , it thus follows that  $\mu$  is given by

$$\mu = \log \rho + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{(\mathbb{R}^d)^k} \left( \prod_{m=1}^k g(x - y_m) - 1 \right) \rho_T^{(k)}(\mathbf{y}_k) \, d\mathbf{y}_k.$$

Furthermore, it is easily checked, that for the Kirkwood-closure process, there holds

$$\tilde{\rho}_T^{(1+k)}(x, \mathbf{y}_k) = \left( \prod_{m=1}^k g(x - y_m) - 1 \right) \rho_T^{(k)}(\mathbf{y}_k), \quad (4.4)$$

thus Assumption A is not needed in this case to identify the limit in Eq. (6.17), although one could also show that it holds here.

**Corollary 4.5** *Let  $g$  satisfy the assumptions above and  $\mathbb{P}$  be the Kirkwood-closure process if*

$$\rho < \frac{1}{(2 + \zeta D)ebC(g)}$$

*then the chemical potential  $\mu$  of the Kirkwood-closure process is given by Eq. (3.3).*

**Remark 4.6** In the case of the Kirkwood-closure it is quite easy to see the connection of the expansion Eq. (3.3) to classic cluster expansions. It follows from Eq. (4.4) and [18, Sect. 3.1] that

$$\tilde{\rho}_T^{(1+k)}(\tilde{\mathbf{y}}_{k+1}) = \rho^k \sum_{C \in \mathcal{C}_{k+1}} \prod_{\{i, j\} \in E(C)} (g(\tilde{y}_i - \tilde{y}_j) - 1)$$

where  $\tilde{\mathbf{y}}_{k+1} = (x, \mathbf{y}_k)$ . Thus, in this case we have

$$\mu = \log \rho + \sum_{k=1}^{\infty} \frac{(-\rho)^k}{k!} \int_{(\mathbb{R}^d)^k} \sum_{C \in \mathcal{C}_{k+1}} \prod_{\{i, j\} \in E(C)} (g(\tilde{y}_i - \tilde{y}_j) - 1) \, d\mathbf{y}_k,$$

where  $\mathcal{C}_{k+1}$  is the set of connected graphs on  $k + 1$  vertices and  $E(C)$  is the set of edges of the graph  $C$ . This has the same structure as the classic cluster expansion for the density in the case of pair-interactions with  $-\rho$  taking the role of the activity (the exponential of the chemical potential) and  $g - 1$  taking the role of the Mayer function.

**Remark 4.7** Using the easier structure of the family  $(\tilde{\rho}_T^{(1+k)})_{k \geq 1}$  in this case, i.e. Eq. (4.4), one can easily improve the bound Eq. (6.7) to

$$\sup_{x \in \mathbb{R}^d} \int_{\Lambda^n} \left| \tilde{\rho}_T^{(1+k)}(x, \mathbf{y}_k) \right| \, d\mathbf{y}_k \leq k! M (\rho ebC(g))^k$$

and obtain a better radius of convergence.

## 5 A First Step: The Case of the Zero-Point Density

Let  $\Lambda \subset \mathbb{R}^d$  be a bounded set, then in the case of  $n = 0$  we get from Eq. (2.1)

$$j_{\Lambda}^{(0)} = \int_{\Gamma_{\Lambda^c}} d\mathbb{P}(\eta).$$

Now furthermore we know that by Eq. (2.2)

$$j_{\Lambda}^{(0)} = 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{\Lambda^k} \rho^{(k)}(\mathbf{y}_k) d\mathbf{y}_k.$$

Our goal now is to write the above equation as an exponential, for that we use the truncated correlation functions defined in Eq. (2.6). Note that in Eq. (2.6) we can easily multiply the left-hand side by  $(-1)^n$  and split this factor into  $\prod_{i=1}^k (-1)^{k_i}$  on the right-hand side so that we get the representation Eq. (2.4) for  $\Phi = (-1)^n \rho^{(n)}$  with  $F_n = (-1)^n \rho_T^{(n)}$ .

Assume that  $\mathbb{P}$  satisfies Assumption B for some  $q < 1/2$ , then in particular

$$\int_{\Lambda} \int_{\Lambda^n} \left| \rho_T^{(1+n)}(x, \mathbf{y}_n) \right| d\mathbf{y}_n dx \leq |\Lambda| n! D \rho q^n. \quad (5.1)$$

This means the assumptions of Theorem A are satisfied, and we can write

$$j_{\Lambda}^{(0)} = \exp \left( \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{\Lambda^k} \rho_T^{(k)}(\mathbf{y}_k) d\mathbf{y}_k \right)$$

and thus

$$\log j_{\Lambda}^{(0)} = \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{\Lambda^k} \rho_T^{(k)}(\mathbf{y}_k) d\mathbf{y}_k. \quad (5.2)$$

**Remark 5.1** Let us consider the same setting as in Example 4.1. If in this case in Eq. (4.1) we have  $\bar{q} < 1/3$ , then  $q = \bar{q}/(1 - \bar{q}) < 1/2$  and thus taking the thermodynamic limit of Eq. (5.2), i.e. dividing by the volume  $|\Lambda|$  and increasing  $\Lambda$  to the whole space  $\mathbb{R}^d$ , gives

$$\lim_{\Lambda \nearrow \mathbb{R}^d} \frac{1}{|\Lambda|} \log j_{\Lambda}^{(0)} = -\rho + \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{(k+1)!} \int_{(\mathbb{R}^d)^k} \rho_T^{(k+1)}(0, \mathbf{y}_k) d\mathbf{y}_k = -p(\mu, u),$$

where  $p$  is the infinite-volume grand-canonical pressure, cf. [24]. The above expression follows immediately from the cluster expansions of the pressure and the truncated correlation functions therein and [25].

