The equivalence of ensembles for classical systems of particles

Hans-Otto Georgii
Mathematisches Institut der Universität München
Theresienstr. 39, D–80333 München

Abstract

For systems of particles in classical phase space with standard Hamiltonian, we consider (spatially averaged) microcanonical Gibbs distributions in finite boxes. We show that infinite–volume limits along suitable subsequences exist and are grand canonical Gibbs measures. On the way, we establish a variational formula for the thermodynamic entropy density, as well as a variational characterization of grand canonical Gibbs measures.

Key words: Classical statistical mechanics; microcanonical ensemble; variational principle; entropy; pressure.

1 Introduction

The equivalence of Gibbs ensembles is one of the central problems of statistical mechanics. As far as the thermodynamic functions are concerned, this question is already well understood. For example, it is well-known that, under suitable conditions on the interaction, the infinite–volume limits of the entropy per volume and of the Gibbs free energy per volume exist and are related to each other by a Legendre–Fenchel transform; see, for example, Refs. 3–5. A much deeper question is the equivalence of ensembles on the level of measures, which (in a possible but not optimal formulation) would mean that the microcanonical and the grandcanonical Gibbs distributions in finite boxes have the same infinite–volume limits. In this paper we prove a version of the last statement which is not affected by the possibility of phase transitions.

We consider the standard setting of classical statistical mechanics. A particle in Euclidean space of dimension $d \geq 1$ is characterized by a pair (x, p), where $x \in \mathsf{R}^d$ is the position and $p \in \mathsf{R}^d$ the momentum. A particle configuration (without multiple occupancies) is thus given by a set $\omega \subset \mathsf{R}^d \times \mathsf{R}^d$ of the form

$$\omega = \{(x, p_x) : x \in \hat{\omega}\},$$

where $\hat{\omega}$, the set of occupied positions, is a locally finite subset of R^d, and $p_x \in \mathsf{R}^d$ is the momentum of the particle at position x. The set of all such configurations ω is denoted by Ω. We assume that the formal Hamiltonian has the standard form

$$\begin{aligned} H(\omega) &\equiv H^{\mathrm{kin}}(\omega) + H^{\mathrm{pot}}(\omega) \\ &= \sum_{x\in\hat{\omega}} |p_x|^2 + \tfrac{1}{2} \sum_{x,y\in\hat{\omega},x\neq y} \varphi(x-y)\,, \end{aligned} \tag{1.1}$$

where φ is a pair potential satisfying suitable stability conditions. (For convenience we assume throughout that the particle mass is equal to $1/2$.)

We are interested in the behavior of microcanonical Gibbs distributions in finite boxes in the infinite volume limit. For simplicity we only consider the half-open cubes $\Lambda_n = [-n - 1/2, n + 1/2[^d$ of volume $v_n = (2n+1)^d, n \geq 0$, which will often considered as tori by assuming periodic boundary conditions. We write $M_{n|N,E,\mathrm{per}}$ for the microcanonical Gibbs distribution for N particles in Λ_n with the energy constraint $H_{n,\mathrm{per}} \leq E$, where $H_{n,\mathrm{per}}$ is the Hamiltonian in Λ_n with periodic boundary condition; see (2.5). We also consider the microcanonical Gibbs distributions $M^u_{n|N,E,\mathrm{per}}$ on the energy shells $\{E-u < H_{n,\mathrm{per}} \leq E\}$ of thickness $u > 0$, and the microcanonical distributions $M^0_{n|N,E,\mathrm{per}}$ on the energy surfaces $\{H_{n,\mathrm{per}} = E\}$; see (2.13), (2.15) and Lemma 6.3 for precise definitions. We shall also deal with other than periodic boundary conditions.

Here is an outline of our results. Under fairly general hypotheses on the pair interaction φ we will show that, in the limit as $n \to \infty$ and $N/v_n \to \rho, E/v_n \to \varepsilon$ for an admissible pair (ρ, ε) of particle and energy densities, the sequence $(M_{n|N,E,\mathrm{per}})_{n\geq 0}$ is relatively (sequentially) compact in a fairly strong topology, and each accumulation point belongs to the class $\mathcal{M}_{\rho,\varepsilon}$ of all translation invariant probability measures P on Ω which have expected particle density ρ and mean energy ε and maximize the mean entropy under this constraint (Theorem 3.3). Clearly, $\mathcal{M}_{\rho,\varepsilon}$ is contained in $\mathcal{G}_{z,\beta}$, the set of all P that minimize the mean free energy for a suitably chosen activity $z > 0$ and inverse temperature $\beta > 0$. Our second main result is a variational principle stating that $\mathcal{G}_{z,\beta}$ consists of all tempered, translation invariant (grand canonical) Gibbs measures for z, β (Theorem 3.4). These two results together imply that every accumulation point of the microcanonical distributions $M_{n|N,E,\mathrm{per}}$ is a Gibbs measure (in the grand canonical sense) for suitable parameters $z, \beta > 0$. This proves the equivalence of ensembles on the level of measures in a formulation which still makes sense in the presence of phase transitions. If there is no phase transition, i.e., if $\mathcal{G}_{z,\beta}$ contains a unique element P, then

$$M_{n|N,E,\mathrm{per}} \to P$$

in the above-mentioned topology. In fact, this convergence already holds when P is the unique tempered Gibbs measure with expected particle and energy densities (ρ, ε). A main step in the derivation of the preceding results will be to show that the thermodynamic entropy density can be characterized by a variational formula involving the mean entropy of translation invariant probability measures on Ω (Theorem 3.2). As a matter of fact, it is the variational characterization of thermodynamic quantities in terms of states which will enable us to lift the equivalence of ensembles from the level of thermodynamic functions to the level of states.

The same results will be obtained for microcanonical distributions on thick or thin energy shells, as well as for free or configurational boundary conditions under an additional spatial averaging. As is easy to see from the proofs, our results can also be extended to the case when the kinetic energy is given by any positive power of the momenta. They remain also true for pure positional ensembles of particles without momentum, at least as long as no microcanonical distributions on energy shells are considered.

Our proofs make essential use of some ideas from large deviation theory and rely, in particular, on the developments in Refs. 1 and 2. These papers will be referred to as I and II. For example, Theorem II.1 stands for Theorem 1 of Ref. 2.

In Section 2 we introduce our conditions on the pair interaction φ and describe the general setup. The main results are stated in Section 3. In Section 4 we investigate the mean entropy of translation invariant states, and in Section 5 we derive the variational characterization of the thermodynamic entropy density. The subject of Section 6 is the asymptotics of microcanonical distributions, whereas the final Section 7 is devoted to the variational priciple for Gibbs measures.

We conclude this introduction with a few bibliographical notes. A classical approach to the equivalence of ensembles (first proposed by Khinchin[(6)]) is to use a local central limit theorem for the particle number and energy. This was carried out in Refs. 7–10 both for continuous systems and for lattice systems. The drawback of this method is that it works only in situations where a good local limit theorem is available — which certainly requires, at least, the absence of phase transitions. For classical lattice systems, Martin–Löf[(11)] developed a "thermodynamic" approach to the equivalence of ensembles. It works also in the presence of phase transitions and makes essential use of translation invariance and variational principles, as we do here. A quite different approach is to define microcanonical resp. canonical Gibbs states directly in infinite volume in analogy to the familiar procedure in the grand canonical case, and to show that these are convex mixtures of the usual grand canonical Gibbs measures with different parameters. This idea was pursued in Refs. 12–17 both in the lattice and the continuous case. It is again not affected by the presence of phase transitions, but the underlying concept of equivalence — though natural in the context of infinite–volume time evolutions — is not the traditional concept which is the subject of the present paper.

2 The setting

2.1 The interaction

We start with our assumptions on the pair interaction φ, an even measurable function from R^d to $\mathsf{R}\cup\{\infty\}$.

(A1) φ is *regular*. That is, there exists a decreasing function $\psi : [0,\infty[\to [0,\infty[$ and a number $r(\varphi) < \infty$ such that

$$\begin{aligned} \varphi(x) &\geq -\psi(|x|) \quad \text{for all } x \in \mathsf{R}^d, \\ \varphi(x) &\leq \psi(|x|) \quad \text{whenever } |x| \geq r(\varphi) \end{aligned} \tag{2.1}$$

and

$$\int_0^\infty \psi(s) s^{d-1}\, ds < \infty \ .$$

(A2) Either $\varphi = \infty$ on a neighborhood of the origin (*hard core case*). Or $\varphi < \infty$ Lebesgue-almost everywhere, and φ is *non-integrably divergent at the origin*, in that there exists a decreasing function $\chi :]0,\infty[\to [0,\infty[$ such that

$$\varphi(x) \geq \chi(|x|) \ \text{ whenever } \ \varphi(x) \geq 0 \tag{2.2}$$

and

$$\int_0^1 \chi(s) s^{d-1}\, ds = \infty\ .$$

As is well-known[(3)], (A1) and (A2) together imply that φ is *superstable*. By definition, this means that there exist constants $a > 0, b < \infty$ such that for all n

$$H_n \geq a\, T_n - b\, N_n. \tag{2.3}$$

In the above, $H_n \equiv H_{n,\text{free}}$, the *Hamiltonian in the box* Λ_n *with free boundary condition*, is given by $H_n(\omega) = H(\omega_n)$, where H is as in (1.1) and $\omega_n = \omega \cap (\Lambda_n \times \mathsf{R}^d)$ is the restriction of $\omega \in \Omega$ to Λ_n; $N_n(\omega) = \text{card}(\hat{\omega}_n)$ is the particle number in Λ_n; and

$$T_n = \sum_{i \in \Lambda_n \cap \mathsf{Z}^d} N_{C+i}^2 \tag{2.4}$$

is the sum of the squared particle numbers $N_{C+i}(\omega) = \text{card}(\hat{\omega} \cap (C+i))$ in the disjoint unit cells $C+i$ in Λ_n, where $C = \Lambda_0 = [-1/2, 1/2[^d$ stands for the centered unit cube.

The superstability and regularity of φ will be sufficient as long as we deal only with free or periodic boundary conditions. The latter lead to the *periodic Hamiltonians*

$$\begin{aligned} H_{n,\text{per}}(\omega) &= \sum_{x \in \hat{\omega}_n} |p_x|^2 + \tfrac{1}{2} \sum_{x \in \hat{\omega}_n, y \in \hat{\omega}^{(n)}, y \neq x} \varphi(y-x) \\ &= H_n(\omega) + \tfrac{1}{2} \sum_{x \in \hat{\omega}_n, y \in \hat{\omega}^{(n)} \setminus \Lambda_n} \varphi(y-x). \end{aligned} \tag{2.5}$$

Here $\omega^{(n)} = \{(x + (2n+1)i, p) : (x,p) \in \omega_n, i \in \mathsf{Z}^d\}$ is the periodic continuation of ω_n.

The non-integrable divergence of φ at the origin will become important as soon as configurational boundary conditions are involved. For $t > 0$ we define

$$\Omega(t) = \{T_n / v_n \leq t \ \text{ for all } \ n \geq 0\}. \tag{2.6}$$

The configurations in $\Omega^* = \bigcup_{t>0} \Omega(t)$ are called *tempered.* If φ even has a hard core we use the same symbols $\Omega(t)$ and Ω^* to denote the set of all admissible hard-core configurations. (This avoids a messy distinction of cases.) For each $\zeta \in \Omega^*$ and $n \geq 0$ we let

$$H_{n,\zeta}(\omega) = H_n(\omega) + \sum_{x \in \hat{\omega}_n, y \in \hat{\zeta} \setminus \Lambda_n} \varphi(y-x) \tag{2.7}$$

denote the *Hamiltonian in* Λ_n *with tempered boundary condition* ζ. The regularity of φ ensures that the last sums exists, cf. Lemma II.4.2.

2.2 Translation invariant states

Next we consider the class of possible states of our particle system. The configuration space Ω is equipped with the σ-algebra $\mathcal{F}$ which is generated by the counting variables $N(B) : \omega \to \text{card}(\omega \cap B)$ for Borel sets $B \subset \mathsf{R}^d \times \mathsf{R}^d$. It is well-known[(18)] that $\mathcal{F}$ is the Borel σ-algebra for the Polish topology on Ω that is induced by the mappings $N(h) : \omega \to \sum_{x \in \hat{\omega}} h(x, p_x), h : \mathsf{R}^d \times \mathsf{R}^d \to \mathsf{R}$ continuous with compact support.

We consider the class $\mathcal{P}$ of all probability measures ("states") P on $(\Omega, \mathcal{F})$ satisfying

$$\int P(d\omega) \sum_{x \in \hat{\omega}_n} (1 + |p_x|^2) < \infty \quad \text{for all } n, \tag{2.8}$$

as well as the class $\mathcal{P}_\Theta$ of all $P \in \mathcal{P}$ which are invariant under the translation group $\Theta = (\vartheta_x)_{x \in \mathsf{R}^d}$ acting on Ω via $\vartheta_x \omega = \{(y - x, p) : (y, p) \in \omega\}$. The mapping $(x, \omega) \to \vartheta_x \omega$ is known[18] to be measurable.

For each $P \in \mathcal{P}_\Theta$ there exists a number $\rho(P) < \infty$, the *intensity* or *mean particle number* of P, such that $P(N_\Lambda) = \rho(P)|\Lambda|$ for all measurable $\Lambda \subset \mathsf{R}^d$; here we write $N_\Lambda = N(\Lambda \times \mathsf{R}^d)$ for the counting variable associated to Λ, $|\Lambda|$ is the Lebesgue measure of Λ, and $P(f) = \int f dP$ stands for the integral of a function f with respect to P.

