A note on the uniqueness result for the inverse Henderson problem

F. Frommer,[1, a)] M. Hanke,[1, b)] and S. Jansen[2, c)]
[1)] *Institut für Mathematik, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany*
[2)] *Institut für Mathematik, Ludwig-Maximilians Universität München, Theresienstr. 39, 80333 München, Germany*

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.

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; cf., e.g., the survey of Potestio, Peters, and Kremer[37], or the special issue[7] of The European Physical Journal, Special Topics.

Among these multiscale ingredients is a technique known as (bottom-up) *coarse-graining*[34,35,39]: 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[38]. 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[34]) 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 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 point-like 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[18], 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[20] 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 (cf., e.g., Hansen and McDonald[11]), and those fail to provide sufficiently accurate potentials in general (Schommers[44]), physicists have started in the 1970's to design iterative algorithms for the numerical solution of the inverse Henderson problem[28,29,39,43,45,46], and this still is a very active area of research[9,19,31,48].

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, cf., e.g., the papers[1,3,30,36], and the references therein.

In his fundamental paper Henderson employs a technique suggested by Hohenberg and Kohn[21], Mermin[32], and others, for studying a similar inverse problem for *external potentials*; see also[5,22]. 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 probabil-

[a)] fafromme@uni-mainz.de
[b)] hanke@math.uni-mainz.de
[c)] jansen@math.lmu.de

ity 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, cf., e.g.,[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[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[13]. 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; cf., e.g., Ruelle[41] 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[17]; see also Caglioti, Kuna, Lebowitz, and Speer[2].

The outline of this note is as follows: In Section 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 Section 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 Section 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 $u : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, 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 $\varphi\colon (0, r_0] \to \mathbb{R}_0^+$ with
$$\int_0^{r_0} r^{d-1} \varphi(r)\,\mathrm{d}r = +\infty$$
and
$$u(x) \geq \varphi(|x|) \qquad \text{for} \quad |x| \leq r_0 .$$

2. There exists a decreasing function $\psi\colon [r_0, \infty) \to \mathbb{R}_0^+$ with
$$\int_{r_0}^{\infty} r^{d-1} \psi(r)\,\mathrm{d}r < \infty$$
and
$$|u(x)| \leq \psi(|x|) \qquad \text{for} \quad |x| \geq r_0 . \tag{1}$$

For this class $\mathscr{U}$ 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
$$H_u(\mathbf{x}_m) = \sum_{1 \leq i < j \leq m} u(x_i - x_j)\,, \qquad \mathbf{x}_m = (x_1, \dots, x_m)\,,$$
is stable (cf. Dobrushin[8]), i.e., for every $u \in \mathscr{U}$ 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
$$\rho_{\Lambda_\ell}^{(m)}(\mathbf{x}_m) \;=\; \frac{e^{\beta\mu m}}{\Xi(\Lambda_\ell, \beta, \mu, u)} \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{N!} \int_{\Lambda_\ell^N} e^{-\beta H_u(\mathbf{x}_m, \mathbf{y}_N)}\,\mathrm{d}\mathbf{y}_N\,, \tag{2}$$
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 $\int_{\Delta^0} c\,\mathrm{d}\mathbf{x}_0$ with bounded domain $\Delta \subset \mathbb{R}^d$ is always taken to be equal to c, and the normalizing constant
$$\Xi(\Lambda_\ell, \beta, \mu, u) \;=\; \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{N!} \int_{\Lambda_\ell^N} e^{-\beta H_u(\mathbf{x}_N)}\,\mathrm{d}\mathbf{x}_N$$

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

We assume that for some sequence $(\ell_k)_k$ going to infinity as $k \to \infty$, these correlation functions converge uniformly on compact subsets to translation-invariant functions $\rho^{(m)} : (\mathbb{R}^d)^m \to \mathbb{R}^+_0$, $m \in \mathbb{N}_0$, 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.,

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

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

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

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 σ-algebra, cf. Ruelle[42]. 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

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

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

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

is the limiting particle counting density.