## 6 Generalization to the One-Point Density

We now want to generalize this approach for the one-particle density  $j_{\Lambda}^{(1)}$  of an infinite-volume Gibbs measure. Again we will use Eq. (2.2) to find an expansion and Eq. (2.1) to identify the terms. By Eq. (2.2) we have

$$j_{\Lambda}^{(1)}(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{\Lambda^k} \rho^{(1+k)}(x, \mathbf{y}_k) d\mathbf{y}_k = \rho \left( 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{\Lambda^k} \frac{\rho^{(k+1)}(x, \mathbf{y}_k)}{\rho} d\mathbf{y}_k \right). \quad (6.1)$$

We introduce the family  $(F_k)_{k \geq 1}$  that is recursively defined by  $F_1(x, y) = \rho^{(2)}(x, y)/\rho$  and for  $k \geq 2$  by

$$F_k(x, \mathbf{y}_k) = \frac{\rho^{(1+k)}(x, \mathbf{y}_k)}{\rho} - \sum_{l=2}^k \sum_{\pi \in \Pi_l(\{\mathbf{y}_k\})} \prod_{i=1}^l F_{\kappa_i}(x, \pi_i). \quad (6.2)$$

The ansatz Eq. (6.1) with Eq. (6.2) was first used by Nettleton and Green in [12] and the division by  $\rho$  lets us use Theorem A to obtain the exponential representation in Eq. (6.1) as the series then starts at one. As previously noted, we have  $j_\Lambda^{(1)} \rightarrow 0$  as  $\Lambda \nearrow \mathbb{R}^d$ , thus we need to identify the divergent part of  $\log j_\Lambda^{(1)}$ .

**Proposition 6.1** *For every  $k \in \mathbb{N}$  there holds*

$$F_k(x, \mathbf{y}_k) = \rho_T^{(k)}(\mathbf{y}_k) + \tilde{\rho}_T^{(1+k)}(x, \mathbf{y}_k), \quad (6.3)$$

where the family  $(\tilde{\rho}_T^{(1+k)})_{k \geq 1}$  is the one defined in Eq. (3.4).

**Proof** The claim holds for  $k = 1$ . By definition Eq. (6.2) of the  $F_k$  there holds

$$F_{k+1}(x, \mathbf{y}_{k+1}) = \frac{\rho^{(k+2)}(x, \mathbf{y}_{k+1})}{\rho} - \sum_{l=2}^{k+1} \sum_{\pi \in \Pi_l(\{\mathbf{y}_{k+1}\})} \prod_{i=1}^l F_{\kappa_i}(x, \pi_i). \quad (6.4)$$

For the second part of the right-hand side above we can use the induction hypothesis to find

$$\sum_{l=2}^{k+1} \sum_{\pi \in \Pi_l(\{\mathbf{y}_{k+1}\})} \prod_{i=1}^l F_{\kappa_i}(x, \pi_i) = \sum_{l=2}^{k+1} \sum_{\pi \in \Pi_l(\{\mathbf{y}_{k+1}\})} \prod_{i=1}^l \left( \rho_T^{(\kappa_i)}(\pi_i) + \tilde{\rho}_T^{(1+\kappa_i)}(x, \pi_i) \right)$$

and thus we get

$$F_{k+1}(x, \mathbf{y}_{k+1}) = \frac{\rho^{(k+2)}(x, \mathbf{y}_{k+1})}{\rho} - \sum_{l=2}^{k+1} \sum_{\pi \in \Pi_l(\{\mathbf{y}_{k+1}\})} \prod_{i=1}^l \left( \rho_T^{(\kappa_i)}(\pi_i) + \tilde{\rho}_T^{(1+\kappa_i)}(x, \pi_i) \right). \quad (6.5)$$

For the first term of the right-hand side of Eq. (6.4), there holds by Eq. (2.6)

$$\frac{\rho^{(k+2)}(x, \mathbf{y}_{k+1})}{\rho} = \frac{1}{\rho} \sum_{l=1}^{k+2} \sum_{\pi \in \Pi_l(\{x, \mathbf{y}_{k+1}\})} \prod_{i=1}^l \rho_T^{(\kappa_i)}(\pi_i).$$

For any partition  $\pi \in \Pi_l(\{x, \mathbf{y}_{k+1}\})$  we denote by  $\nu$  the index such  $\pi_\nu = (x, \pi'_\nu)$ , i.e.  $\pi_\nu$  is the part of the partition  $\pi$  containing  $x$ . We can thus write

$$\frac{\rho^{(k+2)}(x, \mathbf{y}_{k+1})}{\rho} = \sum_{l=1}^{k+2} \sum_{\pi \in \Pi_l(\{x, \mathbf{y}_{k+1}\})} \frac{\rho_T^{(\kappa_\nu)}(x, \pi'_\nu)}{\rho} \prod_{\substack{i=1 \\ i \neq \nu}}^l \rho_T^{(\kappa_i)}(\pi_i). \quad (6.6)$$

We distinguish three cases:

1. If  $\kappa_\nu = k+2$ , then  $\pi$  is the partition into one element and contributes the term  $\frac{\rho_T^{(k+2)}(x, \mathbf{y}_{k+1})}{\rho}$  which only appears in the case  $l = 1$  above;

2. If  $\kappa_v = 1$ , then  $\rho_T^{(\kappa_v)}/\rho = 1$  and we can understand  $\widehat{\pi} := \pi \setminus \pi_v$  as a partition of the elements  $\mathbf{y}_{k+1}$  into  $l - 1$  elements, i.e. an element of  $\widehat{\pi} \in \Pi_{l-1}(\mathbf{y}_{k+1})$ . Summing over all of the partitions where  $\kappa_v = 1$  then gives

$$\sum_{l=2}^{k+2} \sum_{\substack{\pi \in \Pi_l(\{x, \mathbf{y}_{k+1}\}) \\ \kappa_v=1}} \prod_{\substack{i=1 \\ i \neq v}}^l \rho_T^{(\kappa_i)}(\pi_i) = \sum_{l=2}^{k+2} \sum_{\widehat{\pi} \in \Pi_{l-1}(\mathbf{y}_{k+1})} \prod_{i=1}^{l-1} \rho_T^{(\widehat{\kappa}_i)}(\widehat{\pi}_i)$$

where  $\widehat{\kappa}_i$  is the number of elements in  $\widehat{\pi}_i$ . Shifting the index we get

$$\sum_{l=2}^{k+2} \sum_{\widehat{\pi} \in \Pi_{l-1}(\mathbf{y}_{k+1})} \prod_{i=1}^{l-1} \rho_T^{(\widehat{\kappa}_i)}(\widehat{\pi}_i) = \sum_{l=1}^{k+1} \sum_{\widehat{\pi} \in \Pi_l(\mathbf{y}_{k+1})} \prod_{i=1}^l \rho_T^{(\widehat{\kappa}_i)}(\widehat{\pi}_i) = \rho^{(k+1)}(\mathbf{y}_{k+1})$$

where we have used Eq. (2.6) for the last equality.

