Ensemble Equivalence for Mean Field Models and Plurisubharmonicity

Citation for the original published paper (version of record):
Berman, R. (2022). Ensemble Equivalence for Mean Field Models and Plurisubharmonicity. Archive for Rational Mechanics and Analysis, 246(2-3): 771-822.
http://dx.doi.org/10.1007/s00205-022-01829-6

(article starts on next page)

Arch. Rational Mech. Anal. 246 (2022) 771–822
Digital Object Identifier (DOI) https://doi.org/10.1007/s00205-022-01829-6

Ensemble Equivalence for Mean Field Models and Plurisubharmonicity

Robert J. Berman

Communicated by P.-L. Lions

Abstract

We show that entropy is globally concave with respect to energy for a rich class of mean field interactions, including regularizations of the point vortex model in the plane, plasmas and self-gravitating matter in 2D, as well as the higher-dimensional logarithmic interactions appearing in conformal geometry and power laws. The proofs are based on a corresponding "microscopic" concavity result at finite N, shown by leveraging an unexpected link to Kähler geometry and plurisubharmonic functions. Under more restrictive homogeneity assumptions, strict concavity is obtained using a uniqueness result for free energy minimizers, established in a companion paper. The results imply that thermodynamic equivalence of ensembles holds for this class of mean field models. As an application, it is shown that the critical inverse negative temperatures—in the macroscopic as well as the microscopic setting—coincide with the asymptotic slope of the corresponding microcanonical entropies. Along the way, we also extend previous results on the thermodynamic equivalence of ensembles for continuous weakly positive definite interactions, concerning positive temperature states, to the general non-continuous case. In particular, singular situations are exhibited where, somewhat surprisingly, thermodynamic equivalence of ensembles fails at energy levels sufficiently close to the minimum energy level.

1. Introduction

1.1. General Setup

Let X be a topological space, W a symmetric lower semi-continuous (lsc) function on X^2 (the *pair interaction potential*) and V a lsc function on X (the *exterior potential*), both taking values in $]-\infty, \infty]$. The corresponding N-particle

mean field N-particle Hamiltonian is defined by

$$H^{(N)}(x_1, ..., x_N) := \frac{1}{2}\frac{1}{N}\sum_{i\neq j\leq N} W(x_i, x_j) + \sum_{i=1}^{N} V(x_i). \tag{1.1}$$

The self-interactions have, as usual, been excluded in order to render $H^{(N)}$ generically finite in the case when W is singular on the diagonal. The corresponding *(macroscopic) energy* $E(\mu)$ of a probability measure μ on X, i.e., $\mu \in \mathcal{P}(X)$, is defined by

$$E(\mu) := \frac{1}{2}\int_X W\mu \otimes \mu + \int_X V\mu \in]-\infty, \infty] \tag{1.2}$$

when X has compact support. The definition can be extended to non-compactly supported measures (see Sect. 4.2), but for most purposes it will be enough to consider the restriction of $E(\mu)$ to the space of all probability measures on X with compact support, denoted by $\mathcal{P}(X)_0$.

Now, fix also a measure μ_0 on X (the "prior"). Following [15,16,29,30], the *entropy* $S(e)$ (at energy e) is the function on $\mathbb{R}$ defined by $-\infty$ in the case that $\{E(\mu) = e\}$ is empty, and otherwise,

$$S(e) := \sup_{\mu\in\mathcal{P}(X)_0} \{S(\mu) : \ E(\mu) = e\}, \quad S(\mu) := -\int_X \log(\mu/\mu_0)\mu, \tag{1.3}$$

where $S(\mu)$ is the *entropy* of μ relative μ_0, which, by definition, is equal to $-\infty$ if μ is not absolutely continuous wrt μ_0. A measure μ^e maximizing $S(\mu)$ above is called a *maximum entropy measure*. In the case when μ_0 is a probability measure, we shall focus on,

$$\max_{e\in\mathbb{R}} S(e) = S(e_0) = 0, \quad e_0 := E(\mu_0)$$

(since $S(\mu) \leq 0$ with equality iff $\mu = \mu_0$). This setup is modeled on repulsive Hamiltonians (as in the case of identical point vortices described below), but an equivalent setup of "attractive" Hamiltonians is obtained by replacing $H^{(N)}$ with $-H^{(N)}$ and e with $-e$.

1.2. Background: Concavity of $S(e)$ and Thermodynamic Equivalence of Ensembles

In the case when $E(\mu)$ is linear on $\mathcal{P}(X)$ (i.e., $W = 0$), it follows directly from the concavity of $S(\mu)$ on $\mathcal{P}(X)$ that the entropy $S(e)$ is concave with respect to e (see Sect. 5.1). This is the standard setup in information theory and statistical inference, going back to Shannon and Jaynes [38], but here we will be concerned with the case when $E(\mu)$ is quadratic, motivated by mean field models in statistical mechanics (see 7.2 for a comparison between the classical linear setup and the quadratic setup appearing in the context of plasmas). General, nonlinear $E(\mu)$ also appear naturally in engineering optimization [4]. The concavity properties of the entropy $S(e)$ for mean field models and other systems with long-range interactions

have been studied extensively from various points of views in the last decades: theoretical as well as experimental and numerical [15,16,21,29,31]. As stressed in [16], the unusual properties of these systems stem from the lack of additivity (i.e., lack of linearity of $E(\mu)$). In particular, the question whether $S(e)$ is globally concave is crucial in connection with negative temperature states in Onsager's point vortex model for the large time limit of turbulent incompressible non-viscous 2D fluids: classical fluids [15,29], as well as quantum fluids [35,39]. In the case of vortices of equal circulation moving in the whole plane $\mathbb{R}^2$, the vortex–vortex pair interaction potential $W(x, y)$ is proportional to the Green function for the Laplacian in $\mathbb{R}^2$,

$$W(x, y) = -\log(|x - y|). \tag{1.4}$$

As emphasized in [16,30,55], the relevance of the global concavity of $S(e)$ stems from the fact that it equivalently means that $S(e)$ may (under appropriate regularity assumptions discussed in Sect. 3) be expressed as the Legendre–Fenchel transform of the *Helmholtz (scaled) free energy* $F(\beta)$ at inverse temperature β :

$$S(e) = \inf_{\beta \in \mathbb{R}} (-F(\beta) + \beta e). \tag{1.5}$$

Here $F(\beta)$ is defined as the following infimum of the (scaled) free energy functional $F_\beta(\mu)$:

$$F(\beta) = \inf_{\mu \in \mathcal{P}(X)_0} F_\beta(\mu), \quad F_\beta(\mu) := \beta E(\mu) - S(\mu) \tag{1.6}$$

(where $F(\mu)$ is defined to be equal to $+\infty$ if $S(\mu) = -\infty$). Accordingly, when $S(e)$ is globally concave *thermodynamic equivalence of ensembles* is said to hold [16,30,55] (since it amounts to the equivalence between the *microcanonical ensemble* at a fixed energy e and the *canonical ensemble* at a corresponding fixed inverse temperature β, in the large N-limit). More generally, this duality fits into primal–dual formulations of nonlinear optimization problems, where the free energy functional appears as the augmented Lagrangian [4]. If thermodynamic equivalence of ensembles holds, then $S(e)$ is differentiable at almost any energy level e and, by the concavity of $F(\beta)$, the infimum over the β in formula 1.5 is attained precisely at the inverse temperature

$$\beta(e) = \frac{dS(e)}{de},$$

Remarkably, as stressed already by Onsager in the late 40 s [48], this means, since $S(e)$ is decreasing for $e > E(\mu_0)$, that in the *"high-energy regime"*

$$E(\mu_0) < e, \tag{1.7}$$

an energy level e should correspond to a *negative* inverse temperature β. As a consequence, the repulsive vortex interaction should then become effectively attractive, resulting in the aggregation of microscopic vortices of equal circulation into large-scale coherent clusters (as observed in oceanic and atmospheric fluids, notably Jupiter's famous great red spot). A few years after Onsager's prediction the existence of negative temperature states was experimentally verified in nuclear

spin systems [50], while the original prediction was quantitatively experimentally demonstrated only very recently in a 2D quantum superfluid (a Bose Einstein condensate [35,39]). Note that the high-energy region only exists if

$$E(\mu_0) < \infty.$$

This is automatically the case if X is compact and W and V are locally integrable, but it also holds in many non-compact situations, for example, the vortex model in $\mathbb{R}^2$, when μ_0 is taken as a Gaussian probability measure (incorporating conservation of angular momentum).

As shown in [29], if $W(x, y)$ defines a weakly positive definite kernel, as in the point vortex model, then the concavity of $S(e)$ holds in the *"low-energy regime"*

$$e \leq E(\mu_0),$$

which corresponds to positive inverse temperature β (more precisely, in [29] it is assumed that W is continuous; the general case is discussed in Sect. 5.1). However, the concavity may fail in the "high-energy regime," and thus, the correspondence with negative temperature then breaks down (leading to the peculiar phenomenon of negative heat capacity [16,44]). This is illustrated by the mean field Blume–Emery–Griffiths spin model in [16, Section 4.2.4]. In the case of the point vortex model, the global concavity of $S(e)$ has been established when X is the unit disk in $\mathbb{R}^2$ [15] (or a sufficiently small deformation of the unit disk) or all of $\mathbb{R}^2$ [15,19,42], while shown to fail for some domains X (e.g., a sufficiently thin rectangle). The proofs in [15,19] exploit that in the case of the point vortex model any minimizer μ_β of the free energy functional $F_\beta(\mu)$ satisfies a second-order PDE (the Joyce-Montgomery mean field equation/the Liouville equation). This opens the door for the application of various PDE techniques (uniqueness results, concentration/compactness alternatives, symmetrization arguments, etc.).

1.3. Summary of the Main Results

To the best of the author's knowledge, there are, apart from a few special cases—such as the BEG model and the vortex model recalled above—no general global concavity results for mean field Hamiltonians. Even in the case of the regularized vortex model [29], the question of global concavity of $S(e)$ raised in [29] appears to have been left open. Similar questions have also been put forward in the context of self-gravitating matter, where regularizations appear naturally [40]. Allowing regularizations is also crucial when comparing theoretical results with numerical simulations (such as [26,53]) to ensure that the concavity of $S(e)$ is a robust feature of the models in question. The main aim of the present work is to establish the concavity of $S(e)$ for a rich class of potentials W, V and priors μ_0, including the point vortex model in $\mathbb{R}^2$, as well as its regularizations and regularized plasmas and self-gravitating systems in 2D and power laws. However, since neither explicit calculations, nor PDE techniques are available for such interactions W we take a different route. First it is shown that the *(upper) microcanonical entropy* (at energy e).

$$S_+^{(N)}(e) := \frac{1}{N} \log \mu_0^{\otimes N}\{H^{(N)}/N > e\} \tag{1.8}$$

is concave for any finite N. (In the context of the point vortex model, this microcanonical entropy appears in [15, Theorem 4.2].) Then, letting $N \to \infty$, and using the asymptotics for $S_+^{(N)}(e)$ from [29], the concavity of the upper entropy $S_+(e)$ is obtained (defined by replacing the condition $E(\mu) = e$ in the definition 1.3 of $S(e)$ with the condition $E(\mu) \geq e$). Hence, the concavity of $S(e)$ in the high-energy region, $e > E(\mu_0)$, results from the observation that $S_+(e) = S(e)$ there. This derivation of the concavity of $S(e)$ is in the spirit of statistical mechanics; the macroscopic property in question emerges from a microscopic one. The proof of the concavity of $S_+^{(N)}(e)$ leverages some developments in Kähler geometry [10,11], centered around complex analogs of the Brunn–Minkowski inequality. Under more restricted assumptions, $S(e)$ is shown to be *strictly* concave, using a different (macroscopic) approach—which is more in the spirit of [15]—based on a uniqueness result for free energy minimizers of independent interest.