Also, for $P \in \mathcal{P}_\Theta$ there exists[18,1] a unique finite measure P° on $\mathsf{R}^d \times \Omega$, the *Palm measure* of P, such that

$$\int P(d\omega) \sum_{x \in \hat{\omega}} F(x, p_x, \vartheta_x \omega) = \int dx \int P^\circ(dp, d\omega) F(x, p, \omega) \tag{2.9}$$

for all measurable functions $F \geq 0$ on $\mathsf{R}^d \times \mathsf{R}^d \times \Omega$. In particular, P° is supported on the set $\{(p, \omega) : (0, p) \in \omega\}$, and the first marginal $\mu_P = P^\circ(\cdot \times \Omega)$ of P° satisfies

$$\int P(d\omega) \sum_{x \in \hat{\omega}_n} f(p_x) = v_n \int f \, d\mu_P \tag{2.10}$$

for all $n \geq 0$ and measurable $f : \mathsf{R}^d \to [0, \infty[$. We call μ_P the *momentum intensity measure*.

We introduce a topology $\tau_\mathcal{L}$ on $\mathcal{P}$ (which is much finer than the weak topology associated to the above-mentioned Polish topology on Ω) as follows. Let $\mathcal{L}$ denote the class of all measurable functions $f : \Omega \to \mathsf{R}$ which are *local* and *tame* in the sense that $f(\omega) = f(\omega_\ell)$ and

$$|f(\omega)| \leq c\Big(1 + \sum_{x \in \hat{\omega}_\ell} (1 + |p_x|)\Big) \equiv c\left(1 + \tilde{N}_\ell(\omega)\right) \tag{2.11}$$

for some $\ell \geq 0, c < \infty$ and all $\omega \in \Omega$. The *topology $\tau_\mathcal{L}$ of local convergence* is then defined as the weak* topology on $\mathcal{P}$ relative to $\mathcal{L}$, i.e., as the smallest topology making the mappings $P \to P(f) \equiv \int f dP$ with $f \in \mathcal{L}$ continuous. In particular, the mean particle density $\rho(P)$ and the mean momentum per particle $\int p \, \mu_P(dp)/\rho(P)$ are continous functions of $P \in \mathcal{P}_\Theta$.

2.3 The Gibbs ensembles

For a fixed box Λ_n and particle number N we consider the set

$$\Omega_{n|N} = \{\omega \in \Omega : \hat{\omega} \subset \Lambda_n, N_n(\omega) = N\}$$

of all N–particle configurations in Λ_n, and the associated "Lebesgue measure" (or "Liouville measure")

$$L_{n|N}(A) = \frac{1}{N!} \int_{\Lambda_n^N} dx_1 \ldots dx_N \int_{\mathsf{R}^{dN}} dp_1 \ldots dp_N \, 1_A\left(\{(x_1, p_1), \ldots, (x_N, p_N)\}\right), \tag{2.12}$$

$A \in \mathcal{F}$. The *microcanonical distributions* are obtained by conditioning $L_{n|N}$ on an energy shell of the form $\{E - u < H_{n,\mathrm{bc}} \le E\}$ for some $E \in \mathsf{R}, 0 < u \le \infty$, and a boundary condition $\mathrm{bc} \in \Omega^* \cup \{\mathrm{per}\}$. That is, we set

$$M^u_{n|N,E,\mathrm{bc}}(d\omega) = 1_{\{E - u < H_{n,\mathrm{bc}} \le E\}}(\omega)\, L_{n|N}(d\omega)\; /\; Z^u_{n|N,E,\mathrm{bc}} \tag{2.13}$$

whenever the parameters are chosen in such a way that the normalizing constant

$$Z^u_{n|N,E,\mathrm{bc}} = L_{n|N}(E - u < H_{n,\mathrm{bc}} \le E) \tag{2.14}$$

is positive. ((2.3) implies that it is always finite.) In the case $u = \infty$ (which means that there is no lower bound on the energy) we shall omit the upper index u. $Z_{n|N,E,\mathrm{bc}}$ is called the *microcanonical partition function* or the *phase volume.*

We will also consider microcanonical distributions on thin energy shells $\{H_{n,\mathrm{bc}} = E\}$. Formally, these are defined by

$$M^0_{n|N,E,\mathrm{bc}} = \lim_{u \to 0} M^u_{n|N,E,\mathrm{bc}}. \tag{2.15}$$

It is well-known that these limits exists in the vague sense when $H_{n,\mathrm{bc}}$ is a C^2 function (which certainly holds if φ is C^2 with derivatives satisfying an analogue of (2.1)); see, for example, Khinchin[6] or Section 3.5 of Tjur[19]. However, we do not need any smoothness of φ. Indeed, since we are only interested in the asymptotic behavior of $M^0_{n|N,E,\mathrm{bc}}$ as $n \to \infty$, we only need to define $M^0_{n|N,E,\mathrm{bc}}$ for events in a microscopic subregion of Λ_n, and for such events the limit (2.15) exists without any additional assumption on φ (due to the smoothing effect of the kinetic energy). We defer the details until Section 6.

Finally, for the nonperiodic boundary conditions $\mathrm{bc} \in \Omega^*$, we will have to deal with the *averaged microcanonical distributions*

$$\bar{M}^u_{n|N,E,\mathrm{bc}} = v_n^{-1} \int_{\Lambda_n} dx\, M^u_{n|N,E,\mathrm{bc}} \circ \vartheta_x^{-1}. \tag{2.16}$$

This spatial averaging guarantees the asymptotic translation invariance we need.

In the grand canonical ensemble, the parameters N and E are replaced by the positive parameters z, the *activity*, and β, the *inverse temperature.* The *grand canonical Gibbs distribution* in Λ_n with boundary condition $\mathrm{bc} \in \Omega^* \cup \{\mathrm{per}\}$ and parameters z, β is defined by

$$G_{n,z,\beta,\mathrm{bc}}(d\omega) = \Xi^{-1}_{n,z,\beta,\mathrm{bc}}\, z^{N_n(\omega)} \exp[-\beta\, H_{n,\mathrm{bc}}(\omega)]\, L_n(d\omega), \tag{2.17}$$

where $L_n = \sum_{N \ge 0} L_{n|N}$ is the Lebesgue measure on $\Omega_n = \{\omega \in \Omega : \hat{\omega} \subset \Lambda_n\}$, and

$$\Xi_{n,z,\beta,\mathrm{bc}} = L_n(z^{N_n} \exp[-\beta H_{n,\mathrm{bc}}])$$

is the grand canonical partition function (which is finite, cf. the estimates in the proof of Lemma 5.2).

A measure $P \in \mathcal{P}$ is called a *tempered Gibbs measure* for $z, \beta > 0$ if $P(\Omega^*) = 1$ and, for all $n \ge 0$ and measurable functions $f \ge 0$ on Ω,

$$P(f) = \int P(d\zeta) \int G_{n,z,\beta,\zeta}(d\omega)\, f(\omega_n \cup (\zeta \setminus \zeta_n)). \tag{2.18}$$

Clearly, the form of our Hamiltonian implies that, relative to all $G_{n,z,\beta,\zeta}$ and thus all tempered Gibbs measures, the momenta of the particles are conditionally i.i.d. with a Maxwellian (i.e., normal) distribution when the set $\hat{\omega}$ of occupied positions is given. For functions f that depend only on $\hat{\omega}$, (2.18) is equivalent to the equilibrium equations introduced by Ruelle[20].

3 Results

A basic ingredient of our results is the existence of the mean energy and mean entropy of any $P \in \mathcal{P}_\Theta$. The *mean energy* is defined by

$$U(P) = \lim_{n\to\infty} v_n^{-1} P(H_n). \tag{3.1}$$

This limit exists in $\mathsf{R}\cup\{\infty\}$. Indeed, (2.10) and (2.8) show that the kinetic contribution equals

$$U^{\text{kin}}(P) = \int |p|^2 \mu_P(dp) = v_n^{-1} P(H_n^{\text{kin}}) < \infty \tag{3.2}$$

independently of n. On the other hand, it was shown in Theorem II.1 (under the assumptions (2.1) and (2.3)) that the mean potential energy

$$U^{\text{pot}}(P) = \lim_{n\to\infty} v_n^{-1} P(H_n^{\text{pot}}) \tag{3.3}$$

exists in $[a\rho(P)^2 - b\rho(P), \infty]$ and satisfies

$$U^{\text{pot}}(P) = \int P^\circ(dp, d\omega)\, \tfrac{1}{2} \sum_{0\neq x\in\hat{\omega}} \varphi(x) \tag{3.4}$$

if $P(N_C^2) < \infty$ and equals $+\infty$ otherwise. Moreover, (3.2), (3.4) and Theorem II.1 show that the functions U^{kin} and U^{pot}, and thus also U, on $\mathcal{P}_\Theta$ are affine (even measure affine, in that they are affine with respect to convex mixtures formed by arbitrary probability measures on $\mathcal{P}_\Theta$) and lower semicontinuous relative to $\tau_\mathcal{L}$.

The *entropy* of a state $P \in \mathcal{P}$ in Λ_n is defined by

$$S_n(P) = \begin{cases} -P_n(\log f_n) & \text{if } P_n \ll L_n \text{ with density } f_n \\ -\infty & \text{otherwise,} \end{cases} \tag{3.5}$$

where $P_n = P(\{\omega \in \Omega : \omega_n \in \cdot\})$ is the restriction of P to Λ_n. The following proposition on the mean entropy of invariant states will be proved in Section 4.

Proposition 3.1 *For each $P \in \mathcal{P}_\Theta$, the mean entropy*

$$S(P) = \lim_{n\to\infty} v_n^{-1} S_n(P)$$

exists in $\mathsf{R}\cup\{\infty\}$ and is a measure affine, upper semicontinuous function of P. The "energy-bounded" superlevel sets

$$\{P \in \mathcal{P}_\Theta : S(P) \geq -c, U(P) \leq \varepsilon\}$$

of S (with $c, \varepsilon \in \mathsf{R}$) are compact and sequentially compact in $\tau_\mathcal{L}$.

We emphasize that S fails to be upper semicontinuous with respect to the coarser topology $\tau_{\mathcal{L}^b}$ that is associated to the class $\mathcal{L}^b$ of all *bounded* local functions, see Example 4.3.

Our first main result concerns the existence and variational characterization of the thermodynamic entropy density. To state it we define for $\rho \geq 0$

$$\varepsilon_{\min}(\rho) = \inf \{U(P) : P \in \mathcal{P}_\Theta, \rho(P) = \rho, S(P) > -\infty\}. \tag{3.6}$$

By (2.3), $\varepsilon_{\min}(\rho) \geq a\rho^2 - b\rho$, and $\varepsilon_{\min}(\cdot)$ is clearly convex. Also, $\varepsilon_{\min}(\cdot)$ is finite and continuous on a maximal interval $[0, \rho_{\max}[$, where $\rho_{\max} = \infty$ except when φ has a hard core. These facts are proved in Lemma II.7.1 (which extends without difficulties to the present setting of particles with momentum). We introduce the convex set

$$\Sigma = \{(\rho, \varepsilon) : 0 < \rho < \rho_{\max}, \varepsilon > \varepsilon_{\min}(\rho)\} \tag{3.7}$$

and the abbreviation $\log_- u = \min(0, \log u)$.

Theorem 3.2 (a) *Let $(\rho, \varepsilon) \in \Sigma$ and $t > 0$. If $n \to \infty$ and $N, E, u,$ bc run through any sequences such that $N/v_n \to \rho, E/v_n \to \varepsilon, v_n^{-1} \log_- u \to 0$ and bc $\in \Omega(t) \cup \{\mathrm{per}\}$, the limit*

$$s(\rho, \varepsilon) = \lim v_n^{-1} \log Z^u_{n|N,E,\mathrm{bc}} \tag{3.8}$$

exists and admits the variational characterization

$$s(\rho, \varepsilon) = \sup \{ S(P) : P \in \mathcal{P}_\Theta, \rho(P) = \rho, U(P) \leq \varepsilon \}. \tag{3.9}$$

(b) *The function $s(\cdot, \cdot)$ on $[0, \infty[\times \mathrm{R}$ defined by (3.9) is concave and upper semicontinuous. The set Σ is the interior of the effective domain $\{s(\cdot, \cdot) > -\infty\}$, $s(\rho, \cdot)$ is strictly increasing on $]\varepsilon_{\min}(\rho), \infty[$ for each ρ, and the condition $U(P) \leq \varepsilon$ in (3.9) can be replaced by $U(P) = \varepsilon$.*

We note that the existence of the limit in (3.8) is already known for a long time, at least for bc = free [3,4]. So our main point is its variational characterization (3.9); it is the microcanonical counterpart of the more familiar variational formula for the pressure which appears in Proposition 7.1. For $(\rho, \varepsilon) \notin \Sigma$, we still have the relation

$$\limsup v_n^{-1} \log Z^u_{n|N,E,\mathrm{bc}} \leq s(\rho, \varepsilon),$$

as is easy to see from the proof of Theorem 3.2 in Section 5.

Theorem 3.2 is the essential step towards the following main result which will be proved in Section 6.