3. If  $\kappa_v = m + 1$  for some  $1 \leq m \leq k$ , we can use the induction assumption to write

$$\frac{\rho_T^{(1+m)}(x, \pi'_v)}{\rho} = \sum_{j=1}^m \sum_{\widehat{\pi} \in \Pi_l(\{\pi'_v\})} \prod_{s=1}^j \widetilde{\rho}_T^{(1+\widehat{\kappa}_s)}(x, \widehat{\pi}_s).$$

Plugging the three cases above into Eq. (6.6) we find

$$\begin{aligned} \frac{\rho^{(k+2)}(x, \mathbf{y}_{k+1})}{\rho} &= \sum_{l=1}^{k+2} \sum_{\pi \in \Pi_l(\{x, \mathbf{y}_{k+1}\})} \sum_{\kappa_v=1}^{k+2} \frac{\rho_T^{(\kappa_v)}(x, \pi'_v)}{\rho} \prod_{\substack{i=1 \\ i \neq v}}^l \rho_T^{(\kappa_i)}(\pi_i) \\ &= \frac{\rho_T^{(k+2)}(x, \mathbf{y}_{k+1})}{\rho} + \rho^{(k+1)}(\mathbf{y}_{k+1}) \\ &\quad + \sum_{l=2}^{k+2} \sum_{\pi \in \Pi_l(\{x, \mathbf{y}_{k+1}\})} \sum_{\kappa_v=2}^{k+1} \sum_{j=1}^m \sum_{\widehat{\pi} \in \Pi_l(\{\pi'_v\})} \prod_{s=1}^j \widetilde{\rho}_T^{(1+\widehat{\kappa}_s)}(x, \widehat{\pi}_s) \prod_{\substack{i=1 \\ i \neq v}}^l \rho_T^{(\kappa_i)}(\pi_i). \end{aligned}$$

Finally, we plug the above into Eq. (6.5) to get

$$\begin{aligned} F_{k+1}(x, \mathbf{y}_{k+1}) &= \frac{\rho_T^{(k+2)}(x, \mathbf{y}_{k+1})}{\rho} + \rho^{(k+1)}(\mathbf{y}_{k+1}) \\ &\quad + \sum_{l=2}^{k+2} \sum_{\pi \in \Pi_l(\{x, \mathbf{y}_{k+1}\})} \sum_{\kappa_v=2}^{k+1} \sum_{j=1}^m \sum_{\widehat{\pi} \in \Pi_l(\{\pi'_v\})} \prod_{s=1}^j \widetilde{\rho}_T^{(1+\widehat{\kappa}_s)}(x, \widehat{\pi}_s) \prod_{\substack{i=1 \\ i \neq v}}^l \rho_T^{(\kappa_i)}(\pi_i) \\ &\quad - \sum_{l=2}^{k+1} \sum_{\pi \in \Pi_l(\mathbf{y}_{k+1})} \prod_{i=1}^l (\rho_T^{(\kappa_i)}(\pi_i) + \widetilde{\rho}_T^{(1+\kappa_i)}(x, \pi_i)). \end{aligned}$$

Looking closer at the difference of the last two lines above, we see that all terms containing at least one  $\rho_T^{(\kappa_i)}$  and one  $\tilde{\rho}_T^{(1+\kappa_j)}$  cancel and we get

$$\begin{aligned} & \sum_{l=2}^{k+2} \sum_{\pi \in \Pi_l(\{x, \mathbf{y}_{k+1}\})} \sum_{\kappa_v=2}^{k+1} \sum_{j=1}^m \sum_{\hat{\pi} \in \Pi_l(\{\pi'_v\})} \prod_{s=1}^j \tilde{\rho}_T^{(1+\kappa_s)}(x, \hat{\pi}_s) \prod_{\substack{i=1 \\ i \neq v}}^l \rho_T^{(\kappa_i)}(\pi_i) \\ & - \sum_{l=2}^{k+1} \sum_{\pi \in \Pi_l(\{\mathbf{y}_{k+1}\})} \prod_{i=1}^l (\rho_T^{(\kappa_i)}(\pi_i) + \tilde{\rho}_T^{(1+\kappa_i)}(x, \pi_i)) \\ & = - \sum_{l=2}^{k+1} \sum_{\pi \in \Pi_l(\{\mathbf{y}_{k+1}\})} \prod_{i=1}^l \tilde{\rho}_T^{(1+\kappa_i)}(x, \pi_i). \end{aligned}$$

We thus have that

$$F_{k+1}(x, \mathbf{y}_{k+1}) = \rho_T^{(k+1)}(\mathbf{y}_{k+1}) + \frac{\rho_T^{(k+2)}(x, \mathbf{y}_{k+1})}{\rho} - \sum_{l=2}^{k+1} \sum_{\pi \in \Pi_l(\{\mathbf{y}_{k+1}\})} \prod_{i=1}^l \tilde{\rho}_T^{(1+\kappa_i)}(x, \pi_i).$$

Finally, using Eq. (3.4) proves Eq. (6.3). □

We now want to apply Theorem A to Eq. (6.1), in view of Eqs. (6.3) and (3.1) we thus need a bound on the integrals over  $(\tilde{\rho}_T^{(1+k)})_{k \geq 1}$ .

**Theorem 6.2** *Let  $\mathbb{P}$  be a  $(\mu, H)$ -Gibbs measure that satisfies Assumption B, then for any bounded  $\Lambda \subset \mathbb{R}^d$*

$$\sup_{x \in \mathbb{R}^d} \int_{\Lambda^n} \left| \tilde{\rho}_T^{(1+k)}(x, \mathbf{y}_k) \right| d\mathbf{y}_k \leq (k-1)! M (1 + \zeta D)^k q^k \tag{6.7}$$

for some  $M > 0$  and  $\zeta > 1$  independent of  $\Lambda$ .