Before turning to a more precise description of the main present results, it may be worth emphasizing that the concavity of $S_+^{(N)}(e)$ is considerably stronger than the concavity of $S(e)$ and does not require the mean field scaling (nor the permutation symmetry). Thus, it also applies to the microcanonical study of *small systems*, considered in the physics literature (see, for example, [28,33,52]). The relation to the setup in [28,52] becomes clearer in the equivalent setup of "attractive" Hamiltonians obtained by replacing the $H^{(N)}$ with the Hamiltonian $-H^{(N)}$ and e with $-e$. The concavity of $S_+^{(N)}(e)$ then translates into the concavity of

$$S_-^{(N)}(e) := \frac{1}{N} \log \mu_0^{\otimes N}\{H^{(N)}/N < e\}, \tag{1.9}$$

called the *bulk entropy* in [52] and the *microcanonical Gibbs entropy* in [28]. In recent years, it has been debated whether this microcanonical entropy is physically more relevant than the microcanonical Boltzmann entropy, obtained by replacing the volume $\mu_0^{\otimes N}\{H^{(N)}/N < e\}$ with its derivative with respect to e (the surface area of the level set $H^{(N)}/N = e$); see the discussion in [36] and references therein. The present results may, perhaps, be interpreted as a case for bulk/Gibbs entropy as this entropy is shown to be concave in our class of "attractive" Hamiltonians, while the Boltzmann entropy is not always concave in this class (as discussed in connection with Theorem 6.7). On the other hand, in the case when μ_0 is Lebesgue measure on $\mathbb{R}^{2n}$ the bulk/Gibbs and the Boltzmann entropy coincide in the limit when $N \to \infty$. (In the classical thermodynamical limit, this is discussed in [37, Section 6.2], and in the present mean field setup, it can be shown that both limits coincide with $S(e)$.)

Let now X be a (possible non-compact) subset of $\mathbb{R}^{2n}$ and let ϕ be a defining function for X, i.e., a continuous function such that

$$X = \{\phi \leq 0\}.$$

Endow X with a probability measure μ_0 which is absolutely continuous wrt Lebesgue measure $d\lambda$:

$$\mu_0 = e^{-\Psi_0} d\lambda$$

on $\mathbb{R}^{2n}$. We will identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ in the usual way and denote by $(z_1, ..., z_n)$ the standard holomorphic coordinates on $\mathbb{C}^n$.

1.3.1. Concavity of $S_+^{(N)}(e)$ and $S(e)$ in the high-energy region $e > e_0$ The main results, saying that upper microcanonical entropy $S_+^{(N)}(e)$ and the entropy $S(e)$ are concave in the high-energy regime 1.7 (Theorem 6.7 and Theorem 6.9), are shown to hold under appropriate plurisubharmonicity and symmetry properties of the data. Denoting by $PSH_{\boldsymbol{a}}$ the class of all plurisubharmonic functions which are invariant under the action

$$(z_1, ..., z_n) \mapsto (e^{ia_1\theta} z_1, ..., e^{ia_n\theta} z_n) \tag{1.10}$$

for any $\theta \in \mathbb{R}$, for a given "weight vector" $\boldsymbol{a} \in]0, \infty[^n$, the main results hold under the following:

MAIN ASSUMPTIONS: $\phi, \Psi_0, -V$ are in the class $PSH_{\boldsymbol{a}}(\mathbb{C}^n)$ and $-W$ is in $PSH_{\boldsymbol{a},\boldsymbol{a}}(\mathbb{C}^n \times \mathbb{C}^n)$ for some $\boldsymbol{a} \in]0, \infty[^n$

The definition of plurisubharmonicity is recalled in Sect. 2.3. For the moment, we just point out that the class $PSH_{\boldsymbol{a}}$ is very rich. For example, when the weights a_i are positive integers the class $PSH_{\boldsymbol{a}}$ contains the functions

$$\psi(z) = \log\left(\sum_{j=1}^{r} |P_j(z)|^2\right) \tag{1.11}$$

where P_j is a polynomial in $z_1, .., z_n$, which is homogeneous wrt the scaling action by $\mathbb{C}^*$ on $\mathbb{C}^n$ with weights $\boldsymbol{a}$. In particular, for any $\boldsymbol{a}$ the class $PSH_{\boldsymbol{a}}$ contains $\psi(z) = \log|z|$ as well as $\Psi_0(z) := \sum_{i=1}^n \lambda_i |z_i|^2$, for any positive λ_i. Hence, the Main Assumptions apply to the corresponding Gaussian measures

$$\mu_0 = e^{-\sum_{i=1}^n \lambda_i |z_i|^2} d\lambda. \tag{1.12}$$

In the case when the data are invariant under rotations of the z_i-variables, this is—from a physical point of view—the most natural choices of priors, as they incorporate preservation of angular momentum in the z_i-variables (see the discussion in Sect. 3.3).

An important general feature of the class $PSH_{\boldsymbol{a}}(\mathbb{C}^n)$ is that is closed under scaling by positive numbers, taking sums and maxima, as well as under composition with a complex linear map on $\mathbb{C}^n$ or an increasing convex function, defined on the range of a given $\psi \in PSH_{\boldsymbol{a}}(\mathbb{C}^n)$. This means, in particular, that the Main Assumptions are stable under a range of different regularizations of the data. For example, the Main Assumptions apply to the point vortex model in $X := \mathbb{R}^2$ (formula 1.4) endowed with a centered Gaussian measure. But the Main Assumptions also apply to the standard continuous regularization and smooth regularization of the point vortex model where, for a given positive number δ, the pair interaction $W(x, y)$ is, in the continuous case, modified so that it is constant on $|x - y| \leq \delta$, while the smooth regularization is defined by

$$W_\delta(x, y) = -\frac{1}{2\pi} \log(|x - y| + \delta).$$

More generally, they apply to the regularizations obtained by convolution of $-\log|x|$ with a positive sufficiently rapidly decreasing density on $\mathbb{R}^2$, as used in the vortex blob model [45, Section 6.2.1] (or more generally to the convolution of $-\log|x-y|$ with a smooth density on $\mathbb{R}^2 \times \mathbb{R}^2$). An abundance of other examples in $PSH_{\boldsymbol{a}}$ may be obtained by replacing ψ in formula with $\chi \circ \psi$ for any convex increasing function χ.

Imposing translational and rotational symmetry the Main Assumptions apply, in particular, under the

Homogeneous Assumptions:

- X is either a ball of radius R centered at the origin in $\mathbb{R}^{2n}$ or equal to all of $\mathbb{R}^{2n}$
- $W(x, y) = w(|x-y|)$, $V(x) = v(|x|)$ and $\Psi_0(x) = \psi_0(|x|)$ with $w(r)$, $v(r)$ and $-\psi_0(r)$ *concave* functions of $\log r$ (when $0 < r \leq 2R$) and bounded from below as $r \to 0$.

In fact, the special assumptions imply that $w(r)$ is decreasing in r. In other words, the Homogeneous Assumptions equivalently mean that the pair interaction $W(x, y)$ is repulsive and a concave function of $\log|x-y|$. The special assumptions, apply, for example, to the continuous repulsive power laws

$$W_\alpha(x, y) := -|x-y|^\alpha, \quad \alpha > 0. \tag{1.13}$$

Note that the Homogeneous Assumptions apply, in particular, to the standard centered Gaussian probability measure μ_0 on $\mathbb{R}^{2n}$. However, one virtue of the Main Assumptions is that they, as pointed out above, apply to the more general Gaussian measures 1.12 incorporating conservation of angular momentum in the z_i-variables (as discussed in Sect. 3.3).

1.3.2. Global Concavity of $S(e)$ and Thermodynamic Equivalence of Ensembles In Sect. 6, it is shown that if the assumption that $W(x, y)$ be weakly positive definite is added to the Main Assumptions, then $S(e)$ is globally concave, i.e., concave on all of $\mathbb{R}$ (Theorem 7.1) and finite on $]e_{min}, e_{max}[$. For example, as explained in Sect. 7.1, this applies to the logarithmic interaction in $\mathbb{R}^{2n}$, as well as the continuous power laws 1.13 when $a \in]0, 2]$ and to the exponential pair potential

$$W(x, y) = e^{-a|x-y|}, \quad a > 0,$$

when X is taken to be a disk centered at the origin with radius at most $1/2a$ (known as the *Born–Mayer potential* in chemistry). It should be stressed that neither the power laws with $a \in]0, 1[$, nor the exponential pair potential is concave wrt (x, y). (Otherwise, the concavity of $S_+^{(N)}(e)$ could also be deduced from the ordinary Brunn–Minkowski inequality; compare Remark 6.6.)

We then deduce that thermodynamic equivalence of ensembles holds for any energy level e in $]e_{min}, e_{max}[$ using a general result (Theorem 5.4), saying that for a general lower semi-continuous convex energy functionals $E(\mu)$ and prior μ_0 thermodynamic equivalence of ensembles holds in the low-energy region $]e_{min}, e_0[$ iff $E(\mu)$ and μ_0 satisfy a certain compatibility property (the "energy approximation property"). This property has previously appeared in connection with the study of

large deviation principles for the corresponding canonical ensembles at positive inverse temperatures β [7,17].

We also show that the global concavity of $S(e)$ holds for singular repulsive power laws (Prop 5.5). However, in contrast to the continuous power laws 1.13 (and the repulsive logarithmic interaction) the singular power laws do not satisfy the Main Assumptions. In fact, in this case the global concavity of $S(e)$ in high-energy region $e \geq e_0$ holds for a bad reason: $S(e) \equiv S(e_0)$ and, as a consequence, there are no maximum entropy measures μ^e when $e > e_0$. This means that the equivalence of ensembles at the level of *macrostates* then breaks down (see Sect. 4). Similarly, regularized singular power laws are expected to yield non-equivalent ensembles, and thus, the corresponding entropies are expected to be non-concave (as discussed in [40, Page 252]).