Theorem 3.3 *Suppose that $(\rho, \varepsilon) \in \Sigma, t > 0, N/v_n \to \rho$ and $E/v_n \to \varepsilon$ as $n \to \infty$, and $u \in [0, \infty]$ and bc $\in \Omega(t) \cup \{\mathrm{per}\}$ are allowed to vary with n. Then the sequences $(M^u_{n|N,E,\mathrm{per}})_{n \geq 0}$ and $(\bar{M}^u_{n|N,E,\mathrm{bc}})_{n \geq 0}$ are (well-defined for large n and) relatively sequentially compact in the topology $\tau_{\mathcal{L}}$, and every accumulation point belongs to the set*

$$\mathcal{M}_{\rho,\varepsilon} = \{ P \in \mathcal{P}_\Theta : \rho(P) = \rho, U(P) = \varepsilon, S(P) = s(\rho, \varepsilon) \}.$$

Our final task is to identify the "microcanonical equilibrium states" in $\mathcal{M}_{\rho,\varepsilon}$ as grand canonical Gibbs measures. For $(\rho,\varepsilon)\in\Sigma$ let $\beta=\beta(\rho,\varepsilon)>0$ be the derivative of $s(\rho,\cdot)$ at ε. This derivative is known to exist, see Rechtmann and Penrose[21] or Remark 6.5 below. The differentiability with respect to ρ seems to be unknown, but by concavity we can anyway find a number $z>0$ such that, for a suitable constant $p(z,\beta)\in\mathsf{R}$, the plane

$$(\rho',\varepsilon')\to p(z,\beta)+\beta\varepsilon'-\rho'\log z$$

is a tangent to $s(\cdot,\cdot)$ at (ρ,ε). It then follows that for all $P\in\mathcal{P}_\Theta$ with $U(P)<\infty$

$$S(P)\le s(\rho(P),U(P))\le p(z,\beta)+\beta U(P)-\rho(P)\log z$$

with equality when $P\in\mathcal{M}_{\rho,\varepsilon}$. Introducing the *mean free energy*

$$F_{z,\beta}(P)=\beta\, U(P)-\rho(P)\log z-S(P) \tag{3.10}$$

we find that $F_{z,\beta}(P)\ge -p(z,\beta)$ with equality for $P\in\mathcal{M}_{\rho,\varepsilon}$. Since $\mathcal{M}_{\rho,\varepsilon}\ne\emptyset$, we conclude that

$$p(z,\beta)=-\min\{\,F_{z,\beta}(P):P\in\mathcal{P}_\Theta\,\}, \tag{3.11}$$

which is (β times) the *pressure* (see Proposition 7.1), and

$$\mathcal{M}_{\rho,\varepsilon}\subset\mathcal{G}_{z,\beta}\equiv\{\,P\in\mathcal{P}_\Theta:F_{z,\beta}(P)=-p(z,\beta)\,\}\,. \tag{3.12}$$

The following variational principle asserts that $\mathcal{G}_{z,\beta}$ coincides with the set of all translation invariant tempered Gibbs measures. Together with Theorem 3.3 and (3.12) this completes the proof of the equivalence of ensembles.

Theorem 3.4 *For all $z,\beta>0$, the set of all translation invariant tempered Gibbs measures for z,β (as defined around (2.18)) coincides with the set $\mathcal{G}_{z,\beta}$ on which the mean free energy $F_{z,\beta}$ attains its minimum $-p(z,\beta)$.*

This result (which is well-known for lattice systems[22,23,24]) will be proved in Section 7.

Remark 3.5 The preceding results obviously imply that the set $\mathcal{G}_{z,\beta}$ of translation invariant tempered Gibbs measures is non-empty — which was proved first by Ruelle[20] and Dobrushin[21]. In fact, this follows already from (the proof of) Proposition 3.1 together with Theorem 3.4: $F_{z,\beta}$ has compact sublevel sets $\{P\in\mathcal{P}_\Theta:F_{z,\beta}(P)\le c\}, c\in\mathsf{R}$, and thus attains its minimum.

Remark 3.6 $\mathcal{G}_{z,\beta}$ also contains all accumulation points of the relatively compact sequence $(G_{n,z,\beta,\mathrm{per}})_{n\ge0}$. This can be proved in complete analogy to Theorem 3.3; compare Proposition II.7.4. Thus, in the absence of phase transition when $\mathcal{G}_{z,\beta}=\{P\}$ we have

$$P=\lim M^u_{n|N,E,\mathrm{per}}=\lim G_{n,z,\beta,\mathrm{per}}$$

in $\tau_{\mathcal{L}}$, and a similar statement holds for spatially averaged distributions with boundary conditions in $\Omega(t)$ for some t. This expresses the equivalence of ensembles in terms of finite volume Gibbs distributions.

4 The mean entropy of invariant states

In this section we prove Proposition 3.1. It is convenient to replace the infinite reference measures L_n in (3.5) by consistent probability measures, namely the (local restrictions of) Poisson point random fields. Let $\tau : \mathsf{R}^d \to \mathsf{R} \cup \{\infty\}$ be any function satisfying $c(\tau) \equiv \int e^{-\tau(p)} dp < \infty$ and Q^τ the *Poisson point random field* with intensity measure $\mu^\tau(dx, dp) = dx\, e^{-\tau(p)} dp$. That is, Q^τ is the unique measure in $\mathcal{P}_\Theta$ relative to which the particle numbers $N(B_1), \ldots, N(B_k)$ in disjoint phase space regions $B_1, \ldots, B_k \subset \mathsf{R}^d \times \mathsf{R}^d$ are independent and Poisson distributed with parameters $\mu^\tau(B_i), 1 \le i \le k$. By definition, its restriction Q_n^τ to Λ_n has the Radon–Nikodym density

$$q_n^\tau(\omega) = \exp[-\sum_{x \in \hat{\omega}_n} \tau(p_x) - v_n c(\tau)] \tag{4.1}$$

relative to L_n.

For any $P \in \mathcal{P}_\Theta$ we consider the relative entropy

$$I(P_n; Q_n^\tau) = \begin{cases} P_n(\log h_n) & \text{if } P_n \ll Q_n^\tau \text{ with density } h_n \\ \infty & \text{otherwise.} \end{cases}$$

Clearly, $P_n \ll Q_n^\tau$ with a density h_n if and only if $P_n \ll L_n$ with density $f_n = h_n q_n^\tau$. Hence we conclude from (3.5) and (2.10) that

$$\begin{aligned} S_n(P) &= -I(P_n; Q_n^\tau) - P(\log q_n^\tau) \\ &= -I(P_n; Q_n^\tau) + v_n \mu_P(\tau) + v_n c(\tau). \end{aligned} \tag{4.2}$$

We will use the following fact.

Lemma 4.1 *For all $P \in \mathcal{P}_\Theta$, the mean relative entropy*

$$I_\tau(P) = \lim_{n\to\infty} v_n^{-1} I(P_n; Q_n^\tau)$$

exists in $[0, \infty]$. The function I_τ on $\mathcal{P}_\Theta$ is measure affine, and its sublevel sets $\{P \in \mathcal{P}_\Theta : I_\tau(P) \le c\}, c \ge 0$, are compact and sequentially compact in the topology $\tau_{\mathcal{L}^b}$ that is induced by the class $\mathcal{L}^b$ of all bounded local functions. If τ is such that

$$\int e^{c|p| - \tau(p)} dp < \infty$$

for all $c > 0$ then the last assertion also holds in the topology $\tau_{\mathcal{L}}$.

Proof. The first two sentences follow from the analogous result for lattice models by identifying Ω with $(\Omega_0)^{\mathsf{Z}^d}$, cf. Theorems (15.12) and (15.20) of Ref. 24. The last assertion is proved in Proposition I.2.6. □

We are particularly interested in the case when $\tau(p) = \beta |p|^2$ for some $\beta > 0$. We then replace the index τ by β. In particular, Q^β is the ideal gas with particle density $c(\beta) = (\pi/\beta)^{d/2}$ and Maxwellian momenta of variance $1/2\beta$, and (4.2) takes the form

$$v_n^{-1} S_n(P) = -v_n^{-1} I(P_n; Q_n^\beta) + \beta\, U^{\text{kin}}(P) + c(\beta).$$

Since $U^{\rm kin}$ is finite, we conclude from Lemma 4.1 that $v_n^{-1}S_n(P)$ converges to

$$S(P) = -I_\beta(P) + \beta U^{\rm kin}(P) + c(\beta) \tag{4.3}$$

which is a measure affine function of P.

It remains to establish the continuity properties of S. To this end, we note first that a bound on the mean energy implies a bound on the mean kinetic energy. Indeed, since

$$U^{\rm pot}(P) \geq a\rho(P)^2 - b\rho(P) \geq -b^2/4a$$

by (2.3) and (3.3), we have the implication

$$U(P) \leq \varepsilon \Longrightarrow U^{\rm kin}(P) \leq \tilde{\varepsilon} \equiv \varepsilon + b^2/4a. \tag{4.4}$$

The proof of Proposition 3.1 is therefore completed by the following lemma.

Lemma 4.2 *The mean entropy S is upper semicontinous (relative to $\tau_{\mathcal{L}}$), and the restricted superlevel sets $\{S \geq c, U^{\rm kin} \leq \tilde{\varepsilon}\}$, $c, \tilde{\varepsilon} \in \mathsf{R}$, are compact and sequentially compact.*

Proof. In view of (4.3), the set in question is contained in the set $\{I_1 \leq \tilde{c}\}$ with $\tilde{c} = -c + \tilde{\varepsilon} + c(1)$, and the latter set is compact and sequentially compact by Lemma 4.1. We thus only need to show that $\{S \geq c\}$ is closed, i.e., that S is upper semicontinuous. Applying (4.2) to the function $\tau = |\cdot|$ we see that

$$S(P) = -I_{|\cdot|}(P) + \mu_P(|\cdot|) + c(|\cdot|)$$

for all P. But $P \to \mu_P(|\cdot|)$ is continuous by definition of $\tau_{\mathcal{L}}$, and Lemma 4.1 asserts that $I_{|\cdot|}$ is lower semicontinous relative to $\tau_{\mathcal{L}^b} \subset \tau_{\mathcal{L}}$. This gives the result. □

We complete this section with an example showing that S fails to be upper semicontinuous relative to $\tau_{\mathcal{L}^b}$.

Example 4.3 For each $k \geq 1$ let $B_k \subset \mathsf{R}^d$ be a Borel set of volume e^{k^2}/k which is disjoint from the unit cube C. Let τ and τ_k be such that $e^{-\tau} = 1_C$ and $e^{-\tau_k} = 1_C + e^{-k^2}1_{B_k}$. $q_n^{\tau_k}$ converges to q_n^τ L_n-almost everywhere, and thus in $L^1(L_n)$. Hence Q^{τ_k} converges in $\tau_{\mathcal{L}^b}$ (but not in $\tau_{\mathcal{L}}$) to Q^τ. By (4.2),

$$\begin{aligned} S(Q^{\tau_k}) &= -\int \tau_k(p) e^{-\tau_k(p)} dp + c(\tau_k) \\ &= k + 1 + 1/k \to \infty \end{aligned}$$

as $k \to \infty$, but on the other hand $S(Q^\tau) = 1$. □

5 The thermodynamic entropy density

This section contains the proof of Theorem 3.2. To begin we consider the function $s(\rho, \varepsilon)$ on $[0, \infty[\times \mathsf{R}$ defined by (3.9) (with the convention $\sup \emptyset = -\infty$).

Lemma 5.1 (a) *The supremum in (3.9) is attained whenever it is not $-\infty$.*

(b) $s(\cdot,\cdot)$ *is concave and upper semicontinuous, and its energy–bounded superlevel sets* $\{(\rho,\varepsilon) : s(\rho,\varepsilon) \geq c_1, \varepsilon \leq c_2\}, c_1, c_2 \in \mathsf{R}$, *are compact.*

(c) *The set* Σ *in (3.7) is the interior of* $\{s(\cdot,\cdot) > -\infty\}$.

Proof. Assertion (a), the upper semicontinuity, and the compactness of the energy–bounded superlevel sets are immediate consequences of Proposition 3.1 and the continuity of $\rho(\cdot)$. The concavity is clear because S, $\rho(\cdot)$ and U are affine. Assertion (c) follows straight from the definitions. □

Our main task is the proof of the convergence (3.8) towards the limit $s(\rho,\varepsilon)$ defined by (3.9). We shall proceed in three stages. In the first stage (which relies on Refs. I and II) we shall deal with "fattened" partition functions for which the particle number may range in a whole interval. In the second and third stage we shall remove the fattening by controlling the dependence of $Z_{n|N,E,\mathrm{bc}}$ on the parameters N and E.

The first stage is based on techniques from large deviation theory. A central object in this theory is the *translation invariant empirical field*

$$R_{n,\omega} = v_n^{-1} \int_{\Lambda_n} \delta_{\vartheta_x \omega^{(n)}} dx \in \mathcal{P}_\Theta \tag{5.1}$$

of a configuration ω in Λ_n. Here $\omega^{(n)}$ is as in (2.5). We shall take advantage from the formula

$$H_{n,\mathrm{per}}(\omega) = v_n \, U(R_{n,\omega}), \tag{5.2}$$

which follows from (3.2), (3.4) and the explicit formula (I.2.6) for the Palm measure $R^\circ_{n,\omega}$ of $R_{n,\omega}$. The following lemma allows us to proceed from periodic to non-periodic boundary conditions.