**Proof** By Eq. (3.4) there holds

$$\int_{\Lambda^k} \tilde{\rho}_T^{(1+k)}(x, \mathbf{y}_k) d\mathbf{y}_k = \int_{\Lambda^k} \frac{\rho_T^{(1+k)}(x, \mathbf{y}_k)}{\rho} d\mathbf{y}_k - \sum_{l=2}^k \sum_{\pi \in \Pi_l(\{\mathbf{y}_k\})} \prod_{i=1}^l \int_{\Lambda^{\kappa_i}} \tilde{\rho}_T^{(1+\kappa_i)}(x, \pi_i) d\pi_i.$$

Note that because we integrate over all  $\mathbf{y}_k$ , when looking at a particular partition  $\pi \in \Pi_l(\{\mathbf{y}_k\})$  it does not matter which particular elements of  $\mathbf{y}_k$  are contained in which part  $\pi_i$  of  $\pi$  only the sizes of the different parts  $\pi_i$  matter. This is the definition of the  $k$ th Bell polynomial, cf. [26], and allows us to write

$$\int_{\Lambda^k} \tilde{\rho}_T^{(1+k)}(x, \mathbf{y}_k) d\mathbf{y}_k = -B_k \left( I_1, \dots, I_{k-1}, - \int_{\Lambda^k} \frac{\rho_T^{(1+k)}(x, \mathbf{y}_k)}{\rho} d\mathbf{y}_k \right)$$

where  $I_i = \int_{\Lambda^{\kappa_i}} \tilde{\rho}_T^{(1+\kappa_i)}(x, \mathbf{y}_{\kappa_i}) d\mathbf{y}_{\kappa_i}$ . Taking the absolute value above and using the bound from Assumption B we thus get

$$\int_{\Lambda^k} \left| \tilde{\rho}_T^{(1+k)}(x, \mathbf{y}_k) \right| d\mathbf{y}_k \leq w_k \tag{6.8}$$

where the  $w_k$  are recursively defined by  $w_1 = Dq$  and for  $k \geq 2$

$$w_k = B_k \left( w_1, \dots, w_{k-1}, k! Dq^k \right). \tag{6.9}$$

We use the ansatz

$$w_k = q^k \sum_{l=0}^k a_l^{(k)} D^l,$$

and want to calculate the coefficients  $a_l^{(k)}$  for  $0 \leq l \leq k$  in  $w_k$  using the well-known recursion, cf. [26],

$$B_{k+1}(x_1, \dots, x_{k+1}) = \sum_{i=0}^k \binom{k}{i} B_{k-i}(x_1, \dots, x_{k-i}) x_{i+1}, \quad (6.10)$$

where  $B_0 := 1$ . This recursion along with the fact that  $w_1 = Dq$  already implies that  $a_0^{(k)} = 0$  for every  $k$ . Lastly, we note that Eq. (6.10) implies that

$$B_k(x_1, \dots, x_{k-1}, x_k + y_k) = B_k(x_1, \dots, x_{k-1}, x_k) + y_k.$$

Putting  $x_i = w_i$   $1 \leq i \leq k-1$ ,  $x_k = q^k k! D$  and  $y_k = w_k - q^k k! D$  then gives

$$B_k(w_1, \dots, w_k) = q^k \left( 2 \sum_{l=1}^k a_l^{(k)} D^l - k! D \right).$$

Plugging the above into Eq. (6.10) we find

$$\begin{aligned} B_{k+1}(w_1, \dots, w_k, (k+1)! D q^{k+1}) &= (k+1)! D q^{k+1} \\ &+ \sum_{i=0}^{k-1} \binom{k}{i} q^{k-i} \left( 2 \sum_{l=1}^{k-i} a_l^{(k-i)} D^l - (k-i)! D \right) q^{i+1} \left( \sum_{l=1}^{i+1} a_l^{(i+1)} D^l \right). \end{aligned}$$

This already shows that  $a_1^{(k+1)} = (k+1)!$  as in all the terms resulting from the second sum  $D$  has an exponent of at least two. We can thus use the above equality to find

$$\begin{aligned} a_2^{(k+1)} &= \sum_{i=0}^{k-1} \binom{k}{i} (k-i)! (i+1)!, \\ a_m^{(k+1)} &= \sum_{i=0}^{k-1} \binom{k}{i} \left( (k-i)! a_{m-1}^{(i+1)} + 2(i+1)! a_{m-1}^{(k-i)} + 2 \sum_{v=2}^{m-2} a_v^{(i+1)} a_{m-v}^{(k-i)} \right), \quad m \geq 3. \end{aligned}$$

Straightforward computation gives

$$a_2^{(k+1)} = \sum_{i=0}^{k-1} \binom{k}{i} (k-i)! (i+1)! = (k+1)! \frac{k}{2}.$$

We claim that for every  $k, v \in \mathbb{N}$  there holds

$$a_v^{(k)} = \frac{k!}{v!(v-1)!} (k-1) \cdots (k-v+1) b_v = \frac{k!}{v!} \binom{k-1}{v-1} b_v \quad (6.11)$$

where  $(b_v)_{v \geq 0}$  is the sequence defined by  $b_0 = 0, b_1 = b_2 = 1$  and for  $v \geq 2$

$$b_{v+1} = (v+2)b_v + 2 \sum_{j=2}^{v-1} \binom{v}{j} b_j b_{v-j+1}. \quad (6.12)$$

Note that  $a_m^{(k)} = 0$  if  $k < m$ . We have already established that Eq. (6.11) holds for all  $k \geq m$  for  $m = 0, 1, 2$ . We will prove the claim by induction. Let Eq. (6.11) hold for all  $k$  and every  $v \leq m$  for some  $m \geq 2$ , then

$$\begin{aligned} a_{m+1}^{(k+1)} &= \sum_{i=0}^{k-1} \binom{k}{i} (k-i)! a_m^{(i+1)} + 2 \sum_{i=0}^{k-1} \binom{k}{i} \sum_{v=1}^{m-1} a_v^{(i+1)} a_{m+1-v}^{(k-i)} \\ &= \sum_{i=0}^{k-1} \binom{k}{i} (k-i)! \frac{(i+1)!}{m!} \binom{i}{m-1} b_m \end{aligned} \quad (6.13)$$

$$+ 2 \sum_{i=0}^{k-1} \binom{k}{i} \sum_{v=1}^{m-1} \frac{(i+1)!}{v!} \binom{i}{v-1} b_v \frac{(k-i)!}{(m+1-v)!} \binom{k-i-1}{m-v} b_{m+1-v}. \quad (6.14)$$