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

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

where $\gamma' = \gamma \cup \{x_1,\dots,x_N\}$, $\Gamma(\Delta^c) = \bigl\{\gamma \in \Gamma : \gamma \subset \mathbb{R}^d \setminus \Delta\bigr\}$, 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 $j^{(m)}_{\Lambda_\ell} : \Lambda_\ell^m \to \mathbb{R}$ for every $m \in \mathbb{N}_0$ and $\ell > 0$ via

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

These Janossy densities provide the induced probability density on Λ_ℓ, for which

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

for every $F \in L^1(\mathbb{P})$, which satisfies $F(\gamma) = F(\gamma \cap \Lambda_\ell)$, and which is given by functions $f_m : \Lambda^m \to \mathbb{R}$, $m \in \mathbb{N}_0$, 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 $u \in \mathscr{U}$, we denote by

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

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 (cf., e.g., Georgii[12]). We mention for later use that for almost every $x \in \mathbb{R}^d$ and $\mathbb{P}$-almost surely for every $\mathbb{P} \in \mathscr{G}$ the interaction defined in (6) is finite, and there holds

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

see Section 5 in Kondratiev and Kuna[23]. We mention further that there exists some $\mu_0 \in \mathbb{R}$ depending on β and $u \in \mathscr{U}$, such that for $\mu < \mu_0$ – the so-called *gas phase* – the set $\mathscr{G}(\beta,\mu,u)$ consists of a single Gibbs measure $\mathbb{P}$ only[42], and that in this case the correlation functions $\rho^{(m)}_{\Lambda_\ell}$ converge to $\rho^{(m)}$ as $\ell \to \infty$ for every $m \in \mathbb{N}_0$, cf.[41].

III. THE GIBBS VARIATIONAL PRINCIPLE

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

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

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

and the *specific (relative) entropy*

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

where both limits are known to exist in $\mathbb{R}\cup\{+\infty\}$: concerning the specific energy see Proposition 1 below; concerning the specific entropy we refer to Robinson and Ruelle[40] – in fact, using (7) and (3) it is not too difficult to see that $S(\mathbb{P})$ is finite for every $\mathbb{P}\in\mathscr{G}$. The relative entropy differs from the standard entropy by a sign; to simplify language we call $S(P)$ 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 Georgii[13].

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

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

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

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

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

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

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

$$\begin{aligned}
\int_\Gamma \sum_{\substack{x,y\in\gamma\cap\Lambda_\ell\\ x\neq y}} u(x-y)\;\mathrm{d}\mathbb{P}(\gamma)
&= \int_{\Lambda_\ell^2} u(x-y)\rho^{(2)}(x,y)\,\mathrm{d}(x,y)\\
&= \int_{\Lambda_\ell}\biggl(\int_{\Delta_{x,\ell}} u(x-y)\rho^{(2)}(x,y)\,\mathrm{d}y &\text{(13a)}\\
&\qquad + \int_{\Lambda_\ell\setminus\Delta_{x,\ell}} u(x-y)\rho^{(2)}(x,y)\,\mathrm{d}y\biggr)\mathrm{d}x, &\text{(13b)}
\end{aligned}$$

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). Therefore, and since $\rho^{(2)}$ is bounded, compare (3), the integral in (13b) is uniformly bounded, independent of ℓ 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 ℓ, 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

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

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

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

where

$$\begin{aligned}
\Delta_1 &= \{\,(x,y)\in\Lambda_\ell^2 \,:\, x\in\Lambda_{\ell-r}\,,\ |y-x|\leq r\,\}\,,\\
\Delta_2 &= \{\,(x,y)\in\Lambda_\ell^2 \,:\, x\in\Lambda_\ell\setminus\Lambda_{\ell-r}\,,\ |y-x|\leq r\,\}\,,\\
\Delta_3 &= \{\,(x,y)\in\Lambda_\ell^2 \,:\, |x-y|>r\,\}\,.
\end{aligned}$$