Lemma 5.2 (a) *For any* $\varepsilon < \infty$ *and* $\delta > 0$ *there exists an integer* $k = k(\varepsilon,\delta)$ *such that for all* n

$$\{H_n \leq \varepsilon v_n\} \cap \Omega_n \subset \{H_{n+k,\mathrm{per}} \leq (\varepsilon+\delta) v_n\}.$$

(b) *For given* $\varepsilon < \infty$ *and* $\delta, t > 0$ *there exists some* $n(\varepsilon,\delta,t)$ *such that for all* $\zeta \in \Omega(t)$ *and* $n \geq n(\varepsilon,\delta,t)$

$$\{H_{n,\zeta} \leq \varepsilon v_n\} \subset \{H_n \leq (\varepsilon+\delta) v_n\}.$$

Proof. (a) If $H_n \leq \varepsilon v_n$ then, by (2.3),

$$N_n^2 / v_n \leq T_n \leq v_n \, \varepsilon / a + b \, N_n / a$$

and thus $T_n \leq t v_n$ for some $t < \infty$ depending on ε (and a, b). The result thus follows from Lemma II.4.3.

(b) Here we use the non-integrability of φ at 0. By Lemma II.6.1 there exists an increasing function $h : \mathsf{Z}_+ \to [0,\infty[$ with $h(\ell)/\ell^2 \to \infty$ as $\ell \to \infty$, and a number $b' > 0$ such that

$$H_n \geq T_n^h - b' \, N_n \tag{5.3}$$

for all n, where

$$T_n^h = \sum_{i\in\Lambda_n\cap\mathbf{Z}^d} h(N_{C+i}).$$

On the other hand, Lemma II.6.3 asserts that for given $\delta' > 0$ and n larger than some $n_0(\delta', t)$

$$H_{n,\zeta} \geq H_n - \delta' T_n^h - \delta' v_n \tag{5.4}$$

for all $\zeta \in \Omega(t)$. Suppose now that $H_{n,\zeta} \leq \varepsilon v_n$. (5.3) and (5.4) (with $\delta' = 1/2$) together then show that

$$T_n^h \leq (2\varepsilon + 1)v_n + 2b' N_n$$

whenever $n \geq n_0(1/2, t)$. Choosing q such that $h(\ell) \geq \ell^2$ for $\ell \geq q$ we see that

$$N_n^2/v_n \leq T_n \leq T_n^h + q^2 v_n.$$

Combining this with the previous inequality we find that $T_n^h \leq s\, v_n$ for some $s = s(\varepsilon) < \infty$ and all $n \geq n_0(1/2, t)$. Using (5.4) again with $\delta' = \delta/(s+1)$ we can conclude. $\square$

Here is the first stage in the proof of (3.8).

Proposition 5.3 *Let $\varepsilon < \infty$ and $D \subset [0, \infty[$ be a nondegenerate interval such that $(\rho, \varepsilon) \in \Sigma$ for some $\rho \in D$. Let $t > 0$ and consider, for each n,*

$$a_{n|D,\varepsilon,\mathrm{bc}} \equiv L_n(N_n \in v_n D, H_{n,\mathrm{bc}} \leq v_n\varepsilon) = \sum_{N\in v_n D} Z_{n|N,v_n\varepsilon,\mathrm{bc}}\ ,$$

where $\mathrm{bc} \in \Omega(t) \cup \{\mathrm{per}\}$ *may depend on n. Then*

$$\lim_{n\to\infty} v_n^{-1} \log a_{n|D,\varepsilon,\mathrm{bc}} = \sup_{\rho\in D} s(\rho, \varepsilon).$$

Proof. 1) We first derive the upper bound. We start with the case bc = per. Let $\beta > 0$ be arbitrary. Eq. (4.1) with $\tau = \beta|\cdot|^2$, (5.2) and (4.4) yield

$$\begin{aligned} a_{n|D,\varepsilon,\mathrm{per}} &= Q_n^\beta\Big(1/q_n^\beta\,;\, \rho(R_n) \in D, U(R_n) \leq \varepsilon\Big) \\ &\leq \exp[v_n\beta\tilde\varepsilon + v_n c(\beta)]\ Q_n^\beta(R_n \in A), \end{aligned}$$

where $A = \{P \in \mathcal{P}_\Theta : \rho(P) \in \bar D, U(P) \leq \varepsilon\}$ and $\bar D$ is the closure of D. Since A is closed in $\tau_{\mathcal{L}}$, we can use Theorem I.3.1 (with $Q = Q^\beta$) to obtain

$$\begin{aligned} \limsup v_n^{-1} \log a_{n|D,\varepsilon,\mathrm{per}} &\leq \beta\tilde\varepsilon + c(\beta) - \inf_{P\in A} I_\beta(P) \\ &= \beta\tilde\varepsilon + \sup_{P\in A}[S(P) - \beta\, U^{\mathrm{kin}}(P)] \\ &\leq \beta\tilde\varepsilon + \sup_{\rho\in D} s(\rho, \varepsilon). \end{aligned}$$

The second step comes from (4.3), and in the last step we used the positivity of U^{kin} and (3.9) together with the fact that (in view of our assumption on D and Lemma 5.1) the suprema of $s(\cdot, \varepsilon)$ over D and $\bar D$ coincide. Letting $\beta \to 0$ yields the upper bound in the case bc = per.

Turning to the case bc = free, we let D' be any interval containing D in its interior, $\delta > 0$, and k be chosen according to Lemma 5.2 (a). Then

$$a_{n|D,\varepsilon,\mathrm{free}} \leq a_{n+k|D',\varepsilon+\delta,\mathrm{per}}$$

whenever n is so large that v_{n+k}/v_n is sufficiently close to 1. On the other hand, Lemma 5.1 (b) implies that

$$\inf_{D' \supset D, \delta > 0} \sup_{\rho \in D'} s(\rho, \varepsilon + \delta) \leq \sup_{\rho \in D} s(\rho, \varepsilon) \ .$$

The upper bound for bc = free thus follows from that for bc = per.

Finally, suppose bc $= \zeta \in \Omega(t)$ for some $t > 0$, and choose any $\delta > 0$. Lemma 5.2 (b) then shows that

$$a_{n|D,\varepsilon,\zeta} \leq a_{n|D,\varepsilon+\delta,\mathrm{free}}$$

for sufficiently large n, and the desired upper bound follows from the previous case.

2) We now turn to the lower bound. Since $U^{\mathrm{kin}} \geq 0$ we obtain as above

$$a_{n|D,\varepsilon,\mathrm{per}} \geq \exp[v_n c(\beta)]\, Q_n^\beta(R_n \in A^\circ)$$

for any $\beta > 0$. Here $A^\circ = \{\rho(\cdot) \in D^\circ, U < \varepsilon\}$ and D° is the interior of D. Note that A° fails to be open because U is only lower semicontinuous. A direct application of Theorem I.3.1 is therefore impossible. We rather need to extend the lower bound in Lemma II.7.2 to the present case of particles with momentum. To see that such an extension is valid we only need to control the additional kinetic term in Lemma II.5.1. (The Φ there is our U^{pot}.) Using the notation in the proof of this lemma (in particular $m = n + k$ for a fixed k) we can write

$$U^{\mathrm{kin}}(P^{(n)}) = v_m^{-1} P^{(n)}(H_m^{\mathrm{kin}}) = v_m^{-2} \int_{\Lambda_m} \hat{P}^{(n)}(H_m^{\mathrm{kin}} \circ \vartheta_x) dx \ .$$

Since H_m^{kin} is a sum of single–particle terms and $\hat{P}^{(n)}$ is Λ_m-periodic, the integrand above does not depend on x. Hence

$$\begin{aligned} U^{\mathrm{kin}}(P^{(n)}) &= v_m^{-1} \hat{P}^{(n)}(H_m^{\mathrm{kin}}) &&= v_m^{-1} P_n(H_n^{\mathrm{kin}}|\Gamma_{q(n)}) \\ &\leq p_n^{-1} v_m^{-1}\, P_n(H_n^{\mathrm{kin}}) &&= (v_n/p_n v_m)\, U^{\mathrm{kin}}(P) \end{aligned}$$

and therefore $\limsup_{n\to\infty} U^{\mathrm{kin}}(P^{(n)}) \leq U^{\mathrm{kin}}(P)$. This is all what is needed to extend Lemmas II.5.1 and II.7.2 to the present case. The result is

$$\begin{aligned} \liminf_{n\to\infty} v_n^{-1} \log a_{n|D,\varepsilon,\mathrm{per}} &\geq c(\beta) - \inf_{P \in A^\circ} I_\beta(P) \\ &= \sup_{P \in A^\circ} [S(P) - \beta\, U^{\mathrm{kin}}(P)] \\ &\geq \sup_{\rho \in D} s(\rho, \varepsilon) - \beta\, \tilde{\varepsilon} \ . \end{aligned}$$

In the last step we used (4.4) and the fact that the suprema of $s(\cdot, \varepsilon)$ over D° and D coincide. Letting $\beta \to 0$ we obtain the lower bound in the case bc = per. The other boundary conditions can be treated by the same argument, with one addition: For bc = free we also have to use the last estimate in the proof of Proposition II.5.4, whereas in the case bc $= \zeta \in \Omega(t)$ we argue as in the lines leading to Eq. (II.6.6). $\square$

The second step in the proof of (3.8) consists in controlling the variation of the microcanonical partition functions $Z_{n|N,E,\mathrm{bc}}$ with respect to N. We need to distinguish the two cases in assumption (A2).

Lemma 5.4 *Suppose that $\varphi < \infty$ a.e.. Then for any given numbers $\nu, t > 0$ there exist constants $c, u > 0$ and n_0 such that*

$$Z_{n|N,E,\mathrm{bc}} \geq c\, Z_{n|N-1,E-u,\mathrm{bc}}$$

whenever $N \leq \nu v_n$, $E \in \mathsf{R}$, $\mathrm{bc} \in \Omega(t) \cup \{\mathrm{per}\}$ and $n \geq n_0$.

Proof. Let $\varphi_+ = \max(\varphi, 0)$ be the positive part of φ. Since $\varphi < \infty$ a.e., we can find some $q > 0$ such that

$$\int \min(1, \varphi_+(x)/q)dx \leq 1/4\nu \ .$$

Indeed, the integrand in the integral on the left side converges to $1_{\{\varphi_+=\infty\}} = 0$ a.e. as $q \to \infty$, and the regularity assumption (A1) ensures that the dominated convergence theorem is applicable. Next, for each $n, N, \omega \in \Omega_{n|N-1}$ and $(x,p) \notin \omega$ we can write

$$H_{n,\mathrm{bc}}(\omega \cup \{(x,p)\}) = H_{n,\mathrm{bc}}(\omega) + |p|^2 + \sum_{y \in \hat{\omega}} \varphi(y-x) + h_{n,\mathrm{bc}}(x),$$

where, for example,

$$h_{n,\zeta}(x) = \sum_{y \in \hat{\zeta} \setminus \Lambda_n} \varphi(y-x)$$

for $\zeta \in \Omega(t)$. Using (2.1) and the arguments in Lemma II.4.2 we see that there exists some $r \geq r(\varphi)$ such that $h_{n,\mathrm{bc}}(x) \leq 1$ whenever $x \in \Lambda_{n-r}$, $\mathrm{bc} \in \Omega(t) \cup \{\mathrm{per}\}$ and, in the case $\mathrm{bc} = \mathrm{per}$, $\omega \in \Omega_{n|N}$ with $N \leq \nu v_n$. Let n_0 be so large that $v_{n-r} \geq v_n/2$ when $n \geq n_0$. Then we have for arbitrary $n \geq n_0$ and $\omega \in \Omega_{n|N-1}$

$$\begin{aligned} |\{x \in \Lambda_{n-r} : \sum_{y \in \hat{\omega}} \varphi(y-x) \leq q\}| \ &\geq \ \int_{\Lambda_{n-r}} \Big(1 - \sum_{y \in \hat{\omega}} \min(1, \varphi_+(y-x)/q)\Big) dx \\ &\geq \ v_{n-r} - (N-1)/4\nu \ \geq \ v_n/4 \ . \end{aligned}$$

Setting $u = q + 2$ we thus obtain the final estimate

$$\begin{aligned} N\, Z_{n|N,E,\mathrm{bc}} \ &\geq \ \int L_{n|N-1}(d\omega) \int_{\Lambda_{n-r}} dx \int dp \\ &\qquad 1_{\{H_{n,\mathrm{bc}}(\omega) \leq E-u, \sum_{y \in \hat{\omega}} \varphi(y-x) \leq q, |p|^2 \leq 1\}} \\ &\geq \ (c_1 v_n/4)\, Z_{n|N-1,E-u,\mathrm{bc}} \ , \end{aligned}$$

where c_1 is the volume of the unit ball in R^d. The lemma thus follows with $c = c_1/4\nu$. □

If φ has a hard core we have an estimate in the opposite direction.

Lemma 5.5 *Suppose $\varphi = \infty$ on a neighbourhood of 0. For given $\nu > 0$ and $\varepsilon < \infty$ there exist constants $c, u > 0$ and n_0 such that*

$$Z_{n|N,E,\mathrm{bc}} \geq c\, Z_{n|N+1,E-u,\mathrm{bc}}$$

for all $n \geq n_0, N \geq \nu v_n, E \leq \varepsilon v_n,$ *and* $\mathrm{bc} \in \Omega^* \cup \{\mathrm{per}\}$.