1.3.3. Critical Negative Inverse Temperatures and Existence of Maximum Entropy Measures The singularity structure of a pair interaction $W(x, y)$ satisfying the Main Assumptions can be very complicated, even if $W(x, y)$ is taken to be translationally invariant, i.e.,

$$W(x, y) = -\Psi(x - y) \tag{1.14}$$

for a function Ψ in the class $PSH_{\boldsymbol{a}}(\mathbb{C}^n)$. Still, as shown in Section Sect. 8.3, the singularities are mild enough to ensure that both the *microscopic critical inverse temperature*

$$\beta_{c,N} := \left\{\beta \in \mathbb{R} : \; Z_{N,\beta} := \int_{X^N} e^{-\beta H^{(N)}} \mu_0^{\otimes N} < \infty\right\}$$

and the *macroscopic critical inverse temperature*

$$\beta_c := \inf : \left\{\beta \in \mathbb{R} : \; \inf_{\mu} F_\beta(\mu) > -\infty\right\}$$

are strictly negative. As a consequence, we deduce that, when X is compact, there exists a maximum entropy measure μ^e for any $e \in]e_{min}, e_{max}[$. The concavity of $S_+^{(N)}(e)$ and $S(e)$ is exploited to establish "dual" formulas for $\beta_{c,N}$ and β_c, which hold under the Main Assumptions (Corollary 8.1 and Corollary 8.4):

$$\beta_{c,N} = \lim_{e \to \sup_{X^N} E_N} \frac{dS^{(N)}(e)}{de}, \quad \beta_c = \lim_{e \to \sup_{\mathcal{P}(X)} E(\mu)} \frac{dS(e)}{de} \tag{1.15}$$

(which are decreasing limits when using either left or right derivatives). The derivative $\frac{dS^{(N)}(e)}{de}$ corresponds to the *inverse Gibbs temperature* at *energy* e in the context of small systems [28,36] (when $H^{(N)}$ is replaced by $-H^{(N)}$and e with $-e$ so that $\frac{dS^{(N)}(e)}{de}$ is positive).

Applied to the regularized vortex model W_δ in $\mathbb{R}^2$, the second formula in 1.15 confirms the expectations expressed in [29, Page 855], concerning the slope $dS_\delta(e)/de$ of the corresponding entropy: on the one hand, as e converges to the maximum (finite) value of the corresponding regularized energy $E_\delta(\mu)$ the entropy $S_\delta(e)$ and its slope $dS_\delta(e)/de$ both converge toward $-\infty$. On the other hand, for

a *fixed* e the slope $dS_\delta(e)/de$ converges, as $\delta \to 0$, to the slope $dS_0(e)/de$ for the point vortex model, which, in turn, is close to -4 for large e (with our normalizations).

1.4. Acknowledgments

Thanks to Bo Berndtsson for many stimulating discussions on the topic of [10]. Also thanks to the referee for pointing out some typos. This work was supported by grants from the Knut and Alice Wallenberg foundation, the Göran Gustafsson foundation and the Swedish Research Council.

1.5. Organization

We start in Sect. 2 by introducing a very general setup and provide some background on concavity and on plurisubharmonic functions (appearing in the Main Assumptions). In Sect. 3, general properties of the entropy $S(e)$ are studied. In particular, finiteness and monotonicity properties of $S(e)$ are established and relations to the notion of *thermodynamic equivalence of ensembles* are explored. In Sect. 4, the notion of *macrostate equivalence of ensembles* is discussed and existence results for maximum entropy measures are provided. Then, in Sect. 5 we consider the case when $E(\mu)$ is convex and show that thermodynamic equivalence of ensembles holds in the low-energy region $\{e > e_0\}$ iff the *energy approximation property* holds. In the remaining sections, we specialize to the Main Assumptions. First in Sect. 6, we deduce the concavity of the upper microcanonical entropy $S_+^{(N)}(e)$ (Theorem 6.7) from a complex analog of the Brunn–Minkowski inequality. Then, letting $N \to \infty$ the concavity of the entropy $S(e)$ in the high-energy region $\{e > e_0\}$ (Theorem 6.9) is deduced. In Sect. 7, this is shown to yield global concavity of $S(e)$ when the Main Assumptions are complemented with weak positive definiteness and some examples are exhibited. In Sect. 8, applications to slope formulas of *critical inverse temperatures* are given and some connections to algebraic geometry are explained. In Sect. 9, a strict concavity result for $S(e)$ is deduced under the Homogeneous Assumptions from a uniqueness result for free energy minimizers, established in the companion paper [8].

2. Setup and Preliminaries

2.1. Very General Setup and Notation

A very general formulation of the setup that we shall consider, henceforth called the Very General Setup may be formulated as follows. Let X be a topological space endowed with a probability measure μ_0 and $E(\mu)$ a lsc functional $E(\mu)$ on the space $\mathcal{P}(X)$ of all probability measures on X. We then define the corresponding entropy $S(e)$ and free energy $F(\beta)$ as in formula 1.3 and formula 1.6, respectively. Occasionally, when specializing to the General Setup introduced in Sect. 1.1 the

notation $E_{W,V}(\mu)$ will designate an energy functional $E(\mu)$ of the particular form 1.2.

We set

$$e_{min} := \inf_{\mathcal{P}_0(X)} E(\mu), \quad e_0 := E(\mu_0), \quad e_{max} := \sup_{\mathcal{P}_0(X)} E(\mu)$$

(recall that $\mathcal{P}_0(X)$ denotes the space of all probability measures on X with compact support).

We will mainly consider the case when $X \Subset \mathbb{R}^{2n}$ and the Main Assumptions (or the Homogeneous Assumptions) introduced in Sect. 1.3 hold. These assumptions will be recalled in Sect. 6.1, but we first provide some preliminaries on concavity and plurisubharmonicity.

2.2. *Concave Preliminaries*

We will be discussing concavity properties of the entropy $S(e)$, and we provide some general preliminaries on concave functions. First recall that a function ϕ on a convex subset C of $\mathbb{R}^d$ taking values in $]-\infty, \infty]$ is said to be *convex* on C if, for any given two points x_0 and x_1 and $t \in]0, 1[$,

$$\phi(tx_0 + (1-t)x_1) \leq t\phi(x_0) + (1-t)\phi(x_1)$$

and is *strictly convex* on C if the inequality above is strict for any $t \in]0, 1[$. A function f on C is *(strictly) concave* if $-f$ is (strictly) convex. Here we will be mainly concerned with the case when $d = 1$. In this case, if f is concave and finite on a closed interval $C \subset \mathbb{R}$, but not strictly convex, then there exist two points x_0 and x_1 in C such that f is affine on $[x_0, x_1]$. In Sects. 3, 7, we will use some standard properties of convex functions recalled below, translated into the setup of concave functions (for further background see [51] and [56, Section 2.1.3]). If ϕ is a convex function on $\mathbb{R}^d$, then its*subdifferential* $(\partial\phi)$ at a point $x_0 \in \mathbb{R}^d$ is defined as the convex set

$$(\partial\phi)(x_0) := \left\{ y_0 : \phi(x_0) + y_0 \cdot (x - x_0) \leq \phi(x) \ \ \forall x \in \mathbb{R}^d \right\} \tag{2.1}$$

In particular, if $\phi(x_0) = \infty$, then $(\partial\phi)(x_0)$ is empty. Similarly, if f is concave on $\mathbb{R}^d$ then its *superdifferential* $(\partial f)(x_0)$ is defined as above, but reversing the inequality. In other words, $(\partial f)(x_0) := -(\partial(-f)(x_0)$. In the case when f is concave on $\mathbb{R}$ and finite in a neighborhood of x_0

$$(\partial f)(x) = [f'(x+), f'(x-)],$$

where $f'(x+)$ and $f'(x-)$ denote the right and left derivatives of f at x, respectively. In particular, f is differentiable at x iff $(\partial f)(x)$ consists of a single point. If f is a function on $\mathbb{R}^d$ taking values in $[-\infty, \infty]$ its (concave) *Legendre–Fenchel transform* is the usc and concave function on $\mathbb{R}^d$ (taking values in $[-\infty, \infty[$) defined by

$$f^*(y) := \inf_{x \in \mathbb{R}} (x \cdot y - f(x)).$$

It follows readily from the definitions that

$$y \in \partial f(x) \iff x \in \partial f^*(y). \tag{2.2}$$

Moreover, it is well known that

$$\overline{\partial f(\{f > -\infty\})} = \overline{\partial f^*(\{f^* > -\infty\})}. \tag{2.3}$$

Note that, in general, f^{**} is the concave envelope of f :

$$(f^{**})(x) = \inf_{a \text{ affine}} \{a(x) : \ a \geq f\} = \inf_{g \text{ concave, finite}} \{g(x) : \ g \geq f\}. \tag{2.4}$$

Indeed, the first equality follows directly from the definition and the second one is shown by, for a fixed x, taking $a(x)$ to be any affine function coinciding with g at x and with gradient in $\partial g(x)$.

We will also make use of the following lemmas (which are without doubt essentially well known, but for completeness proofs are provided in appendix):

Lemma 2.1. *Let f be a concave function on $\mathbb{R}$ and assume that f is differentiable in a neighborhood of $[x_0, x_1]$. Then f^* is strictly concave in the interior of $[y_0, y_1] := [f'(x_1), f'(x_0)]$.*

Note that, in general, $f^{**} \geq f$. Concerning the strict inequality we have the following

Lemma 2.2. *Let f be a function on $\mathbb{R}$ such that $\sup_{\mathbb{R}} f < \infty$ and $U \Subset \mathbb{R}$ an open set where f is finite and usc. Then $\{f^{**} > f\} \cap U$ is open in U and f^{**} is affine on $\{f^{**} > f\} \cap U$.*

2.3. Background on Plurisubharmonicity and the Class PSH_a

The Main Assumptions introduced in Sect. 1.3 involve the notion of plurisubharmonicity. While this notion is central in the fields of several complex variables and complex geometry, it may not be familiar to readers lacking background in these fields. We thus recall the main definitions and properties that we shall use and refer to [23, Section 5.A.] for further background. We will identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ in the standard way. A function ψ on $\mathbb{C}^n$ is said to *plurisubharmonic* (*psh*, for short) if ψ is upper semi-continuous (usc) taking values in $[-\infty, \infty[$ and subharmonic along complex lines, i.e., if $\zeta \mapsto \psi(z_0 + \zeta a_0)$ is a local subharmonic function on $\mathbb{C}$ for any given $z_0, a_0 \in \mathbb{C}^n$, or equivalently that

$$\psi(z_0) \leq \frac{1}{2\pi} \int \psi(z_0 + e^{i\theta} a_0) d\theta.$$

In particular, ψ is then subharmonic on $\mathbb{R}^{2n}$. If ψ is smooth, then it is psh iff the complex Hessian $\partial\bar{\partial}\psi$ of ψ is a semi-positive Hermitian matrix at any z :

$$\partial\bar{\partial}\psi(z) := \left(\frac{\partial^2 \psi(z)}{\partial z_i \partial \bar{z}_j}\right) \geq 0, \quad \frac{\partial}{\partial z_i} := \frac{1}{2}\frac{\partial}{\partial x_i} - \frac{i}{2}\frac{\partial}{\partial y_i}$$

Equivalently, a function ψ is psh if, locally, it can be expressed as a decreasing limit of smooth psh functions ψ_j. In fact, ψ_j may be taken as a convolution of ψ with any suitably scaled smooth probability density with compact support. If $-u$ is plurisubharmonic, then u is called *plurisuperharmonic*. An open set Ω in $\mathbb{C}^n$ is said to be *pseudoconvex* if Ω admits a continuous psh exhaustion function ρ, i.e., ρ is psh on Ω and such that $\{\rho \leq C\}$ is a compact subset of Ω. We recall the following essentially standard lemma (see appendix for a proof):

Lemma 2.3. *Let ϕ be a psh function on a pseudoconvex open set Ω. Then $\{\phi < 0\} \cap \Omega$ is also pseudoconvex.*

We also recall that the following standard facts [23, Theorem 5.5.], which allows one to construct a range of different types of psh functions:

Lemma 2.4. *If $\psi_1, ..., \psi_r$ are psh functions and $\chi(t_1, ..., t_r)$ is a convex function on $\mathbb{R}^r$ which is increasing in each t_i, then $\chi(\psi_1, ..., \psi_r)$ is psh. In particular, if $\alpha_1, ..., \alpha_r$ are nonnegative functions, then*

$$\sum_{i=1}^{r} \alpha_i \psi_i, \quad \log \sum_{i=1}^{r} e^{\alpha_i \psi_i} \text{ and } \max\{\psi_1, ..., \psi_r\}$$

are psh functions.