Under these assumptions on r and ℓ it follows from the fact that (14) is absolutely convergent that

$$\begin{aligned}
\Bigl|\int_{\Delta_2} u(x-y)\rho^{(2)}(x,y)\,\mathrm{d}(x,y)\Bigr|
&\leq \int_{\Lambda_\ell\setminus\Lambda_{\ell-r}} \int_{|y-x|\leq r} \bigl|u(x-y)\bigr|\rho^{(2)}(x,y)\,\mathrm{d}y\,\mathrm{d}x\\
&\leq \bigl(|\Lambda_\ell| - |\Lambda_{\ell-r}|\bigr)\int_{\mathbb{R}^3} \bigl|u(y)\bigr|\rho^{(2)}(y,0)\,\mathrm{d}y
\end{aligned}$$

and

$$\Big|\int_{\Delta_3} u(x-y)\rho^{(2)}(x,y)\,\mathrm{d}(x,y)\Big|$$
$$\le \int_{\Delta_3} \big|u(x-y)\big|\rho^{(2)}(x,y)\,\mathrm{d}(x,y)$$
$$\le \int_{\Lambda_\ell}\int_{|y-x|>r} \big|u(x-y)\big|\rho^{(2)}(x,y)\,\mathrm{d}y\,\mathrm{d}x$$
$$\le |\Lambda_\ell| \int_{|y|>r} \big|u(y)\big|\rho^{(2)}(y,0)\,\mathrm{d}y\,.$$

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

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

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

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

and hence,

$$\limsup_{\ell\to\infty}\left|\frac{1}{|\Lambda_\ell|}\int_{\Delta_1} u(x-y)\rho^{(2)}(x,y)\,\mathrm{d}(x,y) - \int_{\mathbb{R}^d} u(y)\rho^{(2)}(y,0)\,\mathrm{d}y\right| \quad (16)$$
$$\le \int_{|y|>r} \big|u(y)\big|\rho^{(2)}(y,0)\,\mathrm{d}y\,.$$

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

We mention that the Kirkwood-Salsburg equations (cf., e.g.,[42]) can be used to argue that the integrand of (12) is bounded near the origin when $\mathbb{P}$ is a (β,μ,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[20], and provide a rigorous proof, based on arguments borrowed from[20] and from the proof of Theorem 2.34 in[14].

Theorem 2. *Let $u,v \in \mathscr{U}$, $\beta > 0$, and $\mu,\mu' \in \mathbb{R}$ be given, and assume that $\mathbb{P}_u \in \mathscr{G}(\beta,\mu,u)$ and $\mathbb{P}_v \in \mathscr{G}(\beta,\mu',v)$ 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

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

and

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

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

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

and

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

and adding them we get

$$\begin{aligned}&\Omega_{\beta,\mu}(u,\mathbb{P}_v) - \Omega_{\beta,\mu}(u,\mathbb{P}_u)\\ &+ \Omega_{\beta,\mu'}(v,\mathbb{P}_u) - \Omega_{\beta,\mu'}(v,\mathbb{P}_v) \le 0\,.\end{aligned} \quad (19)$$

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

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

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

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

and

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

Inserting this into (20) we conclude that

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

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

It therefore follows from (5) that

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

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

$$\begin{aligned}H_u(\mathbf{x}_N) &+ W_u(\mathbf{x}_N;\gamma) - N\mu\\ &= H_v(\mathbf{x}_N) + W_v(\mathbf{x}_N;\gamma) - N\mu'\end{aligned} \tag{21}$$

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

$$W_u(x;\gamma) - \mu \;=\; W_v(x;\gamma) - \mu' \tag{22}$$

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

$$u(x-y) = v(x-y) \tag{23}$$

for almost every $x, y \in \Delta$, and since Δ 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 $\mathbb{P} \in \mathscr{G}$ there holds

$$W_u(x;\gamma\cap\Lambda_\ell) = W_v(x;\gamma\cap\Lambda_\ell) \tag{24}$$

$\mathbb{P}$ 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. □

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 $u, v \in \mathscr{U}$, $\beta > 0$, and $\mu \in \mathbb{R}$ be given, and assume that $\mathbb{P}_u \in \mathscr{G}(\beta,\mu,u)$ and $\mathbb{P}_v \in \mathscr{G}(\beta,\mu,v)$ 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 (β,μ,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 $\mathscr{G}(\beta,\mu,u)$ 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[41], compare Kuna[25] – that there exists at least one rotation and translation invariant Gibbs measure $\mathbb{P}_u \in \mathscr{G}(\beta,\mu,u)$, which can be used to define a *radial distribution function* $g : \mathbb{R}^+ \to \mathbb{R}_0^+$ given by