We start by simplifying the term in Eq. (6.13) and get

$$\sum_{i=0}^{k-1} \binom{k}{i} (k-i)! \frac{(i+1)!}{m!} \binom{i}{m-1} b_m = \frac{k! b_m}{m!} m \sum_{i=0}^{k-1} \binom{i+1}{m} = \frac{k! b_m}{m!} \binom{k+1}{m+1} m \quad (6.15)$$

by the Vandermonde identity, cf. [27]. For the term in Eq. (6.14), we first swap the inner and outer sums and rearrange the binomial coefficients to get

$$\begin{aligned} &\sum_{i=0}^{k-1} \binom{k}{i} \sum_{v=1}^{m-1} \frac{(i+1)!}{v!} \binom{i}{v-1} b_v \frac{(k-i)!}{(m+1-v)!} \binom{k-i-1}{m-v} b_{m+1-v} \\ &= \frac{k!}{m!} \sum_{v=1}^{m-1} \binom{m}{v-1} b_v b_{m+1-v} \sum_{i=0}^{k-1} \binom{i+1}{v} \binom{k-i-1}{m-v}. \end{aligned}$$

We note that

$$\sum_{i=0}^{k-1} \binom{i+1}{v} \binom{k-i-1}{m-v} = \binom{k+1}{m+1},$$

which is a version of the Chu-Vandermonde identity, see e.g. [27], and thus

$$\frac{k!}{m!} \sum_{v=1}^{m-1} \binom{m}{v-1} b_v b_{m+1-v} \sum_{i=0}^{k-1} \binom{i+1}{v} \binom{k-i-1}{m-v} = \frac{k!}{m!} \binom{k+1}{m+1} \sum_{v=1}^{m-1} \binom{m}{v-1} b_v b_{m+1-v}. \quad (6.16)$$

Plugging Eqs. (6.15) and (6.16) back into Eqs. (6.13) and (6.14) we get

$$a_{m+1}^{(k+1)} = \frac{k!}{m!} \binom{k+1}{m+1} \left( m b_m + 2 b_m + 2 \sum_{v=2}^{m-1} \binom{m}{v} b_v b_{m+1-v} \right) = \frac{k!}{m!} \binom{k+1}{m+1} b_{m+1}$$

where we used that

$$\sum_{v=1}^{m-1} \binom{m}{v-1} b_v b_{m+1-v} = b_m + \sum_{v=2}^{m-1} \binom{m}{v} b_v b_{m+1-v}.$$

The claim is thus proved, and we arrive at the bound

$$\int_{\Lambda^k} \left| \tilde{\rho}_T^{(1+k)}(x, \mathbf{y}_k) \right| d\mathbf{y}_k \leq q^k \sum_{m=1}^k \binom{k}{m} \frac{(k-1)!}{(m-1)!} b_m D^m.$$

The sequence  $(b_m)_{m \geq 0}$  is the number of total partitions of  $m$  elements, see [28, 29]. Therein we also find the asymptotic behavior of  $(b_m)_m$ , namely, there is an  $M > 0$  such that

$$b_m \leq M \frac{m^{m-1} \zeta^m}{e^m} \quad \text{where} \quad \zeta = \frac{1}{2 \log 2 - 1},$$

giving

$$\sum_{m=1}^k \binom{k}{m} \frac{(k-1)!}{(m-1)!} b_m D^m \leq M(k-1)!(1 + \zeta D)^k.$$

□

**Proof of Theorem 3.1.** Assume that  $\mathbb{P}$  is a  $(\mu, H)$ -Gibbs measure satisfying a Ruelle-condition as well as Assumptions A and B. By Eq. (6.3), the triangle inequality, Eq. (5.1) and Theorem 6.2 we get for  $F_n$  of Eq. (6.2), that

$$\begin{aligned} \int_{\Lambda^n} |F_n(0, \mathbf{y}_n)| \, d\mathbf{y}_n &\leq |\Lambda|(n-1)! D \rho q^{n-1} + M(n-1)!(1 + \zeta D)^n q^n \\ &\leq \left( |\Lambda| \frac{D \rho}{q} + M \right) (n-1)! (q(2 + \zeta D))^n. \end{aligned}$$

If now  $q < q_0$  from Eq. (3.2), then by Theorem A there holds

$$\log j_{\Lambda}^{(1)}(0) = \log \rho + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{\Lambda^k} \left( \rho_T^{(k)}(\mathbf{y}_k) + \tilde{\rho}_T^{(1+k)}(0, \mathbf{y}_k) \right) \, d\mathbf{y}_k.$$

Which, in view of Eq. (5.2), is equivalent to

$$\log j_{\Lambda}^{(1)}(0) - \log j_{\Lambda}^{(0)} = \log \rho + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{\Lambda^k} \tilde{\rho}_T^{(1+k)}(0, \mathbf{y}_k) \, d\mathbf{y}_k. \quad (6.17)$$

By Assumption A and Eq. (2.1) the left-hand side of Eq. (6.17) converges to  $\mu$  as  $\Lambda \nearrow \mathbb{R}^d$  and since the bound in Eq. (6.7) is independent of  $\Lambda$  the right-hand side above converges by dominated convergence, and we arrive at Eq. (3.3). □

**Funding** Open Access funding enabled and organized by Projekt DEAL. The research leading to this work has been done within the Collaborative Research Center TRR 146, corresponding funding by the DFG is gratefully acknowledged.