Proof. The hard–core property and the regularity (2.1) imply the existence of a number $u > 0$ such that

$$\sum_{y \in \hat{\omega}} \varphi(y - x) \geq -u$$

for all $x \in \mathsf{R}^d$ and all admissible hard–core configurations $\hat{\omega}$. This gives the estimate

$$\begin{aligned}(N+1)\, Z_{n|N+1,E-u,\mathrm{bc}} &\leq \int L_{n|N}(d\omega) \int_{\Lambda_n} dx \int dp\; 1_{\{H_{n,\mathrm{bc}}(\omega) + |p|^2 \leq E\}} \\ &= v_n \int dp\; Z_{n|N,E-|p|^2,\mathrm{bc}}\ .\end{aligned}$$

In view of Lemma 5.7 below, the last expression is at most

$$v_n \int dp\; e^{-\gamma|p|^2}\, Z_{n|N,E,\mathrm{bc}}$$

for some $\gamma = \gamma(\nu, \varepsilon) > 0$, provided $N \geq \nu v_n, E \leq \varepsilon v_n$ and n is large enough. The lemma is now obvious. □

The next result marks the second stage in the proof of (3.8).

Proposition 5.6 *Let* $(\rho, \varepsilon) \in \Sigma$ *and* $t > 0$. *In the limit as* $n \to \infty, N/v_n \to \rho, E/v_n \to \varepsilon$ *and* $\mathrm{bc} = \mathrm{bc}(n) \in \Omega(t) \cup \{\mathrm{per}\}$, $v_n^{-1} \log Z_{n|N,E,\mathrm{bc}}$ *converges to the function* $s(\rho, \varepsilon)$ *defined by (3.9).*

Proof. If $D \subset [0, \infty[$ is a non-degenerate interval containing ρ in its interior and $\varepsilon' > \varepsilon$ then $Z_{n|N,E,\mathrm{bc}} \leq a_{n|D,\varepsilon',\mathrm{bc}}$ eventually. Proposition 5.3 and the upper semicontinuity of $s(\cdot, \cdot)$ thus imply that

$$\limsup v_n^{-1} \log Z_{n|N,E,\mathrm{bc}} \leq s(\rho, \varepsilon)\ .$$

For the lower bound we consider first the case when $\varphi < \infty$ a.e.. We fix any $\nu > \rho$ and let u and $0 < c < 1$ be as in Lemma 5.4. Let $\delta > 0$ be so small that $\delta < \rho$ and $\varepsilon - u\delta > \varepsilon_{\min}(\rho)$. For sufficiently large n we have $N/v_n \leq \nu$ and, by k–fold iteration of Lemma 5.4,

$$Z_{n|N,E,\mathrm{bc}} \geq c^k\, Z_{n|N-k,E-ku,\mathrm{bc}}\ .$$

Averaging the right-hand side over all k between 0 and δv_n and choosing any $\varepsilon' \in]\varepsilon_{\min}, \varepsilon - \delta[$ and a non-degenerate closed interval $D \subset]\rho - \delta, \rho[$ we obtain

$$Z_{n|N,E,\mathrm{bc}} \geq c^{\delta v_n}\, (\delta v_n)^{-1}\, a_{n|D,\varepsilon',\mathrm{bc}}$$

when n is large enough. Proposition 5.3 thus gives

$$\liminf v_n^{-1} \log Z_{n|N,E,\mathrm{bc}} \geq \delta \log c + s(\rho', \varepsilon')$$

for any $\rho' \in D$. Letting $\rho' \to \rho, \delta \to 0, \varepsilon' \to \varepsilon$ in this order we can conclude. In the case when φ has a hard core we can proceed analogously, using Lemma 5.5 instead of 5.4. □

The final step in the proof of (3.8) is the consideration of energy shells. To this end we need to control the E–dependence of the $Z_{n|N,E,\mathrm{bc}}$'s. Let $\hat{L}_{n|N}$ be the "Lebesgue measure" on the set $\hat{\Omega}_{n|N} = \{\hat{\omega} \subset \Lambda_n : \mathrm{card}\,\hat{\omega} = N\}$ of position–configurations of N particles in Λ_n (with the standard σ–algebra), that is,

$$\int f\, d\hat{L}_{n|N} = \frac{1}{N!} \int_{\Lambda_n^N} dx_1 \dots dx_N f(\{x_1, \dots, x_N\}) \tag{5.5}$$

for measurable $f \geq 0$ on $\hat{\Omega}_{n|N}$. Note that $H^{\mathrm{pot}}_{n,\mathrm{bc}}$ can be considered as a function of $\hat{\omega}$.

Lemma 5.7 *For all n, N, E and* bc *we have*

$$Z_{n|N,E,\mathrm{bc}} = c_N \int \hat{L}_{n|N}(d\hat{\omega}) \Big(E - H^{\mathrm{pot}}_{n,\mathrm{bc}}(\hat{\omega})\Big)_+^{Nd/2} \;,$$

where c_N is the volume of the unit ball in R^{Nd}. In particular, $Z_{n|N,E,\mathrm{bc}}$ is differentiable with respect to E with derivative

$$Z'_{n|N,E,\mathrm{bc}} = c_N (Nd/2)\; \hat{L}_{n|N}\Big((E - H^{\mathrm{pot}}_{n,\mathrm{bc}})_+^{(Nd/2)-1}\Big) \;. \tag{5.6}$$

which is sometimes called the structure function[(6)]*. Also, for every $\nu > 0$ and $\varepsilon < \infty$ there exists a constant $\gamma = \gamma(\nu, \varepsilon, t) > 0$ such that*

$$Z_{n|N,E-u,\mathrm{bc}} \leq e^{-\gamma u}\, Z_{n|N,E,\mathrm{bc}} \tag{5.7}$$

whenever $N \geq \nu v_n, E \leq \varepsilon v_n, u \geq 0$, $\mathrm{bc} \in \Omega(t) \cup \{\mathrm{per}\}$ for a given $t > 0$, and n is large enough.

Proof. We only need to comment on (5.7). Let $t > 0$ be fixed. By the proof of Lemma 5.2(b) there exists a number $c > -\varepsilon$ (depending on ε and t) such that, for all $\mathrm{bc} \in \Omega(t) \cup \{\mathrm{per}\}$ and sufficiently large n, $H^{\mathrm{pot}}_{n,\mathrm{bc}} \geq -c\, v_n$ whenever $H^{\mathrm{pot}}_{n,\mathrm{bc}} \leq \varepsilon\, v_n$. Hence $(E - H^{\mathrm{pot}}_{n,\mathrm{bc}})_+ \leq (\varepsilon + c) v_n$ and therefore

$$\frac{d}{dE} \log Z_{n|N,E,\mathrm{bc}} = Z'_{n|N,E,\mathrm{bc}} \;/\; Z_{n|N,E,\mathrm{bc}} \geq \gamma \equiv \nu d / 2(\varepsilon + c) \;. \tag{5.8}$$

(5.7) now follows by integration. □

It follows immediately from (5.7) and Proposition 5.6 that $s(\rho, \cdot)$ is strictly increasing on $]\varepsilon_{\min}(\rho), \infty[$. This in turn shows that the condition "$\leq \varepsilon$" in (3.9) can be replaced by "$= \varepsilon$". (5.7) further implies that

$$(1 - e^{-\gamma u})\, Z_{n|N,E,\mathrm{bc}} \leq Z^u_{n|N,E,\mathrm{bc}} \leq Z_{n|N,E,\mathrm{bc}}$$

for a suitable $\gamma > 0$ whenever $N/v_n \to \rho, E/v_n \to \varepsilon, u > 0$, $\mathrm{bc} \in \Omega(t) \cup \{\mathrm{per}\}$ for a given t, and n is large enough. Combining this inequality with Proposition 5.6 we arrive at (3.8). The proof of Theorem 3.2 is therefore complete.

Remark 5.8 Under the conditions of (3.8) we also have that

$$s(\rho, \varepsilon) = \lim v_n^{-1} \log Z'_{n|N,E,\mathrm{bc}} \;;$$

cf. (5.6). This follows from (5.8) and the inequality

$$Z'_{n|N,E,\mathrm{bc}} \leq \int_0^1 du\, Z'_{n|N,E+u,\mathrm{bc}} = Z^1_{n|N,E+1,\mathrm{bc}} \;.$$

□

6 The asymptotics of microcanonical distributions

This section contains the proof of Theorem 3.3. Let $(\rho, \varepsilon) \in \Sigma$ and

$$N/v_n \to \rho, E/v_n \to \varepsilon, \text{ and } \mathrm{bc} \in \Omega(t) \cup \{\mathrm{per}\} \tag{6.1}$$

for some fixed $t > 0$ as $n \to \infty$. In a first step, we shall prove Theorem 3.3 under the additional hypothesis

$$v_n^{-1} \log_- u \to 0 \text{ as } n \to \infty \tag{6.2}$$

which appears in Theorem 3.2. In this case we follow the idea devised in Ref. 26. In a second step we then shall extend the result to the alternate case when u vanishes or tends to 0 exponentially with v_n.

Suppose (6.1) and (6.2) hold. For brevity we write $M_n = M^u_{n|N,E,\mathrm{bc}}$ and $\bar{M}_n = \bar{M}^u_{n|N,E,\mathrm{bc}}$. Theorem 3.2 implies that these measures are well-defined for sufficiently large n. Let $M_n^{\mathrm{per}} \in \mathcal{P}$ denote the measure relative to which the configurations in the disjoint blocks $\Lambda_n + (2n+1)i, i \in \mathsf{Z}^d$, are independent with identical distribution M_n, and

$$\widetilde{M}_n = v_n^{-1} \int_{\Lambda_n} M_n^{\mathrm{per}} \circ \vartheta_x^{-1} \, dx \in \mathcal{P}_\Theta$$

the associated invariant average. We shall derive the asymptotics of $\bar{M}_n$ from that of $\widetilde{M}_n$. First, we get from (3.2) that

$$U^{\mathrm{kin}}(\widetilde{M}_n) = v_n^{-2} \int_{\Lambda_n} M_n^{\mathrm{per}}(H_n^{\mathrm{kin}} \circ \vartheta_x) \, dx = v_n^{-1} M_n(H_n^{\mathrm{kin}}) \,. \tag{6.3}$$

The second equality follows from the periodicity of M_n^{per} and the additive nature of H_n^{kin}. Together with Lemma 5.2(b) and a version of (4.4) this implies that

$$\limsup_{n\to\infty} U^{\mathrm{kin}}(\widetilde{M}_n) < \infty. \tag{6.4}$$

Next we consider the mean entropy of $\widetilde{M}_n$. The following key estimate follows from Theorem 3.2.

Proposition 6.1 *Under conditions (6.1) and (6.2),* $\liminf_{n\to\infty} S(\widetilde{M}_n) \geq s(\rho, \varepsilon)$.

Proof. Let $Q = Q^1$ and $I = I_1$ be as in (4.3) for $\beta = 1$. Lemma I.5.5 asserts that $I(\widetilde{M}_n) \leq v_n^{-1} I(M_n; Q_n)$. Together with (4.3), (6.3) and (4.2) this yields

$$\begin{aligned} S(\widetilde{M}_n) &\geq -v_n^{-1} I(M_n; Q_n) + v_n^{-1} M_n(H_n^{\mathrm{kin}}) + c(1) \\ &= v_n^{-1} S_n(M_n) = v_n^{-1} \log Z^u_{n|N,E,\mathrm{bc}} \,, \end{aligned}$$

and Theorem 3.2 gives the result. □

Let us say that two sequences $(P_{1,n})$ and $(P_{2,n})$ in $\mathcal{P}$ are *asymptotically equivalent*, and write $P_{1,n} \sim P_{2,n}$ as $n \to \infty$, if for all $f \in \mathcal{L}$

$$\lim_{n\to\infty} \Big(P_{1,n}(f) - P_{2,n}(f) \Big) = 0 \,.$$

Considering again the empirical fields R_n in (5.1) we define the measures $M_n R_{n+k} = \int M_n(d\omega) R_{n+k,\omega} \in \mathcal{P}_\Theta$.

Lemma 6.2 *Assuming (6.1) and (6.2), we have for each* $k \geq 0$

$$\widetilde{M}_n \sim M_n R_{n+k} \sim \bar{M}_n \text{ as } n \to \infty .$$

If bc = per *then, in addition,* $\bar{M}_n \sim M_n$ *as* $n \to \infty$.

Proof. We only prove the first asymptotic equivalence. A similar but simpler argument shows that $M_n R_n \sim \bar{M}_n$, and for bc = per one has $M_n R_n(f) = M_n(f)$ when $f \in \mathcal{L}$ and n is so large that f only depends on ω_n.

Eqs. (4.3) and (6.4) together with Proposition 6.1 imply that

$$\limsup_{n\to\infty} I(\widetilde{M}_n) < \infty \,, \tag{6.5}$$

where again $I = I_1$ is the mean relative entropy with respect to the Poisson point random field $Q = Q^1$. Lemma I.5.7 thus yields that $\widetilde{M}_n \sim M_n R_n$ as $n \to \infty$, which is the case $k = 0$. For general k, the argument in the proof of this lemma shows that for each $f \in \mathcal{L}$ and arbitrary $a > 0$

$$|v_{n+k}^{-1} v_n \widetilde{M}_n(f) - M_n R_{n+k}(f)| \leq o(1) + c\, \widetilde{M}_{n,k}(\tilde{N}_\ell; \tilde{N}_\ell \geq a) \ . \tag{6.6}$$

Here ℓ and c are such that f depends only on ω_ℓ and (2.11) holds (where $\tilde{N}_\ell$ was defined), and $\widetilde{M}_{n,k} \in \mathcal{P}_\Theta$ is the average of the measures $M_{n,k}^{\text{per}} \circ \vartheta_x^{-1}$ with $x \in \Lambda_{n+k}$, where $M_{n,k}^{\text{per}}$ is the measure relative to which the configurations in the disjoint blocks $\Lambda_{n+k} + (2n + 2k + 1)i, i \in \mathsf{Z}^d$, are i.i.d. with distribution M_n. (In particular, $\widetilde{M}_{n,0} = \widetilde{M}_n$, and for $k \geq 1$ there are $M_{n,k}^{\text{per}}$-almost surely no particles in the corridors at the boundaries of the blocks.)