In particular, if ψ is psh and χ is a convex increasing function on $\mathbb{R}$, then the composed function $\chi(\phi)$ is psh. Since $|f(z)|^2$ is psh for any holomorphic function $f(z)$ on $\mathbb{C}^n$ (as follows, for example, directly from the characterization), it follows form the previous lemma that

$$\psi(z) := \log \left(\sum_{i=1}^{r} |f_i(z)|^2 \right)$$

is psh for any given holomorphic functions $f_1, ..., f_r$. In particular, $\log |z|^2$ is psh. Moreover, if a function ψ only depends on the absolute values of z_i, then $\psi(z)$ is psh iff it is convex with respect $(\log |z_1|, ..., \log |z_n|) \in \mathbb{R}^n$.

2.3.1. The Class $PSH_{\boldsymbol{a}}$ Given $\boldsymbol{a} = (a_1, .., a_m) \in]0, \infty[^n$, we denote by $\mathcal{V}_{\boldsymbol{a}}$ the vector field on $\mathbb{C}^n$ defined by

$$\mathcal{V}_{\boldsymbol{a}} := \sum_{i=1}^{m} a_i \frac{\partial}{\partial \theta_i}, \tag{2.5}$$

where $\frac{\partial}{\partial \theta_i}$ denotes the generator of the S^1-action on $\mathbb{C}^n$ which rotates the z_i-coordinate and leaves the other coordinates invariant (i.e., $e^{i\theta} \cdot z := (z_1, ..., e^{i\theta} z_i, ..., z_n)$). In other words, $\mathcal{V}_a$ is the Hamiltonian vector field corresponding to the Hamiltonian

$$h_a(z) := \sum_{i=1}^{m} \frac{1}{2a_i} |z_i|^2 \tag{2.6}$$

on $\mathbb{R}^{2n}$, endowed with its standard symplectic form. Note that the Hamiltonian h_a is plurisubharmonic on $\mathbb{C}^n$ (since $|z_i|^2$ is). Now if U is an open connected subset of $\mathbb{C}^n$ then the class $PSH_{\boldsymbol{a}}(U)$ is defined as the class of all psh functions ψ on U, not identically $-\infty$, such that $\mathcal{V}_{\boldsymbol{a}}(\psi) = 0$. More generally, if X is closed connected subset of $\mathbb{C}^n$ we denote by $PSH_{\boldsymbol{a}}(X)$ the class of all functions ψ such that ψ is in $PSH_{\boldsymbol{a}}(U)$ for some open subset U containing X (depending on ψ).

Example 2.5. (The "algebraic and quasi-homogeneous" case). If $P(z_1, ..., z_n)$ is a quasi-homogeneous polynomial, i.e., there exists exist positive integer weights $a_1, ..a_n$ such that P is homogeneous of degree d wrt the corresponding $\mathbb{R}_+$-action

$$P(c^{a_1} z_1, ..., \lambda^{a_n} z_n) = c^d F(z_1, ..., z_n) \tag{2.7}$$

for any $c \in \mathbb{R}_+$, then $\log |P(z)|$ is in $PSH_{\boldsymbol{a}}(\mathbb{C}^n)$. More generally, if P_j are polynomials on $\mathbb{C}^n$ which are quasi-homogeneous of degree d_j for the same weighs $a_1, ..., a_n$ and $\alpha_i > 0$, then

$$\psi(z) := \log \left(\sum_{j=1}^{r} |P_j(z)|^{\alpha_j} \right) \in PSH_{\boldsymbol{a}}(\mathbb{C}^n) \tag{2.8}$$

In the particular case when all $\alpha_i = 1$ and $d_i = d$, we call d the degree of ψ.

By composing the previous examples with convex increasing functions χ on $\mathbb{R}$, one may fabricate an abundance of examples of functions in the class $PSH_{\boldsymbol{a}}(\mathbb{C}^n)$. For example, $\sum_{j=1}^{M} |P_j(z)|^{\alpha_j}$ is in $PSH_{\boldsymbol{a}}$ if $P_j(z)$ is a homogeneous polynomial (wrt $\boldsymbol{a}$) and $\alpha_j > 0$.

3. General Properties of $S(e)$ and Thermodynamic Equivalence of Ensembles

In this section, general properties of the entropy $S(e)$ are studied and the notion of thermodynamic equivalence of ensembles introduced in [30] is recalled. The main new feature in this section, as compared to the setup in [30], is that $E(\mu)$ is not assumed to be continuous. This leads to some subtle aspects that do not seem to have been addressed before. Throughout the section, we will consider the Very General Setup introduced in Sect. 2.1.

3.1. Monotonicity of $S(e)$

The following lemma generalizes [15, Prop 2.2] (with a similar proof) and involves the following ad hoc property:

Definition 3.1. Assume that X is compact. Then a functional $E(\mu)$ on $\mathcal{P}(X)$ has the *affine continuity property* if for any $\mu_1 \in \mathcal{P}(X)$ such that $E(\mu_1) < \infty$ and $S(\mu_1) > -\infty$ the function $t \mapsto E(\mu_0(1-t) + t\mu_1)$ is continuous on $[0, 1]$. For a general X, the affine continuity property is said to hold if it holds for all compact subsets of X.

Lemma 3.2. *(monotonicity of $S(e)$). Assume that X is compact and $e_0 := E(\mu_0) < \infty$.*

- *If $E(\mu)$ is convex on $\mathcal{P}(X)$, then $S(e)$ is increasing for $e \leq e_0$ and strictly increasing in the subinterval where $S(e) > -\infty$. In particular,*

$$S_-(e) := \sup_{E(\mu)\leq e} S(\mu)$$

- *If $E(\mu)$ has the affine continuity property, then $S(e)$ is decreasing for $e \geq e_0$ and strictly decreasing in the subinterval where $S(e) > -\infty$. In particular,*

$$S_+(e) := \sup_{E(\mu)\geq e} S(\mu)$$

More precisely, in the second point there is no need to assume that E is lsc on $\mathcal{P}(X)$, and thus, it also follows that $S(e)$ is increasing for $e \leq E(\mu_0)$.

Proof. To prove the first point first observe that, since E is lsc and X is compact $\{E(\mu) \leq e\}$ is compact (or empty). We may assume that $S(\mu)$ is not identically equal to $-\infty$ on $\{E(\mu) \leq e\}$. (Otherwise, we are done.) Since $S(\mu)$ is usc, the sup of $S(\mu)$ on the set $\{E(\mu) \leq e\}$ is thus attained at some μ_1 in the set. Assume in order to get a contradiction that $E(\mu_1) < e$. Consider the affine segment μ_t in $\mathcal{P}(X)$ connecting μ_0 and μ_1; $\mu_t := \mu_0(1-t)+t\mu_1$ for $t \in [0, 1]$. By the assumed convexity of $E(\mu)$

$$E(\mu_t) \leq (1-t)E(\mu_0) + tE(\mu_1) < e$$

for t sufficiently small, using that $E(\mu_0) < \infty$. But, as is well known, $S(\mu)$ is strictly concave on $\{S(\mu) > -\infty\} \subset \mathcal{P}(X)$ and attains its maximum at μ_0, and hence, $S(\mu_t) < S(\mu_1)$ for any $t \in [0, 1[$ (as follows from Jensen's inequality). This contradicts the assumption that μ_1 is a maximizer, and hence, it must be that $E(\mu_1) = e$, as desired.

To prove the second point, it will be enough to show that for any $\mu_1 \in \mathcal{P}(X)$ such that $E(\mu_1) \geq e$ and $S(\mu_1) > -\infty$ there exists $\mu \in \mathcal{P}(X)$ such that $E(\mu) = e$ and $S(\mu) \geq S(\mu_1)$. To this, it will, in the light of the previous argument, be enough to show that there exists some $t \in [0, 1]$ such that $E(\mu_t) = e$. But, by assumption $E(\mu_0) \leq e$ and $E(\mu_1) \geq e$. We can thus conclude by invoking the assumption that $E(\mu_t)$ is continuous. Since we have not used that E is lsc on $\mathcal{P}(X)$, the same argument applies to $-E$, which proves the last statement of the lemma. □

3.2. Thermodynamic Equivalence of Ensembles

In this section, we consider the Very General Setup. It follows readily from the definitions that the Legendre–Fenchel transform S^* of S coincides with the free energy $F(\beta)$:

$$S^* = F.$$

Following [30,55], we make the following

Definition 3.3. *Thermodynamic equivalence of ensembles* is said to hold *globally* if

$$S = F^*$$

and thermodynamic equivalence of ensembles is said to hold at*an energy level e* if $S(e) > -\infty$ and

$$S(e) = F^*(e).$$

Recall that, in general, a function $S(e)$ is usc and concave iff $S^{**} = S$. It was shown in [30, Prop 3.1a] that S is always usc under the assumption that X is compact and $E(\mu)$ is continuous wrt the weak topology on $\mathcal{P}(X)$. (This is the case if W and V are continuous.) In this case, global thermodynamic equivalence thus holds iff S is concave. But here we need consider the case when the continuity assumptions are not satisfied (and moreover X may be non-compact). We will impose the following compatibility property between μ_0 and $E(\mu)$.

Definition 3.4. A measure μ_0 in X is said to has the *Energy Approximation Property* if for any compactly supported probability measure μ there exists a sequence $\mu_j \in \mathcal{P}(X)$, supported in the same compact set, converging weakly toward μ with the following properties:

- μ_j is absolutely continuous with respect to μ_0
- $\lim_{j\to\infty} E(\mu_j) = E(\mu)$

Remark 3.5. This property was introduced in the context of large deviation theory in[17] and studied from a potential-theoretic point of view in [7] (see the discussion in the end of Sect. 5.2).