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

provided that the density is nonzero. Assuming further that $\mathbb{P}_v \in \mathscr{G}(\beta,\mu,v)$ 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 Λ_ℓ, 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[4] and Navrotskaya[33] 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 Tóth[47]. 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[6]. 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[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 $\rho^{(2)} : (\mathbb{R}^d)^2 \to \mathbb{R}$, such that there exists a pair potential $u \in \mathscr{U}$ for which ρ is the density and $\rho^{(2)}$ is the pair correlation function of a Gibbs measure $\mathbb{P} \in \mathscr{G}(\beta, \mu, u)$. Partial results for this problem have been contributed, e.g., by Caglioti, Kuna, Lebowitz, and Speer[2,26,27] and Koralov[24]. The general existence problem is widely open, though.

ACKNOWLEDGEMENTS

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 number 233630050 – TRR 146.

[1] 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** (2013), pp. 2108–2119.

[2] E. Caglioti, T. Kuna, J.L. Lebowitz, and E.R. Speer, *Point processes with specified low order correlations*, Markov Process. Related Fields, **12** (2006), pp. 257–272.

[3] 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** (2012), pp. 8383–8393.

[4] J.T. Chayes, L. Chayes, *On the validity of the inverse conjecture in classical density functional theory*, J. Stat. Phys., **36** (1984), pp. 471–488.

[5] J.T. Chayes, L. Chayes, E.H. Lieb, *The inverse problem in classical statistical mechanics*, Comm. Math. Phys., **93** (1984), pp. 57–121.

[6] D. Dereudre, *Variational principle for Gibbs point processes with finite range interaction*, Electron. Commun. Probab., **21** (2016), pp. 10–21.

[7] 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. Special Topics, **225** (2016), pp. 1317–1321.

[8] R. Dobrushin, *Investigation of conditions for the asymptotic existence of the configuration integral of Gibbs distribution*, Theor. Probability Appl., **9** (1964), pp. 566–581.

[9] F. Delbary, M. Hanke, and D. Ivanizki, *A generalized Newton iteration for computing the solution of the inverse Henderson problem*, arXiv:1806.11135.

[10] G. Gallavotti, *Statistical Mechanics: A Short Treatise*, Springer, Berlin, 1999.

[11] J.-P. Hansen and I.R. McDonald, *Theory of Simple Liquids*, Academic Press, Oxford, Fourth Ed., 2013.

[12] H.-O. Georgii, *Large deviations and the equivalence of ensembles for Gibbsian particle systems with superstable interaction*, Probab. Theory Related Fields, **99** (1994), pp. 171–195.

[13] ———, *The equivalence of ensembles for classical systems of particles*, J. Stat. Phys., **80** (1995), pp. 1341–1378.

[14] ———, *Gibbs measures and phase transitions*, De Gruyter, Berlin, Second Ed., 2011.

[15] H.-O. Georgii and H. Zessin, *Large deviations and the maximum entropy principle for marked point random fields*, Probab. Theory Related Fields, **96** (1993), pp. 177–204.

[16] J. Gibbs, *Elementary Principles in Statistical Mechanics*, Scribner's Sons, New York, 1902.

[17] R. Griffiths and D. Ruelle, *Strict convexity ("continuity") of the pressure in lattice systems*, Comm. Math. Phys., **23** (1971), pp. 169–175.

[18] M. Hanke, *A Taste of Inverse Problems. Basic Theory and Examples*, SIAM, Philadelphia, 2017.

[19] M. Heinen, *Calculating particle pair potentials from fluid-state pair correlations: Iterative Ornstein-Zernike inversion*, J. Comput. Chem., **39** (2018), pp. 1531–1543.

[20] R. Henderson, *A uniqueness theorem for fluid pair correlation functions*, Phys. Lett. A, **49** (1974), pp. 197–198.

[21] P. Hohenberg and W. Kohn, *Inhomogeneous electron gas*, Phys. Rev., **136** (1964), pp. B 864–B 871.

[22] N.M. Hugenholtz, *On the inverse problem in statistical mechanics*, Comm. Math. Phys., **85** (1982) pp. 27–38.