To complete the proof we therefore need to show that a can be chosen in such a way that the last term in (6.6) becomes arbitrarily small, uniformly in n. By Lemma I.5.2, this would follow if we knew that (6.5) holds for the sequence $(\widetilde{M}_{n,k})_{n\geq 0}$ in the place of $(\widetilde{M}_n)_{n\geq 0}$. But this is indeed the case because

$$dQ_n \,/\, dQ_{n+k} = 1_{\{N_{n+k} - N_k = 0\}} \exp[(v_{n+k} - v_n)c(1)]$$

and therefore

$$v_n I(\widetilde{M}_{n,k}) \leq I(M_n; Q_{n+k}) = I(M_n; Q_n) + (v_{n+k} - v_n)c(1) \ .$$

Assertion (6.5) for the sequence $(\widetilde{M}_{n,k})_{n\geq 0}$ thus follows as in the case $k = 0$. □

We are now prepared for the first step in the proof of Theorem 3.3.

Proof of Theorem 3.3 under condition (6.2). By (6.4) and Lemmas 6.1 and 4.2, the sequence $(\widetilde{M}_n)_{n\geq 0}$ is sequentially compact, and every accumulation point P satisfies $S(P) \geq s(\rho, \varepsilon)$. Since $\rho(\widetilde{M}_n) = N/v_n \to \rho$ and $\rho(\cdot)$ is continuous, we have $\rho(P) = \rho$.

To show that $U(P) \leq \varepsilon$ we use Lemma 6.2. For each k, P is also an accumulation point of $(M_n R_{n+k})_{n\geq 0}$, and

$$U(M_n R_{n+k}) = M_n(U(R_{n+k})) = v_{n+k}^{-1} M_n(H_{n+k,\text{per}})$$

because of (5.2). By Lemma 5.2, for each $\delta > 0$ we can find some k such that $M_n(H_{n+k,\text{per}}) \leq (\varepsilon + \delta)v_n$ for sufficiently large n. Since U is lower semicontinuous, it follows that $U(P) \le \varepsilon$. In fact, $U(P) = \varepsilon$ because otherwise $S(P) < s(\rho, \varepsilon)$ by the strict monotonicity of $s(\rho, \cdot)$. Hence $P \in \mathcal{M}_{\rho,\varepsilon}$.

By Lemma 6.2, the above properties of the sequence $(\widetilde{M}_n)$ carry over to the asymptotically equivalent sequence $(\bar{M}_n)$ and, in the case bc = per, to (M_n). This establishes Theorem 3.3 under assumption (6.2). □

We now turn to the case when (6.2) fails, which means that u tends to zero exponentially with v_n. Since this includes the case $u = 0$, we first need a proper definition of the microcanonical distributions $M^0_{n|N,E,\mathrm{bc}}$ on the energy surfaces $\{H_{n,\mathrm{bc}} = E\}$. As we have said in Section 2.3, it is sufficient to have a definition of $M^0_{n|N,E,\mathrm{bc}}(f)$ when $f \in \mathcal{L}$ is fixed and n is sufficiently large. Such a definition can be obtained in analogy to (5.6). We still assume (6.1).

Lemma 6.3 *For each $m \ge 1$ there exists a number $n(m)$ such that, for all $n \ge n(m)$ and $0 < u \le 1$, $M^u_{n|N,E,\mathrm{bc}}$ is well-defined and the following holds: For any $f \in \mathcal{L}$ which only depends on the particle configuration in the union $\Lambda = \Lambda(f)$ of at most m unit cells, the limit*

$$M^0_{n|N,E,\mathrm{bc}}(f) = \lim_{u \to 0} M^u_{n|N,E,\mathrm{bc}}(f)$$

exists and equals $Z'_{n|N,E,\mathrm{bc}}(f)/Z'_{n|N,E,\mathrm{bc}}$, where

$$\begin{aligned} Z'_{n|N,E,\mathrm{bc}}(f) &= \sum_{k \le N-4/d} \int L_{\Lambda \cap \Lambda_n|k}(d\omega)\, f(\omega)\, c_{N-k}\, (N-k)(d/2) \\ &\times \int \hat{L}_{\Lambda_n \setminus \Lambda|N-k}(d\hat{\eta}) \Big(E - H^{\mathrm{kin}}(\omega) - H^{\mathrm{pot}}_{n,\mathrm{bc}}(\hat{\omega} \cup \hat{\eta})\Big)_+^{(N-k)d/2-1}. \end{aligned} \tag{6.7}$$

In the above, the Lebesgue measures $L_{\Lambda \cap \Lambda_n|k}$ and $\hat{L}_{\Lambda_n \setminus \Lambda|N-k}$ are defined in obvious analogy to (2.12) and (5.5), and c_{N-k} is as in (5.6).

Proof. Note first that, by Remark 5.8, $Z^u_{n|N,E,\mathrm{bc}} \ge u\, Z'_{n|N,E-u,\mathrm{bc}} > 0$ when $0 < u \le 1$ and n is sufficiently large. This means that for large n the measures $M^u_{n|N,E,\mathrm{bc}}$ are well-defined.

It follows straight from the definitions that

$$\int_{E-u}^{E} Z'_{n|N,s,\mathrm{bc}}(f)\, ds = L_{n|N}\,(f; E-u < H_{n,\mathrm{bc}} \le E, N_\Lambda \le N - 4/d)$$

for all $u > 0$, and the condition $k \le N - 4/d$ in (6.7) ensures that $Z'_{n|N,s,\mathrm{bc}}(f)$ is a continuous function of s. The lemma will therefore be proved once we have shown that

$$\{H_{n,\mathrm{bc}} \le E\} \cap \Omega_{n|N} \subset \{N_\Lambda \le N - 4/d\}$$

for n large enough. So let $\omega \in \Omega_{n|N}$ and suppose that $N_\Lambda(\omega) > N - 4/d \equiv \underline{N}$.Then $T_n(\omega) \ge \underline{N}^2/m$ and therefore $H_n(\omega) \ge a\,\underline{N}^2/m - b\,N$. By an estimate similar to Lemma II.4.2, we find a constant $c > 0$ such that

$$H_{n,\mathrm{bc}}(\omega) - H_n(\omega) \;\ge\; \begin{cases} -cNt & \text{if bc } \in \Omega(t) \\ -cN^2/v_n & \text{if bc } = \text{per}. \end{cases}$$

Combining these estimates with (6.1) we see that $H_{n,\mathrm{bc}}(\omega) > E$ when n is large enough, and the lemma follows. □

Our strategy for the extension of Theorem 3.3 to the case when (6.2) fails is a comparison of $M^u_{n|N,E,\mathrm{bc}}$ with $M^s_{n|N,E,\mathrm{bc}}$ for some $s \geq u$ satisfying $v_n^{-1} \log_- s \to 0$. This comparison is based on the monotonicity properties of $Z'_{n|N,E,\mathrm{bc}}(f)$ and the next lemma.

Lemma 6.4 *Suppose that $u \to 0$ as $n \to \infty$ and (6.1) holds. Then*

$$Z'_{n|N,E+u,\mathrm{bc}} \;/\; Z'_{n|N,E,\mathrm{bc}} \to 1 \;\text{ as }\; n \to \infty.$$

Proof. The ratio under consideration is certainly larger than 1. To obtain an upper estimate we consider the function

$$\sigma_n(s) = \log Z'_{n|N,s,\mathrm{bc}}$$

for values of n and s for which $Z'_{n|N,s,\mathrm{bc}} > 0$. Computing the derivatives as in Lemma 5.7 and using the Cauchy–Schwarz inequality we obtain the estimate

$$(1/\sigma'_n)' \leq 2/(Nd-2) \equiv a_N \;;$$

cf. Ref. 21. Hence

$$1/\sigma'_n(s) - 1/\sigma'_n(t) \leq a_N(s-t) \tag{6.8}$$

when $s \geq t$, and the mean value theorem gives

$$\delta v_n \Big/ \Big(\sigma_n(E+\delta v_n) - \sigma_n(E)\Big) - 1/\sigma'_n(E) \leq \delta v_n a_N$$

for each $\delta > 0$. Letting $n \to \infty$ and $\delta \to 0$ we thus obtain from Remark 5.8

$$\limsup \sigma'_n(E) \leq \partial_+ s(\rho, \varepsilon)/\partial\varepsilon \;. \tag{6.9}$$

The expression on the right side is the right derivative with respect to ε (which exists by concavity). Averaging over t in (6.8) we obtain

$$1/\sigma'_n(E+u) - u \;/\; (\sigma_n(E+u) - \sigma_n(E)) \leq u\, a_N \,. \tag{6.10}$$

Together with (6.9) (for $E+u$ in place of E) this gives

$$\limsup(\sigma_n(E+u) - \sigma_n(E)) \leq \limsup u\, \sigma'_n(E+u) = 0$$

which is the desired estimate. □

Remark 6.5 (6.10) also gives a counterpart to (6.9) which, together with the concavity of $s(\rho,\cdot)$, implies that $s(\rho,\cdot)$ is differentiable and

$$\partial s(\rho,\varepsilon)/\partial\varepsilon = \lim \frac{d}{dE} \log Z'_{n|N,E,\mathrm{bc}} \;.$$

The same argument also shows that the last identity holds without the prime on the righthand side. □

The following proposition asserts that microcanonical distributions on thin and on slightly thicker energy shells are asymptotically equivalent.

Proposition 6.6 *Suppose (6.1) holds and $u, s \in [0,\infty]$ vary with n in such a way that $0 \le u \le s \to 0$ and $v_n^{-1} \log s \to 0$. Then*

$$\bar{M}^u_{n|N,E,\mathrm{bc}} \sim \bar{M}^s_{n|N,E,\mathrm{bc}} \text{ and } M^u_{n|N,E,\mathrm{per}} \sim M^s_{n|N,E,\mathrm{per}} \text{ as } n \to \infty .$$

Proof. We drop all indices which are not necessary. We need to show that, for given $f \in \mathcal{L}$, $\bar{M}^u_{n|\mathrm{bc}}(f) - \bar{M}^s_{n|\mathrm{bc}}(f) \to 0$ and $M^u_{n|\mathrm{per}}(f) - M^s_{n|\mathrm{per}}(f) \to 0$ as $n \to \infty$. We can assume without loss that $f \ge 0$. Let ℓ and c be such that $f(\omega) = f(\omega_\ell)$ and (2.11) holds, $n(v_{\ell+1})$ be chosen according to Lemma 6.3, and $n \ge n(v_{\ell+1})$. We define

$$a_n(u,f) \equiv a_{n|E}(u,f) = \int_0^1 Z'_{n|E-ru}(f)\, dr$$

and $a_n(u) = a_n(u,1)$ for the constant function 1. Then $M^u_n(f) = a_n(u,f)/a_n(u)$. (Note that $a_n(u) \ge Z'_{n|E-u} > 0$ for large n, by Remark 5.8.) It follows from (6.7) that $Z'_{n|E}(f)$ is increasing in E and its increments are increasing in f. Hence

$$\begin{aligned} 0 \;&\le\; \Big(a_n(u,g) - a_n(s,g)\Big) \Big/ a_n(u) \\ &\le\; \Big(a_n(0,h) - a_n(s,h)\Big) \Big/ a_n(s) \end{aligned}$$

whenever $g \le h$. Let $a > 0$ be a constant to be chosen later. We apply the preceding inequality to the two functions $g_1 = f\, 1_{\{\tilde{N}_\ell \le a\}}$ and $g_2 = f\, 1_{\{\tilde{N}_\ell > a\}}$ and the corresponding h_1 and h_2 that are obtained by estimating f according to (2.11). After a spatial averaging we arrive at the estimate

$$\begin{aligned} 0 \;&\le\; \bar{M}^u_n(f) - r_n \bar{M}^s_n(f) \\ &\le\; c(1+a)(q_n - 1) + c\,\tilde{q}_n\, \bar{M}^s_{n|E+s}(\tilde{N}_\ell; \tilde{N}_\ell > a) . \end{aligned}$$

Here $1 \ge r_n \equiv a_n(s)/a_n(u) \ge Z'_{n|E-s}/Z'_{n|E} \to 1$ as $n \to \infty$ by Lemma 6.4, and similarly

$$1 \le q_n \equiv a_n(0)/a_n(s) \le Z'_{n|E}/Z'_{n|E-s} \to 1$$

and $\tilde{q}_n \equiv Z'_{n|E+s}/Z'_{n|E-s} \to 1$ as $n \to \infty$. By Lemma 6.2,

$$\limsup \bar{M}^s_{n|E+s}(\tilde{N}_\ell; \tilde{N}_\ell > a) \le \limsup \widetilde{M}^s_{n|E+s}(\tilde{N}_\ell; \tilde{N}_\ell > a) ,$$

and (6.5) and Lemma I.5.2 imply that the last expression tends to zero as $a \to \infty$. Combining these estimates we arrive at the desired conclusion. In the case bc = per we simply omit the spatial averaging in the preceding argument. □

Rest of the proof of Theorem 3.3. It is sufficient to consider suitable sub-subsequences of arbitrary subsequences. We can therefore assume that $\min(u,1)$ converges to a limit q. If $q > 0$ we are in the case (6.2) which was already treated. In the alternative case, we can choose a sequence (s) as in the hypothesis of Proposition 6.6, and the validity of the theorem for the sequences $(\bar{M}^s_{n|\mathrm{bc}})$ and $(M^s_{n|\mathrm{per}})$ carries over to the sequences $(\bar{M}^u_{n|\mathrm{bc}})$ and $(M^u_{n|\mathrm{per}})$. □

7 The variational principle for Gibbs measures

This final section contains the proof of Theorem 3.4. To obtain a convenient expression for the excess mean free energy we first need the grand canonical analogue of Theorem 3.2, namely the existence and variational characterization of the pressure.