The energy approximation property ensures that $S(e)$ is finite on $]e_{min}, e_{max}[$:

Lemma 3.6. *Assume that μ_0 has the energy approximation property and the affine continuity property on compact subspaces of X. Then $S(e)$ is finite on $]e_{min}, e_{max}[$.*

Proof. By Lemma 3.2, we just have to verify the claim that there exists some $\mu \in \mathcal{P}(X)_0$ such that $E(\mu) \leq e$ and $S(\mu) > -\infty$. To this end, take $\delta > 0$ such that $e-\delta > e_{min}$. By the verify definition of e_{min}, there exists μ such that $E(\mu) \leq e-\delta$. Moreover, by the monotone convergence theorem μ may be chosen to have compact support. Now take a sequence $\mu_j (= \rho_j \mu_0)$ converging weakly toward μ with the energy approximation property. Replacing ρ_j with $\max\{\rho_j, R\}/\int\{\rho_j, R\}\mu_0$ for a given $R > 0$ and using a diagonal argument, we may as well assume that $\rho_j \in L^\infty$. In particular,

$$E(\mu_j) \leq e, \quad S(\mu_j) > -\infty$$

for j sufficiently large, proving the claim when $e \in]e_{min}, e_0[$. A similar approximation argument applies if instead $e \in]e_0, e_{max})[$ (again using Lemma 3.2). Finally, if $e = E(\mu_0)$ then $S(\mu) \geq S(\mu_0) = 0$, which concludes the proof of the claim above. □

Proposition 3.7. *In the Very General Setup, the following holds:*

- *If the entropy $S(e)$ is concave on $]e_{min}, e_{max}[$ and μ_0 has the energy approximation property and the affine continuity property, then $S(e)$ is continuous on $]e_{min}, e_{max}[$ and thermodynamic equivalence of ensembles holds for any $e \in]e_{min}, e_{max}[$.*
- *If the entropy $S(e)$ is concave and continuous on $[e_0, e_{max}[$, then thermodynamic equivalence of ensembles holds for any $e \in [e_0, e_{max}[$ and moreover for any $e \in [e_0, e_{max}[$*
$$S(e) = \inf_{\beta \leq 0} (\beta e - F(\beta)) \tag{3.1}$$
- *If the entropy $S(e)$ is concave and continuous on $]e_{\min}, e_0]$, then thermodynamic equivalence of ensembles holds for any $e \in]e_{\min}, e_0]$*

Proof. In order to show that $S(e_1) = S^{**}(e_1)$ at a given point e_1 in $]e_{max}, e_{min}[$, it is enough to find an affine function s on $\mathbb{R}$ such that $s \geq S$ and $s(e_1) = S(e_1)$ (by formula 2.4). But since s is concave and finite on $]e_{min}, e_{max}[$ its superdifferential ∂S is non-empty, i.e., contains some $\beta \in \mathbb{R}$. This means that the affine function
$$s(e) := \beta(e - e_1) + S(e_1) \tag{3.2}$$
coincides with S at e and has the property that $s \geq S$ on $]e_{min}, e_{max}[$. Hence, by Lemma 3.2, $s \geq S$ on all of $\mathbb{R}$, which proves the first point.

To prove the second point in the proposition, fix $e_1 \in]E(\mu_0), e_{max}[$. By formula 2.4, it will be enough to find an affine function s on $\mathbb{R}$ such that $s \geq S$ and $s(e_1) = S(e_1)$. To this end, first define the function $f(e)$ to be equal to $S(e)$ on $[e_0, e_{max}[$ and e_0 when $e < e_0$. Thus, $f(e) = \max\{e_0, S(e)\}$ is continuous and convex on $]-\infty, e_{max}[$.We then obtain the desired affine function s by picking an element β in the superdifferential ∂f of f at e_1 and again defining $s(e)$ by formula 3.2. Finally, to prove the last formula we have to show that the infimum in formula 3.1 is attained for some $\beta \leq 0$. But this follows from the fact that, in the previous step, β in formula 3.2 is non-positive, since f is decreasing (by Lemma 3.2). The third point is shown in essentially the same way as the second one. □

Remark 3.8. If $e_{max} < \infty$, then it could happen that $S(e_{max}) \neq S^{**}(e_{max})$ in the first point of the previous proposition. Also note that in the case when $E(\mu)$ is of the form $E = E_{W,V}$ (as in formula1.2) then $e_{max} = \infty$ holds if either there exists x_0 such that $V(x_0) = \infty$ or (x_0, y_0) such that $W(x_0, y_0) = \infty$. Indeed, then $E(\mu) = \infty$ for $\mu = \delta_{x_0}/2 + \delta_{x_1}/2$.

As shown in Theorem 5.4, the energy approximation property is not merely a technical assumption, but essential.

3.3. Priors Versus Linear Constraints

Now consider the Very General Setup in the case when X is a domain in $\mathbb{R}^d$ and $\mu_0 = dx$. Given a continuous function ψ_0 and $\lambda \in \mathbb{R}$, we may then replace μ_0 with the prior defined by the probability measure
$$\mu_\lambda := e^{-\lambda\psi_0} dx / Z_\lambda, \quad Z_\lambda := \int_X e^{-\lambda\psi_0} dx,$$

assuming that $Z_\lambda < \infty$. The corresponding entropy function $S_{\mu_\lambda}(e)$ is closely related to the multivariable entropy function $S(e, l)$ on $\mathbb{R}^2$ defined by

$$S(e, l) := \sup_{\mu \in \mathcal{P}(X)_0} \{S(\mu) : \ E(\mu) = e, \ \ L(\mu) = l\}, \ \ L(\mu) := \int_X \psi_0 \mu,$$

obtained by imposing the linear constraint $L(\mu) = l$ (where $S(\mu)$ denotes the entropy of μ relative to dx). Indeed, it follows readily from the definition that, for a fixed e, the Legendre–Fenchel transform of the function $\lambda \mapsto S_{\mu_\lambda}(e)$ is given by $-S(e, l) - \log Z_\lambda$. Hence, under the hypothesis that $S(e, l)$ *is concave and lower semi-continuous wrt* l, inverting the Legendre–Fenchel transform gives

$$S(e, l) = \inf_\lambda \left(S_{\mu_\lambda}(e) + \lambda l + \log Z_\lambda\right).$$

As a consequence, if $S_{\mu_\lambda}(e)$ is globally concave with respect to e, for any fixed λ such that Z_λ is finite, then $S(e, l)$ is globally concave on $\mathbb{R}^2$. Multivariable entropy functions are studied in [30], from the point of view of equivalence of ensembles, but here we will focus on one-variable entropy functions defined with respect to appropriate priors. Note that in the non-compact case when $X = \mathbb{R}^d$ the inclusion of a function ψ_0 with sufficient growth at infinity is crucial in order to get a prior measure with finite total mass. In the presence of rotational symmetry, the standard choice of a prior is a centered Gaussian measure.

Remark 3.9. More generally, given r functions $\psi_1, ..., \psi_r$ on $\mathbb{R}^d$ and $\lambda_1, ..., \lambda_r \in \mathbb{R}^d$ one can consider the prior $\mu_{\boldsymbol{\lambda}} = e^{-\sum \lambda_i \psi_i}/Z_{\boldsymbol{\lambda}}$ and the corresponding entropy function $S(e, \boldsymbol{l})$ on $\mathbb{R}^{1+d}$. Then the previous considerations still apply if λl is replaced by the scalar product between $\boldsymbol{\lambda}$ and $\boldsymbol{l}$.

4. Macrostate Equivalence of Ensembles and Existence of Maximum Entropy Measures

An important motivation for the notion of thermodynamic equivalence of ensembles is that it implies that any maximum entropy measure μ^e (representing an equilibrium macrostate in the microcanonical ensemble) minimizes the free energy $F_\beta(\mu)$ at an inverse temperature β corresponding to the energy level e. This is made precise by the following result (essentially contained in [30]):

Lemma 4.1. *(macrostate equivalence of ensembles). Consider the Very General Setup. Assume that $S^{**}(e) = S(e) > -\infty$ and assume that $\partial S(e)$ is non-empty (this is the case if, for example, $S^{**} = S > -\infty$ in a neighborhood of e). If μ^e is a maximal entropy measure with energy e, i.e., $S(\mu^e) = S(e)$, then μ^e minimizes the free energy functional $F_\beta(\mu)$ for any $\beta \in \partial S(e)$.*

Proof. By assumption $S(e) > -\infty$. Hence, the assumption that $\beta \in \partial S(e)$ means that $\beta \in (\partial F^*)(e)$. Since $F = (F^*)^*$ it follows from the definition of ∂F^* that

$$F(\beta) = -F^*(e) + \beta e$$

(since $0 \in \partial(-F^*(e) + \beta e)$). In other words,

$$\inf_{\mu \in \mathcal{P}(X)} F_\beta(\mu) = -S(\mu^e) + \beta E(\mu^e),$$

which means that μ^e minimizes $F_\beta(\mu)$, as desired. □

Remark 4.2. Without the property that $S(e) = S^{**}(e)$ a maximal entropy measure μ^e will, in general, not minimize $F_\beta(\mu)$. This is discussed in the context of BEG model in the final section of [31] (where it is pointed out that μ^e may be merely a local minimizer of $F_\beta(\mu)$ or even a saddle point). Moreover, even if $S(e) = S^{**}(e)$ there may, in general, exists minimizers of $F_\beta(\mu)$,for $\beta \in \partial S(e)$, which are not maximum entropy measures (at energy e), unless $S(e)$ is strictly concave at e (see [30]).

As shown in [30], the existence of μ^e is automatic for any $e \in]e_0, e_{max}[$, when X is compact and $E(\mu)$ is a continuous functional on $\mathcal{P}(X)$. However, since we do not impose these assumptions in the Main Assumptions we next provide some general existence result for μ^e that will be applied to the Main Assumptions in Sect. 8.4.

4.1. Existence of μ^e when X is Compact

We start with the low-energy region:

Proposition 4.3. *Consider the Very General Setup. Assume that X is compact and that the energy approximation property and the affine continuity property holds. Then, for any $e \in]e_{min}, e_0]$ there exists a maximum entropy measure μ^e.*

Proof. Fix $e \in]e_{min}, e_0]$. First recall that by Lemma 3.6 $S(e)$ is finite. Next, by Lemma 3.2 (and its proof) it is enough to prove that the functional $S(\mu)$ admits a maximizer on $\{E(\mu) \leq e\}$. But since E is lsc, $\{E(\mu) \leq e\}$ is closed in the compact space $\mathcal{P}(X)$, hence compact. The existence of μ^e thus follows from the upper semi-continuity of $S(\mu)$ on $\mathcal{P}(X)$. □

In order to ensure the existence of maximum entropy measures in the high-energy region, we introduce the following stability property:

Definition 4.4. In the Very General Setup, the *thermal stability property* is said to hold if there exists $\beta_0 < 0$ such that

$$\inf_{\mathcal{P}(X)} (\beta_0 E - S) > -\infty.$$

In other words, this property says that the critical inverse temperature β_c (discussed in Sect. 8) is strictly negative. Turning to the General Setup, we will use the following result, shown in the course of the proof of [6, Lemma 2.13, formula 2.12]):

Lemma 4.5. *Consider the General Setup and assume that X is compact. If the thermal stability property holds, then the functional $E_{V,W}$ is continuous on $\{\mu : S(\mu) \geq -C\} \Subset \mathcal{P}(X)$ for any given constant $C > 0$.*

The following result generalizes the existence result in [15], concerning the case when $W(x, y)$ has a logarithmic singularity along the diagonal:

Proposition 4.6. *Consider the General Setup. Assume that X is compact and that the energy approximation property and the thermal stability property hold. Then $S(e)$ is usc on $]e_{min}, e_{max}[$, and for any e in $]e_{min}, e_{max}[$, there exists a maximum entropy measure μ^e.*

Proof. Take $e_j \to e \in]e_{min}, e_{max}[$ and let μ_j be a sequence in $\mathcal{P}(X)$ such that $E(\mu_j) = e_j$ and $S(\mu_j) \geq s(e_j) - 1/j$. In particular, there exists a constant C such that $S(\mu_j) \geq -C$. By the previous lemma, we may, after perhaps passing to a subsequence, assume that $\mu_j \to \mu_\infty$ in $\mathcal{P}(X)$ and $E(\mu_j) \to E(\mu_\infty)$. Hence, $E(\mu_\infty) = e$ and since S is usc on $\mathcal{P}(X)$ $S(\mu_\infty) \geq \limsup_{j\to\infty} S(\mu_j)$. This shows that $S(e) \geq S(\mu_\infty) \geq \limsup_{j\to\infty} S(e_j)$, i.e., that S is usc. Similarly, the *existence* of μ^e also follows from the previous lemma, since it shows that $\{E(\mu) = e\} \cap S(\mu) \geq -C$ is closed (and thus S attains its maximum value there for C sufficiently large). □

If the thermal stability property does not hold, then there may not be no maximum entropy measures,where $S(e)$ is globally concave. In fact, we have the following converse to the previous proposition when $S(e)$ is concave and continuous on $[e_0, e_{max}[$.