[23] Y.G. Kondratiev and T. Kuna, *Correlation functionals for Gibbs measures and Ruelle bounds*, Methods Funct. Anal. Topology, **9** (2003), pp. 9–58.

[24] L. Koralov, *An inverse problem for Gibbs fields with hard core potential*, J. Math. Phys., **48** (2007), 053301.

[25] T. Kuna, *Studies in configuration space analysis and applications*, PhD thesis, Rheinische Friedrich-Wilhelms-Universität, Bonn, 1999.

[26] T. Kuna, J.L. Lebowitz, and E.R. Speer, *Realizability of point processes*, J. Stat. Phys., **129** (2007), pp. 417–439.

[27] T. Kuna, J.L. Lebowitz, and E.R. Speer, *Necessary and sufficient conditions for realizability of point processes*, Ann. Appl. Probab., **21** (2011), pp. 1253-1281.

[28] D. Levesque, J.J. Weis, and L. Reatto, *Pair interaction from structural data for dense classical liquids*, Phys. Rev. Lett., **54** (1985), pp. 451–454.

[29] A.P. Lyubartsev and A. Laaksonen, *Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach*, Phys. Rev. E, **52** (1995), pp. 3730–3737.

[30] A.P. Lyubartsev, A. Naômé, D.P. Vercauteren, and A. Laaksonen, *Systematic hierarchical coarse-graining with the inverse Monte Carlo method*, J. Chem. Phys., **143** (2015), 243120.

[31] S.Y. Mashayak, L. Miao, and N.R. Aluru, *Integral equation theory based direct and accelerated systematic coarse-graining approaches*, J. Chem. Phys., **148** (2018), 214105.

[32] N. Mermin, *Thermal properties of the inhomogeneous electron gas*, Phys. Rev., **137** (1965), pp. A 1471–A 1473.

[33] I. Navrotskaya, *Inverse problem in classical statistical mechanics*, PhD Thesis, University of Pittsburgh, 2016.

[34] W.G. Noid, *Perspective: Coarse-grained models for biomolecular systems*, J. Chem. Phys., **139** (2013), 090901.

[35] C. Peter and K. Kremer, *Multiscale simulation of soft matter systems*, Faraday Discuss., **144** (2010), pp. 9–24.

[36] B.L. Peters, K.M. Salerno, A. Agrawal, D. Perahia, G.S. Grest, *Coarse-grained modeling of polyethylene melts: effect on dynamics*, J. Chem. Theor. Comput., **13** (2017), pp. 2890–2896.

[37] R. Potestio, C. Peter, and K. Kremer, *Computer simulations of soft matter: Linking the scales*, Entropy, **16** (2014), pp. 4199–4245.

[38] 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** (2005), 224106.

[39] D. Reith, M. Pütz, and F. Müller-Plathe, *Deriving effective mesoscale potentials from atomistic simulations*, J. Comput. Chem., **24** (2003), pp. 1624–1636.

[40] D.W. Robinson and D. Ruelle, *Mean entropy of states in classical statistical mechanics*, Comm. Math. Phys., **5** (1967), pp. 288–300.

[41] D. Ruelle, *Statistical Mechanics: Rigorous Results*, W.A. Benjamin Publ., New York, 1969.

[42] ———, *Superstable interactions in classical statistical mechanics*, Comm. Math. Phys., **18** (1970), pp. 127–159.

[43] W. Schommers, *A pair potential for liquid rubidium from the pair correlation function*, Physics Letters, **43A** (1973), pp. 157–158.

[44] ———, *Pair potentials in disordered many-particle systems: A study for liquid gallium*, Phys. Rev. A, **28** (1983), pp. 3599–3605.

[45] M.S. Shell, *The relative entropy is fundamental to multiscale and inverse thermodynamic problems*, J. Chem. Phys., **129** (2008), 144108.

[46] A.K. Soper, *Empirical potential Monte Carlo simulation of fluid structure*, Chemical Physics, **202** (1996), pp. 295–306.

[47] G. Tóth, *Interactions from diffraction data: historical and comprehensive overview of simulation assisted methods*, J. Phys.: Condens. Matter, **19** (2007), 335220.

[48] 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** (2017), 395.