**Data Availability** Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

## Declarations

**Conflict of interest** The author states that there is no Conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Chayes, J.T., Chayes, L.: On the validity of the inverse conjecture in classical density functional theory. *J. Stat. Phys.* **36**, 471–488 (1984)
2. Chayes, J.T., Chayes, L.L., Lieb, E.: The inverse problem in classical statistical mechanics. *Commun. Math. Phys.* **93**, 57–121 (1984)
3. Henderson, R.L.: A uniqueness theorem for fluid pair correlation functions. *Phys. Lett. A* **49**, 197–198 (1974)
4. Korolov, L.: An inverse problem for Gibbs fields with hard core potential. *J. Math. Phys.* **48**, 053301 (2007)
5. Frommer, F., Hanke, M., Jansen, S.: A note on the uniqueness result for the inverse Henderson problem. *J. Math. Phys.* **60**, 093303 (2019)
6. Takacs, R.: Estimator for the pair-potential of a Gibbsian point process. *Statistics* **17**, 233–429 (1986)
7. Fiksel, T.: Estimation of parameterized pair potentials of marked and non-marked Gibbsian point processes. *Elektron. Inform. Kybernet.* **20**, 270–278 (1984)
8. Coeurjolly, J.F., Dereudre, D., Drouilhet, R., Lavancier, F.: Takacs-Fiksel method for stationary marked Gibbs point processes. *Scand. J. Statist.* **39**, 416–443 (2011)
9. Baddeley, A., Dereudre, D.: Variational estimators for the parameters of Gibbs point process models. *Bernoulli* **19**, 905–930 (2013)
10. Lebowitz, J.L., Penrose, O.: Convergence of virial expansions. *J. Math. Phys.* **5**, 841–847 (1964)
11. Jansen, S., Kuna, T., Tsagkarogiannis, D.: Virial inversion and density functionals. *J. Funct. Anal.* **284**, 109731 (2023)
12. Nettleton, R.E., Green, M.S.: Expression in terms of molecular distribution functions for the entropy density in an infinite system. *J. Chem. Phys.* **29**, 1365–1370 (1958)
13. Nguyen, T.X., Fernández, R.: Convergence of cluster and virial expansions for repulsive classical gases. *J. Stat. Phys.* **179**, 448–484 (2020)
14. Dereudre, D., Vasseur, T.: Existence of Gibbs point processes with stable infinite range interaction. *J. Appl. Prob.* **57**, 775–791 (2020)
15. Ruelle, D.: Superstable interactions in classical statistical mechanics. *Comm. Math. Phys.* **18**, 127–159 (1970)
16. Ruelle, D.: *Statistical Mechanics: Rigorous Results*. W.A. Benjamin Publishing, New York (1969)
17. Boluh, V.A., Rebenko, A.: An exponential representation for some integrals with respect to Lebesgue-Poisson measure. *Methods Funct. Anal. Topol.* **20**, 186–192 (2014)
18. Kuna, T., Lebowitz, J.L., Speer, E.R.: Realizability of point processes. *J. Stat. Phys.* **129**, 417–439 (2007)
19. Moraal, H.: The Kirkwood-Salsburg equation and the virial expansion for many-body potentials. *Phys. Lett. A* **59**, 9–10 (1976)
20. Skrypnik, V.I.: Solutions of the Kirkwood-Salsburg equation for particles with finite-range nonpairwise repulsion. *Ukr. Mat. J.* **60**, 1329–1334 (2008)
21. Ambartzumian, R.V., Sukiasian, H.S.: Inclusion-exclusion and point processes. *Acta Appl. Math.* **22**, 15–31 (1991)
22. Kirkwood, J.G.: Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **3**, 300–313 (1935)
23. Glötzl, E.: Lokale energien und potenziale für punktprozesse. *Math. Nachr.* **96**, 195–206 (1980)
24. Penrose, O.: Convergence of fugacity expansions for fluids and lattice gases. *J. Math. Phys.* **4**, 1312–1320 (1963)
25. Ruelle, D.: Cluster property of the correlation functions of classical gases. *Rev. Mod. Phys.* **36**, 580–583 (1964)
26. Comtet, L.: *Advanced Combinatorics: The Art of Finite and Infinite Expansions*. Reidel Publishing Company, Dordrecht, Holland/Boston (1974)
27. Spieß, J.: Some identities involving harmonic numbers. *Math. Comp.* **55**, 839–863 (1990)
28. Flajolet, P.: A problem in statistical classification theory <https://algo.inria.fr/libraries/autocombschroeder-html/schroeder.html> (1997). Accessed 12 Oct 2023
29. Flajolet, P., Sedgewick, R.: *Analytic Combinatorics*. Cambridge University Press, Cambridge (2009)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 6

# Conclusion and Outlook

In this thesis we have investigated multiple questions around the inverse Henderson problem in the thermodynamic limit. In Chapter 3 we have used an infinite volume version of the Gibbs variational principle to prove Henderson's Theorem, i.e. the injectivity of the operator mapping the chemical potential and pair interaction to the density and pair correlation function. However, in Chapter 4 we have shown that this uniqueness is only true when restricting to Gibbs measures depending only on a pair interaction, otherwise one can quite easily construct a multitude of other solutions.

As another consequence of the Gibbs variational principle we showed in Chapter 4 that the grand-canonical pressure is strictly convex and we have used this result to reformulate the inverse Henderson problem as a minimization problem for the strictly convex relative entropy functional. We have also shown that even when choosing the chemical potential as a functional of the pair potential, in order to work at a fixed density, this strict convexity is retained.

For this relative entropy we calculated the first two derivatives on an appropriate space of perturbed potentials. The first derivative is an operator related to the specific energy of the system and the second derivative is the limit of a covariance operator and thus symmetric, bilinear and positive semidefinite and can be extended from the space of perturbations to  $L^2(\mathbb{R}^d)$ . It remains an open problem however, how to treat the „pair potentials“ in  $L^2(\mathbb{R}^d) \setminus \mathcal{V}_{u_*}$  as they will in general not have the decay properties of  $\mathcal{V}_{u_*}$ . Furthermore we expect the operator  $\Phi''$  to be positive definite at least on a nice subspace of  $\mathcal{V}_{u_*}$ , this property is related to *nonhyperuniformity* of point processes, which is satisfied for Gibbs measures with pair interactions, cf. [8].