Proposition 4.7. *Consider the Very General Setup and assume that X is compact and that there exists a maximum entropy measure μ^e for some $e \in]e_0, e_{max}[$. Then the thermal stability property holds.*

Proof. The assumed concavity of $S(e)$ implies that the right derivative of $S(e)$ tends to β_c as $e \to e_{max}$ (see Cor 8.4 and its proof). Hence, if we assume that the thermal stability property does not hold, i.e., that $\beta_c = 0$ it follows, since $S(e)$ attains its maximum at e and is assumed continuous and concave on $[e_0, e_{max}]$ that $S(e) \equiv S(e_0)$. But $S(\mu) = S(\mu_0)$ iff $\mu = \mu_0$ (which implies $E(\mu) = e_0$), and hence, there exists no maximum entropy measure μ^e when $e > e_0$. □

The previous proposition is illustrated by the case of singular power laws in Sect. 5.3. Before turning to the non-compact case, we point out that the following concrete bound implies the thermal stability property (see Lemma 8.5):

$$\sup_{x\in X} \int e^{-\beta_0\left(\frac{1}{2}W(x,y)+V(y)\right)}\mu_0(y) < \infty, \quad \int_X e^{-\beta_0 V}\mu_0 < \infty, \tag{4.1}$$

for some $\beta_0 < 0$, which will turn out to be satisfied if the Main Assumptions are complemented with the assumption that W is translationally invariant, up to a bounded term.

4.2. Existence of μ^e when X is Non-compact

In order to discuss maximum entropy measures in the case when X is non-compact, we first need to replace the space $\mathcal{P}(X)_0$ of all probability measures with compact support, appearing in the definition1.3 of $S(e)$, with probability measures satisfying an appropriate growth assumption "at infinity." Indeed, if, for example, $E = E_V$ for a lsc function V which is unbounded both from above *and* from below (say, $V(x) = -\log|x|$ in $\mathbb{R}^d$), then it is not a priori clear how to define $E_V(\mu)$ if μ have unbounded support. To handle this issue, we will make the following *growth assumption:* exists a continuous nonnegative function ϕ_0 of X such that

$$-W(x,y) - \frac{1}{2}V(x) - \frac{1}{2}V(y) \leq \frac{1}{2}\phi_0(x) + \frac{1}{2}\phi_0(y) + C_0. \tag{4.2}$$

Then we can decompose

$$\begin{aligned} E(\mu) &= E_{\phi_0}(\mu) - \int \mu\phi_0, \quad E_{\phi_0}(\mu) \\ &:= \int \left(W(x,y) + \frac{1}{2}V(x) + \frac{1}{2}V(y) + \frac{1}{2}\phi_0(x) + \frac{1}{2}\phi_0(y) \right) \mu \otimes \mu \end{aligned} \tag{4.3}$$

where the first term has a well-defined value in $]-\infty, \infty]$, since the corresponding integrand is bounded from below. This means that if we replace $\mathcal{P}(X)$ with the subspace

$$\mathcal{P}_{\phi_0}(X) := \left\{ \mu \in \mathcal{P}(X) : \int_X \phi_0 \mu < \infty \right\}$$

then $S(e)$ may be expressed as

$$S(e) := \sup_{\mu \in \mathcal{P}_{\phi_0}(X)} \{ S(\mu) : \ E(\mu) = e \}, \tag{4.4}$$

where $E(\mu)$ is defined by formula4.3. According to the following result the existence of a maximizer μ^e is guaranteed if ϕ_0 has slower growth then an appropriate exhaustion function ψ_0 of X (i.e., the sublevel sets $\{\psi_0 \leq R\}$ are compact and exhaust X when $R \to \infty$):

Proposition 4.8. *Consider the General Setup and assume that there exists a continuous exhaustion function ψ_0 of X such that the following growth properties hold:*

- $\int e^{\delta\psi_0}\mu_0 < \infty$ *for some* $\delta > 0$
- *The growth assumption 4.2 holds for a ϕ_0 such that $\phi_0/\psi_0 \to 0$ uniformly as $\psi_0 \to \infty$ (e.g., for $\phi_0 = \psi_0^{(1-\epsilon)}$ for some $\epsilon \in]0, 1[$).*

If the thermal stability property holds (i.e., $\beta_c < 0$), then there exists a measure μ^e realizing the sup in formula 4.4 for any given $e \in [e_0, e_{max}[$.

Proof. Setting $\tilde{W}(x,y) := W(x,y) + \frac{1}{2}V(x) + \frac{1}{2}V(y) - \frac{1}{2}\phi_0(x) + \frac{1}{2}\phi_0(y)$, we can express $E_{\phi_0}(\mu) = \int \tilde{W}(x,y)\mu \otimes \mu$. Now fix $e \in [e_0, e[$ and recall that $S(e)$ is finite. Since, by assumption, $\tilde{W}(x,y)$ is lsc on $X \times X$ and bounded from below it extends to a lsc function on $\tilde{X} \times \tilde{X}$, where $\tilde{X}$ denotes the one-point compactification of X. Moreover, we identify ψ_0 with a lsc function on $\tilde{X}$, taking the value ∞ at the point at infinity and μ_0 with a probability measure on $\tilde{X}$, not charging the point at infinity. Accordingly, we can identify $E_{\phi_0}(\mu)$ and $S(\mu)$ with functionals on $\mathcal{P}(\tilde{X})$. Denote by $\tilde{S}(e)$ the corresponding entropy function. Since $\int_{\tilde{X}} \mu\psi_0 < \infty$ implies that μ does not charge the point at infinity, it will, in order to prove the proposition, be enough to show that the sup defining $\tilde{S}(e)$ is attained. To this end, take a sequence $\mu_j \in \mathcal{P}(X)$ such that $E(\mu_j) = e$ and $S(\mu_j)$ increases to $\tilde{S}(e)$. Decompose $\mu = e^{-\delta\Psi_0}\mu_\delta$ for $\delta > 0$ such that $\mu_\delta := e^{\delta\Psi_0}\mu_0$ has finite total mass. Then there exists a constant C such that

$$S(\mu_j) = S_{\mu_\delta}(\mu) - \delta \int \Psi_0 \mu_j \geq -C. \tag{4.5}$$

Since $S_{\mu_\delta}(\mu)$ is uniformly bounded from above on $\mathcal{P}(X)$ (using that μ_δ has total finite mass), this means that there exists a finite constant C_δ such that

$$\int \psi_0 \mu_j \leq C_\delta < \infty. \tag{4.6}$$

Now, since $\tilde{X}$ compact we may, after perhaps passing to a subsequence, assume that $\mu_j \to \mu_\infty$ weakly in $\mathcal{P}(\tilde{X})$ for some μ_∞ (which, by the bound 4.6, does not charge the point at infinity). Moreover, combining the bound 4.6 with the growth assumption on the continuous function ϕ_0 gives (using Markov's inequality) that

$$\lim_{j\to\infty} \int \phi_0 \mu_j = \int \phi_0 \mu_\infty.$$

Since $S(\mu)$ is usc on $\mathcal{P}(\tilde{X})$, all that remains is to verify that

$$\lim_{j\to\infty} E_{\phi_0}(\mu_j) = E_{\phi_0}(\mu_\infty) \tag{4.7}$$

To this end, we rewrite the assumed thermal stability property as

$$\beta_0 E_{\phi_0}(\mu) - \beta_0 \int \phi_0 \mu - S(\mu) \geq -C_0, \quad \beta_0 < 0 \tag{4.8}$$

Note that

$$-\beta_0 \int \phi_0 \mu - S(\mu) = -S_{\mu_{\beta_0}}(\mu), \quad \mu_{\beta_0} := e^{\beta_0\phi_0}\mu_0, \tag{4.9}$$

where the measure μ_{β_0} has finite mass (since $\beta_0 \leq 0$ and $\phi \geq 0$) and thus identifies with a measure on $\tilde{X}$. Accordingly, can view 4.8 as an inequality on $\mathcal{P}(\tilde{X})$, saying that lsc functional $E_{\phi_0}(\mu)$ has the thermal stability property wrt the measure μ_{β_0} on the compact space $\tilde{X}$. Thus, it follows from Lemma 4.5 that E_{ϕ_0} is continuous on $\{S_{\mu_{\beta_0}}(\mu) \geq -C\}$. Finally, combining 4.9, 4.6 and 4.5 reveals that $S_{\mu_{\beta_0}}(\mu_j) \geq -C$ for some constant C, and hence, the desired convergence 4.7 follows. □

Remark 4.9. To see that the growth properties in the previous proposition are essential consider the case when $X = \mathbb{R}^d$, $\mu_0 = e^{-|x|}dx$ and $V(x) = -|x|^p$ for $p > 0$. Then the thermal stability property does hold (in fact, $\beta_c = -\infty$, since $Z_\beta := \int e^{-\beta V}\mu_0 < \infty$ for any $\beta < 0$). Moreover, $\int e^{\delta\psi_0}\mu_0 < \infty$ for $\psi_0 := |x|^2$. However, for $e \leq e_0$ a maximum entropy measure μ^e only exists under the assumption that $p < 2$, i.e., precisely when $-V/\psi_0 \to \infty$ (indeed, if μ^e exists, then $\mu^e = e^{-\beta V}/\int e^{-\beta V}\mu_0$ for some $\beta > 0$ (see Sect. 5.1.1).

5. Concavity of $S(e)$ in the Low-energy Region for Convex $E(\mu)$

5.1. Concavity and Monotonicity of $S(e)$ in the Low-energy Region $e \leq e_0$ when $E(\mu)$ is Convex.

We now consider the entropy $S(e)$ in the low-energy region $e \leq e_0$ under the assumption that $E(\mu)$ is convex. By way of motivation, we start with the case when $E(\mu)$ is affine.