Proposition 7.1 *Let $z, \beta > 0$ and $p(z,\beta)$ be defined by (3.11). Then*

$$\lim_{n\to\infty} v_n^{-1} \log \Xi_{n,z,\beta,\mathrm{bc}} = p(z,\beta)$$

for each $\mathrm{bc} \in \Omega^* \cup \{\mathrm{per}\}$, *and the convergence is uniform in* $\mathrm{bc} \in \Omega(t) \cup \{\mathrm{per}\}$ *for each* $t > 0$.

Proof. In the pure positional setting of particles without momentum, this result was established in Ref. II. The present case can be reduced to the positional one as follows. For $P \in \mathcal{P}_\Theta$ we let $\hat{P} = P(\omega : \hat{\omega} \in \cdot)$ denote the distribution of the configuration of particle positions in $\hat{\Omega}$, the set of all locally finite subsets $\hat{\omega}$ of R^d. In particular, if $Q^{z,\beta}$ is the Poisson point random field on $\mathsf{R}^d \times \mathsf{R}^d$ with intensity measure $\mu(dx, dp) = z\, dx\, e^{-\beta|p|^2} dp$ (which amounts to setting $\tau(p) = -\log z + \beta|p|^2$ in (4.1)) then $\hat{Q}^{z,\beta}$ is the Poisson point random field on R^d with particle density $c(z,\beta) = z\, c(\beta)$, and we can write

$$\Xi_{n,z,\beta,\mathrm{bc}} = \exp[v_n c(z,\beta)]\, \hat{Q}_n^{z,\beta}(\exp[-\beta\, H_{n,\mathrm{bc}}^{\mathrm{pot}}])$$

because $H_{n,\mathrm{bc}}^{\mathrm{pot}}(\omega)$ is a function of $\hat{\omega}$. Theorems II.2 and II.3 therefore imply that

$$\lim_{n\to\infty} v_n^{-1} \log \hat{Q}_n^{z,\beta}(\exp[-\beta H_{n,\mathrm{bc}}^{\mathrm{pot}}]) = -\min_{P\in\mathcal{P}_\Theta} \Big[\beta\, U^{\mathrm{pot}}(\hat{P}) + \hat{I}_{z,\beta}(\hat{P})\Big]\,, \tag{7.1}$$

where $\hat{I}_{z,\beta}(\hat{P}) = \lim_{n\to\infty} v_n^{-1} I(\hat{P}_n; \hat{Q}_n^{z,\beta})$ is defined in analogy to Lemma 4.1. On the other hand, (3.10) and (4.2) show that

$$p(z,\beta) = c(z,\beta) - \min_{P\in\mathcal{P}_\Theta} \Big[\beta\, U^{\mathrm{pot}}(P) + I_{z,\beta}(P)\Big]\,, \tag{7.2}$$

where $I_{z,\beta}$ is the mean relative entropy with respect to $Q^{z,\beta}$. We thus only need to show that the minima on the right-hand sides of (7.1) and (7.2) coincide. This will follow once we have shown that, for each $P \in \mathcal{P}_\Theta$, $\hat{I}_{z,\beta}(\hat{P}) \le I_{z,\beta}(P)$ with equality when P is Maxwellian, in that

$$P = \int \hat{P}(d\hat{\omega})\, \Gamma_\beta(\hat{\omega}, \cdot)\ .$$

Here $\Gamma_\beta(\hat{\omega}, \cdot)$ is the conditional distribution of $Q^{z,\beta}$ given $\hat{\omega}$, i.e., the distribution of $\{(x, Y_x) : x \in \hat{\omega}\}$, where the Y_x are i.i.d. centered normal with variance $1/2\beta$.

To verify this inequality we can assume that $I_{z,\beta}(P) < \infty$. Then, for all n, $P_n \ll Q_n^{z,\beta}$ with a density f_n, whence $\hat{P}_n \ll \hat{Q}_n^{z,\beta}$ with density $\hat{f}_n(\hat{\omega}) = \Gamma_\beta(\hat{\omega}, f_n)$. Thus, by Jensen's inequality,

$$\begin{aligned} I(\hat{P}_n; \hat{Q}_n^{z,\beta}) &= \hat{Q}_n^{z,\beta}(\hat{f}_n \log \hat{f}_n) \\ &\le \int \hat{Q}_n^{z,\beta}(d\hat{\omega})\, \Gamma_\beta(\hat{\omega}, f_n \log f_n) \ = \ I(P_n; Q_n^{z,\beta}) \end{aligned}$$

with equality when $f_n(\omega)$ only depends on $\hat{\omega}$. This gives the desired relationship between $\hat{I}_{z,\beta}$ and $I_{z,\beta}$. $\square$

The preceding proposition shows that the excess mean free energy can be characterized as a mean relative entropy relative to the grand canonical Gibbs distributions with free boundary condition.

Corollary 7.2 *For all $z, \beta > 0$ and $P \in \mathcal{P}_\Theta$,*

$$\delta F_{z,\beta}(P) \equiv F_{z,\beta}(P) + p(z,\beta) = \lim_{n\to\infty} v_n^{-1} I(P_n; G_{n,z,\beta,\text{free}}) \ .$$

Proof. Since $G_{n,z,\beta,\text{free}}$ is equivalent to L_n with the density appearing in (2.17), we have

$$I(P_n; G_{n,z,\beta,\text{free}}) = -S_n(P) - P(N_n)\log z + \beta\, P(H_n) + \log \Xi_{n,z,\beta,\text{free}}$$

for all n, so that the corollary follows from Proposition 3.1, Eqs. (3.1) and (3.10), and Proposition 7.1. $\square$

We now turn to the proof of Theorem 3.4. We can assume without loss that $z = \beta = 1$ (because this only amounts to a rescaling of the position and momentum spaces together with a rescaling of φ), and we shall drop all indices referring to these parameters. In particular, we write G_n resp. $G_{n,\zeta}$ for the grand canonical Gibbs distribution in Λ_n with free boundary condition resp. bc $= \zeta$. We shall also need the Gibbs distributions in more general sets Λ in place of Λ_n, which will be denoted by G_Λ resp. $G_{\Lambda,\zeta}$. The sets Λ and Δ considered below will always be finite unions of the unit cells $C + i, i \in \mathsf{Z}^d$.

Our first aim is the proof of that part of the variational principle which is essential for the equivalence of ensembles, namely that all minimizers of the free energy are Gibbsian. For this we shall need an estimate of the expected variational distance

$$\int P(d\zeta)\, \|G_{0,\zeta} - G_{0,\zeta_k}\|$$

(relative to certain measures P) between the Gibbs distributions in $\Lambda_0 = C$ with boundary conditions ζ and $\zeta_k = \zeta \cap (\Lambda_k \times \mathsf{R}^d)$. Clearly, if φ has finite range then this distance vanishes as soon as k exceeds the range of φ. The general case will be treated in Lemma 7.4; the next lemma serves as preparation.

For each $i \in \mathsf{Z}^d$ we write $\psi_i = \psi(d(C, C+i))$, where ψ is as in assumption (A1) and $d(C, C+i)$ is the distance of the cells C and $C+i$. We also set $\Psi = \sum_{i\in\mathsf{Z}^d} \psi_i$ and $\tilde{\psi}_i = \psi_i/2\Psi$ so that $\sum_{i\in\mathsf{Z}^d} \tilde{\psi}_i = 1/2$. Note that $\tilde{\psi}_0 = 0$.

Lemma 7.3 *For every exponent $\gamma \ge 1$ there exists a constant $a_\gamma < \infty$ such that*

$$G_{0,\zeta}(N_0^\gamma) \le a_\gamma + \sum_{i\in\mathsf{Z}^d} \tilde{\psi}_i\, N_{C+i}(\zeta)^\gamma$$

for all $\zeta \in \Omega^*$.

Proof. This is a version of Lemma 2 of Dobrushin[25]. For completeness we indicate the proof. Let $b_\zeta = \sum_{i\in \mathsf{Z}^d} \tilde{\psi}_i\, N_{C+i}(\zeta)$. It is easy to see that $b_\zeta < \infty$ for $\zeta \in \Omega^*$ (cf. Lemma II.4.2), and we only need to show that

$$G_{0,\zeta}(N_0^\gamma; N_0 > b_\zeta) \le a_\gamma\ . \tag{7.3}$$

In view of (5.3) and (2.1),

$$H_{0,\zeta}^{\text{pot}} \ge h(N_0) - b'\, N_0 - 2\Psi\, N_0\, b_\zeta\ .$$

On the other hand, $\Xi_{0,\zeta} \ge 1$. The left-hand side of (7.3) is therefore at most

$$\sum_{\ell > b_\zeta} \ell^\gamma \exp\Big[-h(\ell) + b'\ell + 2\Psi\ell^2\Big]\, c(1)^\ell\ /\ \ell!\,,$$

and the corresponding sum over all $\ell \ge 0$ is finite because $h(\ell)/\ell^2 \to \infty$ as $\ell \to \infty$. □

Lemma 7.4 *Let $P \in \mathcal{P}_\Theta$ be such that $t_P \equiv P(N_C^2) < \infty$. Then for any $\delta > 0$ we can find a number k_0 such that for all $k \ge k_0$*

$$\int P(d\zeta)\, \|G_{0,\zeta} - G_{0,\zeta_k}\| < \delta$$

and, for each Λ, the same inequality holds when P is replaced by

$$PG_\Lambda \equiv \int P(d\eta)\, G_{\Lambda,\eta}(\zeta : \zeta \cup (\eta \setminus \Lambda \times \mathsf{R}^d) \in \cdot)\ .$$

Proof. A standard estimate (see Ref. 24, p. 33, proof of (b)) and a symmetry argument yield for each $\zeta \in \Omega^*$

$$\|G_{0,\zeta} - G_{0,\zeta_k}\| \le 2\, L_0\Big(|\exp[-H_{0,\zeta}] - \exp[-H_{0,\zeta_k}]\Big)\ \Big/\ \max(\Xi_{0,\zeta}, \Xi_{0.\zeta_k}).$$

Decomposing the integral above into parts according to the sign of $H_{0,\zeta} - H_{0,\zeta_k}$ and using (2.1) we obtain the upper bound

$$2\,[G_{0,\zeta}(N_0) + G_{0,\zeta_k}(N_0)] \sum_{i \in \mathsf{Z}^d \setminus \Lambda_k} \psi_i N_{C+i}(\zeta)$$

when $k \ge r(\varphi)$. Applying Lemma 7.3 for $\gamma = 1$, integrating over ζ and using the Cauchy-Schwarz inequality we find

$$\int P(d\zeta)\, \|G_{0,\zeta} - G_{0,\zeta_k}\| \le (4a_1 t_P^{1/2} + 2t_P)\delta_k\,,$$

where $\delta_k \equiv \sum_{i\in \mathsf{Z}^d \setminus \Lambda_k} \psi_i$ tends to zero as $k \to \infty$. This proves the first assertion.