In Chapter 5 we have modified an ansatz used by Nettleton and Green for the grand-canonical ensemble for infinite volume Gibbs measures to find an expansion of the chemical potential. Using our approach we have been able to show that the resulting expansion converges even in the thermodynamic limit. We have also shown that this result is valid for a wide range of models. It would be interesting to see whether this same technique can be used to construct an expansion of the pair interaction of a Gibbs measure.



# Bibliography

- [1] R. V. Ambartzumian and H. S. Sukiasian. “Inclusion-exclusion and point processes.” In: *Acta Applicandae Mathematicae* 22 (1991), pp. 15–31.
- [2] G. Benfatto, C. Gruber, and P. A. Martin. “Exact decay of correlations for infinite range continuous systems.” In: *Helvetica Physica Acta* 57 (1984), pp. 63–85.
- [3] V. A. Boluh and A. Rebenko. “An exponential representation for some integrals with respect to the Lebesgue-Poisson measure.” In: *Methods of Functional Analysis and Topology* 20 (2014), pp. 186–192.
- [4] J. T. Chayes and L. Chayes. “On the validity of the inverse conjecture in classical density functional theory.” In: *Journal of Statistical Physics* 36 (1984), pp. 471–488.
- [5] J. T. Chayes, L. Chayes, and R. Kotecký. “The analysis of the Widom-Rowlinson model by stochastic geometric methods.” In: *Communications in Mathematical Physics* 172 (1995), pp. 551–569.
- [6] J. T. Chayes, L. Chayes, and E. H. Lieb. “The inverse problem in classical statistical mechanics.” In: *Communications in Mathematical Physics* 93 (1984), pp. 57–121.
- [7] D. Dereudre. “Variational principle for Gibbs point processes with finite range interaction.” In: *Electronic Communications in Probability* 21 (2016), pp. 1–11.
- [8] D. Dereudre and D. Flimmel. “Non-hyperuniformity of Gibbs point processes with short-range interactions.” In: *Journal of Applied Probability* 61 (2024), pp. 1380–1406.
- [9] D. Dereudre and C. Renaud-Chan. *Liquid-Gas phase transition for Gibbs point process with Quermass interaction*. arXiv: 2309.08338. 2023.
- [10] R. L. Dobrushin. “Investigation of conditions for the asymptotic existence of the configuration integral of Gibbs distribution.” In: *Theory of Probability and Its Applications* 9 (1964), pp. 566–581.
- [11] R. L. Dobrushin and R. A. Minlos. “Existence and continuity of pressure in classical statistical physics.” In: *Theory of Probability and Its Applications* 12 (1967), pp. 535–559.
- [12] M. E. Fisher and D. Ruelle. “The Stability of Many-Particle Systems.” In: *Journal of Mathematical Physics* 7 (1966), pp. 260–270.
- [13] F. Frommer. “An Inverse Cluster Expansion for the Chemical Potential.” In: *Journal of Statistical Physics* 191. 112 (2024).
- [14] F. Frommer and M. Hanke. “A variational framework for the inverse Henderson problem of statistical mechanics.” In: *Letters in Mathematical Physics* 112. 71 (2022).
- [15] F. Frommer, M. Hanke, and S. Jansen. “A note on the uniqueness result for the inverse Henderson problem.” In: *Journal of Mathematical Physics* 60 (2019). 093303.
- [16] G. Gallavotti and S. Miracle-Sole. “A variational principle for the equilibrium of hard sphere systems.” In: *Annales de l’institut Henri Poincaré. Section A, Physique Théorique* 8 (1968), pp. 287–299.

- [17] H.-O. Georgii. “Canonical and Grand Canonical Gibbs States for Continuum Systems.” In: *Communications in Mathematical Physics* 48 (1976), pp. 31–51.
- [18] H.-O. Georgii. “Large deviations and the equivalence of ensembles for Gibbsian particle systems with superstable interaction.” In: *Probability Theory and Related Fields* 99 (1994), pp. 171–195.
- [19] H.-O. Georgii. “The equivalence of ensembles for classical systems of particles.” In: *Journal of Statistical Physics* 80 (1995), pp. 1341–1378.
- [20] H.-O. Georgii. *Gibbs measures and phase transitions*. De Gruyter, Berlin, Second Ed., 2011.
- [21] H.-O. Georgii and H. Zessin. “Large deviations and the maximum entropy principle for marked point random fields.” In: *Probability Theory and Related Fields* 96 (1993), pp. 177–204.
- [22] J. Gibbs. *Elementary Principles in Statistical Mechanics*. Scribner’s Sons, 1902.
- [23] J. Ginibre. “Rigorous lower bound on the compressibility of a classical system.” In: *Physics Letters A* 24 (1967), pp. 223–224.
- [24] J. Groeneveld. “Estimation methods for Mayer’s graphical expansions.” PhD thesis. University of Amsterdam, 1967.
- [25] J. Groeneveld. “Rigorous bounds for the equation of state and pair correlation function of classical many-particle systems.” In: *Statistical Mechanics. Foundations and Applications*. Benjamin, New York, 1967, pp. 110–145.
- [26] M. Hanke. *A Taste of Inverse Problems. Basic Theory and Examples*. SIAM, Philadelphia, 2017.
- [27] M. Hanke. “Fréchet differentiability of molecular distribution functions I.” In: *Letters in Mathematical Physics* 108 (2018), pp. 285–306.
- [28] M. Hanke. “Fréchet differentiability of molecular distribution functions II: the Ursell function.” In: *Letters in Mathematical Physics* 108 (2018), pp. 307–329.
- [29] M. Hanke. “Well-Posedness of the Iterative Boltzmann Inversion.” In: *Journal of Statistical Physics* 170 (2018), pp. 536–553.
- [30] J.-P. Hansen and I.R. McDonald. *Theory of Simple Liquids, Fourth Ed.* Academic Press, Oxford, 2013.
- [31] R. L. Henderson. “A uniqueness theorem for fluid pair correlation functions.” In: *Physics Letters A* 49 (1974), pp. 197–198.
- [32] P. Hohenberg and W. Kohn. “Inhomogeneous electron gas.” In: *Physical Review B* 136 (1964), pp. 864–871.
- [33] S. Jansen. *Gibbsian point processes, Lecture notes*. [mathematik.uni-muenchen.de/jansen/gibbspp.pdf](http://mathematik.uni-muenchen.de/jansen/gibbspp.pdf). [Online; accessed 28-September-2023]. 2018.
- [34] J. G. Kirkwood. “Statistical mechanics of fluid mixtures.” In: *Journal of Chemical Physics* 3 (1935), pp. 300–313.
- [35] J. G. Kirkwood and Z. W. Salsburg. “The statistical mechanical theory of molecular distribution functions in liquids.” In: *Discussions of the Faraday Society* 15 (1953), pp. 28–34.
- [36] T. Kuna. “Studies in configuration space analysis and applications.” PhD thesis. Rheinische Friedrich-Wilhelms-Universität Bonn, 1999.
- [37] T. Kuna, J. L. Lebowitz, and E. R. Speer. “Realizability of point processes.” In: *Journal of Statistical Physics* 129 (2007), pp. 417–439.
- [38] G. Last and M. Penrose. *Lectures on the Poisson process (Vol. 7)*. Cambridge University Press, 2017.