5.1.1. The Case of $E(\mu)$ Affine In the case when $E(\mu)$ is affine on $\mathcal{P}(X)$ it follows directly from the definition of $S(e)$ that $S(e)$ is globally concave, using the concavity of $S(\mu)$ on $\mathcal{P}(X)$. Moreover, if X is compact and $E(\mu) = \langle V, \mu\rangle$ for $V \in C^0(X)$, then a duality argument reveals that $S(e)$ is finite and strictly concave on $]e_{min}, e_{max}[$. In fact,

$$S(e) = F_V^*(e), \quad F_V(\beta) = -\log \int_X e^{-\beta V}\mu_0,$$

where $F_V(\beta) < \infty$ for all β, since X is compact and V is bounded. Indeed, in this case it follows from Jensen's inequality that the free energy $F(\beta)$ is of the form $F_V(\beta)$ above.[1] Since $F_V(\beta)$ is differentiable on all of $\mathbb{R}$ and its derivative tends to $\inf_X V (= e_{min})$ and $\sup_X V (= e_{max})$ as $\beta \to \infty$ and $\beta \to -\infty$, respectively, it thus follows from Lemma 5.2 that $S_V(e)$ is strictly concave on $]e_{min}, e_{max}[$. However, if X is non-compact, then the strict concavity of $S_V(e)$ may fail as illustrated by the following simple example:

$$X = \mathbb{R}, \quad \mu_0 = e^{-|x|}dx, \quad V(x) = |x|^2.$$

In this case, $E(\mu_0) < \infty$, but $\int e^{-\beta V}\mu_0 < \infty$ iff $\beta \geq 0$. It follows that $S(e) = S(e_0) = 0$ for $e > e_0$, and thus, $S(e)$ is not strictly concave. Indeed, applying the second point in Prop 3.7, we get, for $e \geq e_0$,

$$S(e) = \inf_{\beta \leq 0} (\beta e - F_V(\beta)).$$

However, since $F_V(\beta) = \infty$ for $\beta < 0$ the right-hand side above is attained at $\beta = 0$, showing that $S(e) = 0$. Also note that replacing V with $-V$ yields an example where $S(e)$ fails to be strictly concave in the low-energy region. Note

[1] This is an instance of the classical Gibbs variational principle.

also that in this example, the sup defining $S(e)$ is not attained in the region where $S(e) = S(e_0)$, if $e \neq e_0$. Indeed, if the sup is attained at μ^e satisfying $E(\mu) = e$, then $S(\mu^e) = S(\mu_0)$, and hence, $\mu^e = \mu_0$, which forces $e = E(\mu_0) := e_0$.

5.1.2. The Case of $E(\mu)$ Convex Using Lemma 3.2, we observe that similar arguments apply in the low-energy region when $E(\mu)$ is convex, under some further regularity assumptions.

Proposition 5.1. *Let X be a topological space and $E(\mu)$ a lsc convex functional on $\mathcal{P}(X)$ and $\mu_0 \in \mathcal{P}(X)$.*

- *If X is compact and $e_0 := E(\mu_0) < \infty$, then $S(e)$ is concave on $]-\infty, e_0]$.*
- *If X is σ-compact (i.e., a countable union of compact space) and $E(1_K \mu_0) < \infty$ for any compact subspace K of X, then, if the energy approximation property holds, $S(e)$ is concave, increasing and finite (hence continuous) on $]e_{min}, e_0[$.*

Proof. Given e_1 and e_2 in $]-\infty, e_0]$ and $t \in [0, 1]$ set $e_t := (1-t)e_0 + te_1$. Let μ_1 and μ_2 be contenders for the sup defining $S(e_1)$ and $S(e_2)$, respectively. Set $\mu_t := (1-t)\mu_1 + t\mu_2$. Since $E(\mu)$ is assumed convex, $E(\mu_t) \leq e_t$. Hence, if X is compact and $E(\mu_0) < \infty$, then Lemma 3.2 gives, $S(e_t) \geq S(\mu_t) \geq (1-t)S(\mu_1) + tS(\mu_2)$, using that S is concave on $\mathcal{P}(X)$. This proves the first point. To prove the second one, we write X is an increasing union of compact subspaces X_R. Denoting by S_R the entropy corresponding to X_R, it follows directly from the definition that $S_R(e) \leq S(e)$. Now, by the energy approximation property in Lemma 3.6, $-\infty < S_R(e) \leq S(e)$. A slight variant of the argument in the end of the proof of Theorem 6.9 then shows that $S_R(e)$ increases toward $S(e)$ as $R \to \infty$. Hence, we can conclude by invoking the first point. □

Next, a different duality argument yields *strict* concavity and continuity up to $e = e_0$ when X is compact. The proof uses the following duality criterion:

Lemma 5.2. *Consider the Very general setup and assume that X is compact and that the energy approximation property holds. If $F(\beta)$ is differentiable in a neighborhood of $[\beta_0, \beta_1]$ and $[F'(\beta_1), F'(\beta_0)] \subset]e_{min}, e_{max}[$, then $S(e)$ is strictly concave and equal to F^* on $[F'(\beta_1), F'(\beta_0)]$. Moreover, in general, if F is differentiable at β, then $F'(\beta) = E(\mu_\beta)$ for any minimizer of F_β.*

Proof. Since $F(\beta)$ is concave and $F = S^*$ Lemma 2.1 implies that S^{**} is strictly concave on $[F'(\beta_1), F'(\beta_0)]$. Next, by Prop 4.1 S is usc on $U :=]e_{min}, e_{max}[$, and hence, Lemma 2.2 forces $S^{**} = S$ on $[F'(\beta_1), F'(\beta_0)]$, which concludes the proof of the first statement. The last statement follows directly from letting δ tend to zero (from left and from right) in the inequality

$$F(\beta + \delta) - F(\beta) \leq F_{\beta+\delta}(\mu_\beta) - F_\beta(\mu_\beta) = \delta E(\mu_\beta). \tag{5.1}$$

□

Proposition 5.3. *Assume that X is compact, $E(\mu)$ is lsc and convex on $\mathcal{P}(X)$. Then $S(e)$ is strictly concave and $S(e) = F^*(e)$ on $]e_{min}, e_0[$. Moreover, $S(e)$ is continuous on $]e_{min}, e_0]$.*

Proof. The concavity was shown in [29] under the extra assumption that $E(\mu)$ be continuous on $\mathcal{P}(X)$. Here we note that an alternative argument yields *strict* concavity under the more general assumptions in the proposition. The starting point is the observation that $F_\beta(\mu)$ is convex on $\mathcal{P}(X)$ for $\beta \geq 0$ and strictly convex on $\{F_\beta < \infty\}$. Indeed, since $E(\mu)$ is assumed convex this follows directly from the corresponding property of $-S(\mu)$ (i.e., from the case $\beta = 0$), which is well known [22]. It then follows from general principles that $F(\beta)$ is differentiable with derivative at β given by $e(\beta) := E(\mu_\beta)$, where μ_β is the unique minimizer of F_β. Indeed, this follows from the general statement in appendix of [9], using that $E(\mu_\beta)$ is continuous in β by the argument below. Hence, by Lemma 5.2 $S(e)$ is strictly concave and equal to F^* on the interval $]\lim_{\beta\to\infty} e(\beta), \lim_{\beta\to 0} e(\beta)[$. By the concavity of $F(\beta)$ the function $e(\beta)$ is decreasing. Moreover, the energy approximation property implies, in a rather straightforward manner, that

$$\lim_{\beta\to 0} e(\beta) = e_{min}$$

(see [7]). All that remains is thus to verify that

$$\lim_{\beta\to 0} e(\beta) = e_0.$$

But since $e(\beta)$ is decreasing, this follows readily from the lower semi-continuity of $E(\mu)$ (see [7]). To prove that that $S(e)$ is continuous on $]e_{min}, e_0]$ it will be enough, by the previous step, to show that $F^*(e)$ is continuous on $]e_{min}, e_0]$ and $F^*(e_0) = 0$. Since F^* is concave it is enough to show that $F^*(e)$ is finite on $]e_{min}, e_{max}[$. But

$$S \leq S^{**} = F^* \leq 0,$$

where the last inequality follows from restricting the inf defining F^* to $\beta = 0$. Since S is finite (by the previous proposition) it follows that is F^* is also finite and thus continuous on $]e_{min}, e_{max}[$. Hence, by the continuity of F^* at e_0 we get $S(e) \to F^*(e_0)$ as $e \to e_0$. But

$$F^*(e_0) = \inf_{\beta\in\mathbb{R}} (\beta e_0 - F(\beta)), \quad F(\beta) = \inf_{\mu\in\mathcal{P}(X)} \beta E(\mu) - S(\mu) \leq \beta E(\mu_0) - S(\mu_0) = \beta e_0$$

Hence,

$$F^*(e_0) = \inf_{\beta\in\mathbb{R}} (\beta e_0 - F(\beta)) \geq \inf_{\beta\in\mathbb{R}} (\beta e_0 - \beta e_0) = 0.$$

which gives $\liminf_{e\to e_0} S(e) \geq 0$. Since, trivially, $S(e) \leq S(e_0) = 0$ it follows that $S(e) \to 0 = S(e_0)$, as desired. □

5.2. The Necessity of the Energy Approximation Property for Thermodynamic Equivalence of Ensembles

We next show that the assumption that μ_0 has the energy approximation property, used in the previous section is necessary for having thermodynamic equivalence of ensembles:

Theorem 5.4. *Let X be a compact topological space endowed with a measure μ_0 such that $E(\mu_0) < \infty$ and assume that $E(\mu)$ is a lsc convex functional on $\mathcal{P}(X)$ and $V \in C^0(X)$. Denote by $S_V(e)$ entropy $S_V(e)$ associated to $E_V(\mu) := E(\mu) + \langle V, \mu \rangle$ and the measure μ_0. Then $S_V(e)$ is concave and finite on $]e_{min}, e_0]$ for any $V \in C^0(X)$ iff μ_0 has the energy approximation property. In other words, thermodynamic equivalence of ensembles holds in the low-energy regions $]e_{min}, e_0]$ for all $V \in C^0(X)$ iff μ_0 has the energy approximation property.*

Proof. First assume that μ_0 has the energy approximation property. Since $E_V(X)$ is lsc and convex it then follows from the previous proposition that $S_V(e)$ is concave on $]e_{min}, e_0[$ for any $V \in C^0(X)$. To prove the converse first note that, by the third point in Prop 3.7, the restriction of S_V to $]e_{min}, e_0]$ is equal to the Legendre–Fenchel transform of $F_V(\beta)$. Hence, since $S_V(e)$ is assumed finite on $]e_{min}, e_0[$ it follows from the property of gradient images in formula 2.3 that $dF_V(\beta)/d\beta \to e_{min}$ as $\beta \to \infty$ (using either left or right derivatives). Since $F_V(\beta)$ is concave this means that

$$\lim_{\beta \to \infty} F_V(\beta)/\beta = \inf_{\mathcal{P}(X)} E(\mu).$$

Now, by definition, $F_V(\mu)/\beta = E(\mu) - S(\mu)/\beta$, and hence,

$$\lim_{\beta \to \infty} \inf_{\mathcal{P}(X)} (E(\mu) - S(\mu)/\beta) = \inf_{\mathcal{P}(X)} E(\mu).$$

However, as shown in [7] the latter convergence holds for all $V \in C^0(X)$ iff μ_0 has the approximation property (briefly, the point is that the convergence in question is, since $E(\mu)$ is convex equivalent to the Γ-convergence of $E(\mu) - S(\mu)/\beta$ toward $E(\mu)$, which, in turn, is equivalent to the energy approximation property of μ_0). □

In the case when $W(x, y)$ is the repulsive logarithmic interaction in $\mathbb{R}^2$ or $W(x, y) = |x - y|^{-s}$ in $\mathbb{R}^n$ for $s \in]d-2, d[$ (specializing to the Coulomb interaction when $s = d - 2$) a potential-theoretic characterization of measures μ_0 satisfying the energy approximation property was given in [7]. In particular, it was shown that any compact domain X with smooth boundary admits probability measures μ_0 with support X and a density in $L^1(X, dx)$, for which the energy approximation property fails. Hence, by the previous theorem thermodynamic equivalence of ensembles also fails. On the other hand, Lebesgue measure on a compact domain X has the energy approximation property, if X is non-thin at all boundary points, in the sense of classical potential theory. For example, this is the case if any point $x \in \partial X$ is the vertex of a cone contained in X (e.g., if X is a Lipschitz domain).