The same kind of estimate also gives the second assertion, provided we can find a constant $t < \infty$ such that $PG_\Lambda(N_{C+i}^2) \le t$ for all Λ and all $i \in \mathsf{Z}^d$. To show this we use again an idea of Dobrushin[25]. Since $PG_\Lambda(N_{C+i}^2) = t_P$ when $i \in \mathsf{Z}^d \setminus \Lambda$ we only need to consider the case when $i \in \mathsf{Z}^d \cap \Lambda$. We define

$$t_\Lambda = \max_{i \in \mathsf{Z}^d \cap \Lambda} PG_\Lambda(N_{C+i}^2)\ .$$

In view of the consistency of the Gibbs distributions and by Lemma 7.3 we have for each $i \in \mathsf{Z}^d \cap \Lambda$

$$\begin{aligned} PG_\Lambda(N_{C+i}^2) &= \int PG_\Lambda(d\zeta)\, G_{C+i,\zeta}(N_{C+i}^2) \\ &\leq a_2 + \sum_{j\in\mathsf{Z}^d} \tilde{\psi}_{i-j}\, PG_\Lambda(N_{C+j}^2) \end{aligned}$$

and therefore $t_\Lambda \leq a_2 + \frac{1}{2}\max(t_\Lambda, t_P)$. On the other hand, $t_\Lambda < \infty$ because

$$t_\Lambda \leq \int P(d\zeta)\, G_{\Lambda,\zeta}(N_\Lambda^2) \leq a' + \sum_{i\in\mathsf{Z}^d\setminus\Lambda} \psi(d(C+i,\Lambda))\, P(N_C^2)$$

for some $a' = a'(\Lambda) < \infty$ by an obvious variant of Lemma 7.3. It follows that $t_\Lambda \leq t \equiv \max(t_P, 2a_2)$, and the proof is complete. $\Box$

After these preparations we can enter into the proof that the minimizers of the mean free energy are tempered Gibbs measures. We follow the well-known idea of Preston[23,24]. Let $P \in \mathcal{P}_\Theta$ be such that $\delta F(P) = 0$. Then $U^{\rm pot}(P) < \infty$ and thus $t_P = P(N_C^2) < \infty$. Hence Lemma 7.4 is applicable. Moreover, the ergodic theorem (for the multidimensional discrete translation group $(\vartheta_i)_{i\in\mathsf{Z}^d}$) shows that $P(\Omega^*) = 1$, i.e., P is tempered. On the other hand, we conclude from Corollary 7.3 that $P_n \ll G_n$ when n is large enough. By translation invariance this means that, for all sufficiently large cubes Λ, P_Λ is absolutely continuous with respect to G_Λ with a density g_Λ. Here P_Λ is the restriction of P to the events in Λ, i.e., the image of P under the mapping $\omega \to \omega_\Lambda = \omega \cap (\Lambda \times \mathsf{R}^d)$, and G_Λ is the Gibbs distribution in Λ with free boundary condition. In particular, for any $\Delta \subset \Lambda$ we consider the restriction $G_{\Lambda,\Delta} = (G_\Lambda)_\Delta$ of G_Λ to the events in Δ, and we have that $P_\Delta \ll G_{\Lambda,\Delta}$ with density

$$g_{\Lambda,\Delta}(\zeta) = \int G_{\Lambda\setminus\Delta,\zeta_\Delta}(d\omega)\, g_\Lambda(\omega \cup \zeta_\Delta)\ .$$

The crucial consequence of the assumption $\delta F(P) = 0$ is the following.

Lemma 7.5 *Let $P \in \mathcal{P}_\Theta$ be such that $\delta F(P) = 0$, and let $k \geq 1$ and $\delta > 0$ be given. Then there exist two sets $\Delta, \Lambda \subset \mathsf{R}^d$ such that $\Lambda_k \subset \Delta \subset \Lambda$ and*

$$G_\Lambda\Big(|g_{\Lambda,\Delta} - g_{\Lambda,\Delta\setminus C}|\Big) < \delta\ .$$

Proof. In view of the argument in Ref. 24, p. 324, Step 2, it is sufficient to show that for any $\delta > 0$ we can find Δ and Λ such that $\Lambda_k \subset \Delta \subset \Lambda$ and

$$I_\Delta(P; G_\Lambda) - I_{\Delta\setminus C}(P; G_\Lambda) < \delta\ . \tag{7.4}$$

Here we write $I_\Delta(\cdot\,;\cdot)$ for the relative entropy of measures that are restricted to the events in Δ. Using Corollary 7.2 we see that there exists some $n \geq k$ such that $v_n^{-1} I(P_n; G_n) \leq \delta/2^d v_k$. As in Ref. 24, p. 324, Step 1, we choose an m with $m^d v_k \leq$

$v_n \le (2m)^d v_k$ and disjoint translates $\Lambda_k(j) = \Lambda + i(j) \subset \Lambda_n, i(j) \in \mathsf{Z}^d, 1 \le j \le m^d$. Writing $\Delta(j) = \Lambda_k(1) \cup \ldots \cup \Lambda_k(j)$ we find

$$\begin{aligned} m^{-d} \sum_{j=1}^{m^d} \Big[I_{\Delta(j)}(P;G_n) - I_{\Delta(j)\setminus(C+i(j))}(P;G_n) \Big] \\ \le \; m^{-d} I_{\Delta(m^d)}(P;G_n) \; < \; \delta \; . \end{aligned}$$

So for at least one j the corresponding term in the sum above is less than δ, and by translation invariance we obtain (7.4) for $\Lambda = \Lambda_n - i(j)$ and $\Delta = \Delta(j) - i(j)$. □

We are now ready for the first direction of the variational principle.

Proposition 7.6 *Each $P \in \mathcal{P}_\Theta$ with $\delta F(P) = 0$ is a tempered Gibbs measure.*

Proof. We saw already that P is tempered and $t_P < \infty$. We need to establish Eq. (2.18) for any measurable $f \ge 0$. We can clearly assume that $f \in \mathcal{L}$ and $0 \le f \le 1$. For given $\delta > 0$ we let k be so large that f only depends on the particles in Λ_k and the conclusions of Lemma 7.4 hold, and we determine Δ and Λ according to Lemma 7.5.

Consider the term on the left-hand side of Eq. (2.18) for $n = 0$. We replace its inner integral

$$G_0 f(\zeta) \equiv \int G_{0,\zeta}(d\omega)\, f(\omega_0 \cup \zeta \setminus \zeta_0)$$

by $\tilde{f}_k(\zeta) \equiv G_0 f(\zeta_k)$. Since f only depends on the particles in Λ_k, this replacement leads to an error of at most $\int P(d\zeta)\, \|G_{0,\zeta} - G_{0,\zeta_k}\| < \delta$. Since $\tilde{f}_k$ only depends on the particles in $\Lambda_k \setminus C \subset \Delta \setminus C$ and $P \ll G_\Lambda$ on the events which occur in this set,

$$P(\tilde{f}_k) = G_\Lambda(g_{\Lambda,\Delta\setminus C} \tilde{f}_k) \; .$$

Changing $\tilde{f}_k$ in the last term back into $G_0 f$ gives an error of at most

$$\begin{aligned} \int G_\Lambda(d\zeta)\, g_{\Lambda,\Delta\setminus C}(\zeta) \|G_{0,\zeta} - G_{0,\zeta_k}\| \\ = \; \int P G_{\Lambda\setminus(\Delta\setminus C)}(d\zeta) \|G_{0,\zeta} - G_{0,\zeta_k}\| \; < \; \delta \; . \end{aligned}$$

(For the equality above we first use the fact that $G_{\Lambda\setminus(\Delta\setminus C),\omega_{\Delta\setminus C}}$ is the conditional distribution of G_Λ given the configuration $\omega_{\Delta\setminus C}$ in $\Delta \setminus C$, and then that P has density $g_{\Lambda,\Delta\setminus C}$ relative to G_Λ on events occurring in $\Delta \setminus C$.) As a next step we observe that

$$G_\Lambda(g_{\Lambda,\Delta\setminus C} G_0 f) = G_\Lambda(g_{\Lambda,\Delta\setminus C} f)$$

because $g_{\Lambda,\Delta\setminus C}$ does not depend on the particles in $\Lambda_0 = C$ and $G_{0,\zeta}$ is the suitable conditional distribution of G_Λ. According to the choice of Δ and Λ we may finally replace $g_{\Lambda,\Delta\setminus C}$ by $g_{\Lambda,\Delta}$ with an error less than δ. In this way we arrive at the term $G_\Lambda(g_{\Lambda,\Delta} f) = P(f)$. We have thus shown that

$$|P(G_0 f) - P(f)| < 3\delta \; ,$$

and this gives (2.18) for $n = 0$. To obtain (2.18) for general n we can either apply Theorem (1.33) of Ref. 24, or simply repeat the preceding argument on a new spatial scale. □

We finally come to the converse part of the variational principle.

Proposition 7.7 *δF vanishes on the set of tempered Gibbs measures.*

Proof. We might deduce this result from the large deviation principle for tempered Gibbs measures in Ref. II by exploiting the fact that a good rate function in a large deviation principle vanishes at the underlying measure whenever this ergodic. But we prefer a more direct argument which requires only the somewhat weaker Proposition 7.1.

Let P be a tempered Gibbs measure. We need that $t_P = P(N_C^2) < \infty$. This follows either from subtle direct analysis, namely the superstability estimates of Ruelle[20]. Or from general theory by noting that (by the ergodic decomposition of translation invariant tempered Gibbs measures[24] and the fact that δF is measure affine) P can be assumed to be ergodic, in which case $t_P < \infty$ follows from the temperedness by means of the ergodic theorem.

We need to show that the limit in Corollary 7.2 vanishes. We fix any $n \geq 0$. By (2.17) and (2.18), $P_n \ll G_n$ with density

$$\begin{aligned} g_n(\omega) &= \int P(d\zeta)\, dG_{n,\zeta}/dG_n(\omega) \\ &= \int P(d\zeta)\, \exp[H_n(\omega) - H_{n,\zeta}(\omega)]\; \Xi_n \,/\, \Xi_{n,\zeta}\ . \end{aligned}$$

Therefore we can write, using Jensen's inequality, (2.18), and the lower regularity of φ,

$$\begin{aligned} I(P_n; G_n) &= G_n(g_n \log g_n) \\ &\leq G_n\Big(\int P(d\zeta)\, dG_{n,\zeta}/dG_n \log dG_{n,\zeta}/dG_n\Big) \\ &= \int P(d\omega)\, [H_n(\omega) - H_{n,\omega}(\omega) + \log \Xi_n - \log \Xi_{n,\omega}] \\ &\leq t_P \sum_{i \in \mathsf{Z}^d \cap \Lambda_n, j \in \mathsf{Z}^d \setminus \Lambda_n} \psi_{i-j} + \log \Xi_n - \int P(d\zeta)\, \log \Xi_{n,\zeta}\ . \end{aligned}$$

After division by v_n, the first of the last three terms converges to 0, and Proposition 7.1 ensures that the second term converges to $p(1,1)$. On the other hand we have $\Xi_{n,\zeta} \geq 1$, so that the same proposition together with Fatou's lemma and the temperedness of P implies that

$$\liminf_{n\to\infty} v_n^{-1} \int P(d\zeta) \log \Xi_{n,\zeta} \geq p(1,1)\ .$$

It follows that $\delta F(P) = 0$, and the proof is complete. □

References

1. 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).
2. 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).
3. D. Ruelle, *Statistical Mechanics. Rigorous results* (Benjamin, New York, 1969).

4. O.E. Lanford, Entropy and equilibrium states in classical statistical mechanics, in *Statistical Mechanics and Mathematical Problems*, A. Lenard, ed., Lect. Notes Phys. 20 (Springer, Berlin, 1973).
5. A. Martin-Löf, *Statistical Mechanics and the Foundations of Thermodynamics*, Lect. Notes Phys. 101 (Springer, Berlin, 1979).
6. A.E. Khinchin, *Mathematical foundations of statistical mechanics* (Dover, 1949).
7. A.M. Halfina, The limiting equivalence of the canonical and grand canonical ensembles (low density case), *Math. USSR Sbornik* **9**:1–52 (1969).
8. R.A. Minlos and A. Haitov, Limiting equivalence of thermodynamic ensembles in case of one-dimensional systems, *Trans. Moscow Math. Soc.* **32**:143–180 (1975).
9. R.L. Dobrushin and B. Tirozzi, The central limit theorem and the problem of equivalence of ensembles, *Commun. Math. Phys.* **54**:173–192 (1977).
10. M. Campanino, G. Del Grosso, and B. Tirozzi, Local limit theorem for Gibbs random fields of particles and unbounded spins, *J. Math. Phys.* **20**:1752-1758 (1979).
11. A. Martin–Löf, The equivalence of ensembles and the Gibbs phase rule for classical lattice systems, *J. Statist. Phys.* **20**:557-569 (1979).
12. R.L. Thompson, *Equilibrium states on thin energy shells*, Memoirs Amer. Math. Soc. 150 (1974).
13. H.O. Georgii, *Canonical Gibbs measures*, Lect. Notes Math. 760 (Springer, Berlin, 1979)
14. M. Aizenman, S. Goldstein, and J.L. Lebowitz, Conditional equilibrium and the equivalence of microcanonical and and grandcanonical ensembles in the thermodynamic limit, *Commun. Math. Phys.* **62**:279–302 (1978).
15. C. Preston, Canonical and microcanonical Gibbs states, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **46**:125–158 (1979).
16. M. Pirlot, Generalized canonical states, *Ann. Sci. Univ. Clermont-Ferrand II. Probab. Appl.* **4**:69–91 (1985).
17. P. Vanheuverzwijn, Discrete lattice systems and the equivalence of microcanonical, canonical and grand canonical Gibbs states, *Commun. Math. Phys.* **101**:153–172 (1985).
18. K. Matthes, J. Kerstan, and J. Mecke, *Infinitely Divisible Point Processes* (Wiley, Chichester, 1978).
19. T. Tjur, *Probability based on Radon measures* (Wiley, Chichester, 1980).
20. D. Ruelle, Superstable interactions in classical statistical mechanics, *Commun. Math. Phys.* **18**:127–159 (1970).
21. R. Rechtmann and O. Penrose, Continuity of the temperature and derivation of the Gibbs canonical distribution in classical statistical mechanics, *J. Statist. Phys.* **19**:359–366 (1978).
22. O.E. Lanford and D. Ruelle, Observables at infinity and states with short range correlations in statistical mechanics, *Commun. Math. Phys.* **9**:327–338 (1968).
23. C. Preston, *Random fields*, Lect. Notes Math. 534 (Springer, Berlin, 1976)
24. H.O. Georgii, *Gibbs Measures and Phase Transitions* (de Gruyter, Berlin, 1988).
25. R.L. Dobrushin, Gibbsian random fields for particles without hard core, *Theor. Math. Phys.* **4**:705–719 (1970).
26. H.O. Georgii, Large deviations and maximum entropy principle for interacting random fields on $\mathbb{Z}^d$, *Ann. Probab.* **21**:1845–1875 (1993).