- [39] J. L. Lebowitz, A. Mazel, and E. Presutti. “Liquid–vapor phase transitions for systems with finite-range interactions.” In: *Journal of Statistical Physics* 94 (1999), pp. 955–1025.
- [40] A. P. Lyubartsev and A. Laaksonen. “Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach.” In: *Physical Review E* 52 (1995), pp. 3730–3737.
- [41] J. E. Mayer. “Integral equations between distribution functions of molecules.” In: *Journal of Chemical Physics* 15 (1947), pp. 187–201.
- [42] J. E. Mayer and E. Montroll. “Molecular Distribution.” In: *Journal of Chemical Physics* 9 (1941), pp. 2–16.
- [43] N. Mermin. “Thermal properties of the inhomogeneous electron gas.” In: *Physical Review A* 137 (1965), pp. 1471–1473.
- [44] T. Murtola, M. Karttunen, and I. Vattulainen. “Systematic coarse graining from structure using internal states: Application to phospholipid/cholesterol bilayer.” In: *Journal of Chemical Physics* 131 (2009). 055101.
- [45] R. E. Nettleton and M. S. Green. “Expression in Terms of Molecular Distribution Functions for the Entropy Density in an Infinite System.” In: *Journal of Chemical Physics* 29 (1957), pp. 1365–1370.
- [46] X. X. Nguyen and H. Zessin. “Punktprozesse mit Wechselwirkung.” In: *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 37 (1976), pp. 91–126.
- [47] X. X. Nguyen and H. Zessin. “Integral and Differential Characterizations of the Gibbs Process.” In: *Mathematische Nachrichten* 88 (1979), pp. 105–115.
- [48] W. G. Noid. “Biomolecular Simulations. Methods and Protocols.” In: Springer, New York, 2013. Chap. Systematic methods for structurally consistent coarse-grained models.
- [49] O. Penrose. “Convergence of Fugacity Expansions for Fluids and Lattice Gases.” In: *Journal of Mathematical Physics* 4 (1963), pp. 1312–1320.
- [50] C. Peter and K. Kremer. “Multiscale simulation of soft matter systems - from the atomistic to the coarse-grained level and back.” In: *Soft Matter* 5 (2009), pp. 4357–4366.
- [51] M. Praprotnik, L. Delle Site, and K. Kremer. “Adaptive resolution molecular-dynamics simulation: Changing the degrees of freedom on the fly.” In: *Journal of Chemical Physics* 123 (2005). 224106.
- [52] A. Procacci and S. Yuhjtman. “Convergence of Mayer and virial expansions and the Penrose tree-graph identity.” In: *Letters in Mathematical Physics* 107 (2017), pp. 31–46.
- [53] D. W. Robinson and D. Ruelle. “Mean entropy of states in classical statistical mechanics.” In: *Communications in Mathematical Physics* 5 (1967), pp. 288–300.
- [54] D. Ruelle. “Classical statistical mechanics of a system of particles.” In: *Helvetica Physica Acta* 36 (1963), pp. 183–197.
- [55] D. Ruelle. *Statistical Mechanics: Rigorous Results*. W. A. Benjamin Publ., 1969.
- [56] D. Ruelle. “Superstable interactions in classical statistical mechanics.” In: *Communications in Mathematical Physics* 18 (1970), pp. 127–159.
- [57] D. Ruelle. “Existence of a phase transition in a continuous classical system.” In: *Physical Review Letters* 27 (1971), pp. 1040–1041.
- [58] M.S. Shell. “The relative entropy is fundamental to multiscale and inverse thermodynamic problems.” In: *Journal of Chemical Physics* 129 (2008). 144108.
- [59] G. Stell. “Cluster expansion for classical systems in equilibrium.” In: *Frisch, H., Lebowitz, J. (eds.) Classical Fluids*. Benjamin, New York, 1964.

- [60] H. Wang, F. H. Stillinger, and S. Torquato. “Sensitivity of pair statistics on pair potentials in many-body systems.” In: *Journal of Chemical Physics* 153 (2020). 124106.

# Acknowledgements

I am grateful for the opportunity to work on this project over the last years. This never would have been possible without a lot of support and help and I want to express my gratitude:

First of all, I want to thank my NAME. After writing my master's thesis with him, NAME asked me all these years ago if I was interested in doing a PhD with him. Already during my master's thesis I could always knock on his office door to discuss any ideas with him and he always gave helpful input. This continued during my PhD and I am very happy I chose this path.

I also want to thank NAME and NAME who welcomed me in L'aquila for a research stay and who have both become collaborators and very good colleagues of mine. I am very much looking forward to future projects with them.

Further thanks go to NAME who gave NAME and I many helpful comments during my first paper and for the fruitful discussions over the years.

Many thanks to NAME who has become a good friend and colleague with whom I spent more than a few enjoyable days (and nights) at conferences.

I do not want to forget NAME who started her term as student speaker with me and NAME who helped us as an organiser for the many events we did with the IRTG.

I also want to thank my colleagues from the numerics department who were always helpful and made many exam gradings bearable.

Lastly, I want to thank my parents and family for supporting and always believing in me.



# Curriculum vitae