5.3. The Catastrophic Case of Singular Power Laws

Consider now the case when X is compact and $W(x, y)$ is a repulsive power-law singularity

$$W(x, y) := |x - y|^{-\alpha} + H(x, y), \quad \alpha > 0 \tag{5.2}$$

for H continuous on $X \times X$. In particular, $e_\infty = \infty$. We will say that a compact set X is *strictly star-shaped* if for any point $x \in X$ and $c \in [0, 1[$ the scaled point cx is contained in the interior of X.

Proposition 5.5. *Consider a repulsive power-law singularity W on a compact strictly star-shaped subset X of $\mathbb{R}^d$ and let μ_0 be proportional (or comparable) to Lebesgue measure on X. Then $S(e)$ is concave on $\mathbb{R}$ and finite (hence continuous) on $]e_{min}, \infty[$. Moreover, $S(e) = S(e_0)$ for any $e \geq e_0$, and as a consequence, there exists no maximum entropy measure μ^e when $e > e_0$.*

Proof. For simplicity we will assume that $H = 0$, but the general case is shown in essentially the same way. First observe that E has the energy approximation property. Indeed, using that X is assumed strictly star-shaped and $W(e^t)$ is monotone in t it is, by the argument in the proof of Lemma 6.3, enough to show this when the support of μ is contained in the interior of X. Let μ_ϵ be defined as in formula 6.2. First using that E is convex and then that E is translationally invariant gives

$$E(\mu_\epsilon) \geq \int_{a \in B_\epsilon} \sigma_\epsilon E\left((T_a)_* \mu\right) = \int_{a \in B_\epsilon} \sigma_\epsilon E\left(\mu\right) = E(\mu).$$

The reversed asymptotic inequality follows directly from the lower semi-continuity of E, resulting from the assumed lower semi-continuity of w. Next note that the affine continuity property appearing in Lemma 3.2 holds, as is seen by modifying the proof of Lemma 6.2. Indeed, using the Cauchy–Schwartz inequality the finiteness in formula 6.1 follows from the positive definiteness of $W(x, y)$ and that $W \in L^1(X^2)$. Hence, by Prop 5.1 $S(e)$ is concave and continuous on $]e_{min}, e_0]$ and $S(e) > -\infty$ for all $e \in]e_{min}, \infty[$.

Next, we will show that $S(e) = S(e_0)$ for any $e > e_0$. To this end it will, thanks to Lemma 3.2, be enough to show that there exists a family $\nu_\epsilon \in \mathcal{P}(X)$ parametrized by $\epsilon > 0$ such that, as $\epsilon \to 0$,

$$(i)\ E(\nu_\epsilon) \to \infty, \quad (ii)\ S(\nu_\epsilon) \to S(\mu_0) = 0. \tag{5.3}$$

We will take $\nu_\epsilon = \epsilon^{\alpha/4}(T_\epsilon)_* \mu_0 + (1 - \epsilon^{\alpha/4})\mu_0$, where, as before, T_ϵ denotes the scaling map $x \mapsto \epsilon x$. First observe that since $S((T_\epsilon)_* \mu_0) = d \log \epsilon$ for some (as seen by making the change of variables $x \mapsto T_\epsilon(x)$ in the integrals) we get, using the concavity of $S(\mu)$ on $\mathcal{P}(X)$,

$$S(\nu_\epsilon) \geq \epsilon^{\alpha/4} S((T_\epsilon)_* \mu_0) + (1 - \epsilon^{\alpha/4}) S(\mu_0) \geq \epsilon^{\alpha/4} d \log \epsilon + 0,$$

which verifies the second item in formula 5.3. To prove the first one, observe that, making the change of variables $x \mapsto T_\epsilon(x)$ in the integrals, reveals that $E((T_\epsilon)_* \mu_0) = \epsilon^{-\alpha} E(\mu_0)$, and hence, $E(\nu_\epsilon) = \epsilon^{\alpha/2} E((T_\epsilon)_* \mu_0) = \epsilon^{\alpha/2} \epsilon^{-\alpha} E(\mu_0)$,

which proves the first item in formula 5.3. Hence, $S(e) = S(e_0)$ for any $e > e_0$. Since we have shown that $S(e)$ is concave, increasing and continuous on $]e_{min}, e_0]$ and $S(e) = S(e_0)$ it follows that $S(e)$ if concave and continuous on $]e_{min}, e_0]$. Moreover, since $S(\mu) = S(\mu_0)$ iff $\mu = \mu_0$ (which implies $E(\mu) = e_0$) it follows that there exists no maximum entropy measure for $e > e_0$. □

More precisely, the proof of the previous proposition reveals that, for any given $e > e_0$ there exists $\mu \in \mathcal{P}(X)$ with energy e, i.e., $E(\mu) = e$, whose entropy $S(\mu)$ can be taken to be arbitrarily close to the maximal entropy and such that μ has a "core–halo" structure, i.e., μ is a convex combination

$$(1 - \lambda)\mu_0 + \lambda\mu_1$$

for some $\lambda \in]0, 1[$, where μ_1 (the "core") can be taken to be a uniform measure of arbitrarily large density on a ball with arbitrarily small radius, centered at a given point in the interior of X.

Remark 5.6. The assumption the X be star-shaped was imposed to ensure the energy approx property and can certainly be relaxed. For example, if $W(x, y)$ is the Coulomb interaction in $\mathbb{R}^n$, then, as pointed out in Sect. 5.2, the energy approximation property in question holds if the interior of X is non-thin at all boundary points.

The previous proposition also applies to the corresponding singular power laws obtained by switching the sign of W, if at the same time e is replaced by $-e$ (using that $S(e)$ is concave iff $S(-e)$ is). In the case of the Newtonian pair interaction in $\mathbb{R}^3$, the nonexistence of the corresponding maximum entropy measure is closely related to the *gravitational catastrophe* (Antonov instability) which plays a central role in astrophysics [13, Section 4.10.1].

6. Concavity in the High-Energy Region under the Main Assumptions

We start by recalling the Main/Homogeneous Assumptions stated in the introduction of the paper.

6.1. The Main and Homogeneous Assumptions

Let X is a (possible non-compact) subset of $\mathbb{R}^{2n}$, and let ϕ be a defining function for X, i.e., a continuous function such that

$$X = \{\phi \leq 0\}.$$

Endow X with a measure μ_0 which is absolutely continuous wrt Lebesgue measure $d\lambda$:

$$\mu_0 = e^{-\Psi_0} d\lambda.$$

on $\mathbb{R}^{2n}$. As pointed out above, we will identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ and denote by $(z_1, ..., z_n)$ the standard holomorphic coordinates on $\mathbb{C}^n$.
MAIN ASSUMPTIONS: $\phi \in PSH_{\boldsymbol{a}}(\mathbb{C}^n)$, Ψ_0, $-V \in PSH_{\boldsymbol{a}}(X)$ and $-W \in PSH_{\boldsymbol{a},\boldsymbol{a}}(X \times X)$ for some $\boldsymbol{a} \in]0, \infty[^n$

The class $PSH_{\boldsymbol{a}}(X)$ was defined in Sect. 2.3 and the class $PSH_{\boldsymbol{a},\boldsymbol{a}}(X \times X)$ is defined similarly, by identifying $\mathbb{C}^n \times \mathbb{C}^n$ with $\mathbb{C}^{2n}$ and using the weight vector $(\boldsymbol{a}, \boldsymbol{a})$. Recall that we also introduced the Homogeneous Assumptions in Sect. 1.3, which according to the following lemma is a special case of the Main Assumptions:

Lemma 6.1. *If the Homogeneous Assumptions are satisfied, then so are the Main Assumptions.*

Proof. First note that $-v(r)$ is increasing in r. Indeed, since $\phi(t) := v(e^t)$ is convex in t the limit, denoted by $\dot{\phi}(-\infty)$, of the one sided derivative $\phi'(t+)$ exists as $t \to \infty$. Since $\phi(t)$ is assumed bounded from above as $t \to -\infty$ it follows that $\dot{\phi}(-\infty) \geq 0$. Hence, by convexity, $\phi'(t+) \geq 0$ for all t, showing that $\phi(t)$ is increasing in t, as desired. Since $\log|z|$ is psh this means that $V(z)$ is an increasing convex function of the psh function $\log|z|$ when $z \neq 0$ and bounded from above in a punctured neighborhood of the origin in $\mathbb{C}^n$. But any psh function which is locally bounded from a above on the complement of a pluripolar set A (i.e., a set which is locally the $-\infty$-set of a psh function) extends over A to a unique psh function [23, Thm 5.24]. Thus, $-V$ indeed defines a psh function on X and the same argument applies to $\Psi_0(z)$. Similarly since $(z, \zeta) \mapsto (z - \zeta)$ is holomorphic, the function $\log|z - \zeta|$ is psh on $\mathbb{C}^{2n}$, and thus, $-W(z, \zeta)$ is an increasing convex function of a psh function when $\log|z - \zeta| \neq -\infty$ and thus psh. All in all this means that the Main Assumptions are satisfied with, for example, $a_0 = ... = a_n = 1$. □

We next show that the "affine continuity property" and energy approximation property introduced in Sect. 3 both hold under the Main Assumptions.

Lemma 6.2. *Under the Main Assumptions the affine continuity property holds.*

Proof. Since X may be assumed compact and W is lsc we may after perhaps replacing W with $W + C$, i.e., E with $E + C$, as well assume that $W \geq 0$ on $X \times X$. Hence, by the dominated convergence theorem it is enough to verify that if $E(\mu) < \infty$, then

$$\int_{X \times X} W \mu \otimes \mu_0 < \infty. \tag{6.1}$$

Set $u_\mu(x) := \int_X W(x, y)\mu(y)$. Since $-W$ is psh on a neighborhood of $X \times X$ the function $-u_\mu(x)$ psh on X. Now since X is connected, as shown in the course of the proof of Theorem 6.7, any psh function (or more generally, subharmonic function) is either identically equal to $-\infty$ or in L^1_{loc} (as follows from the submean property of subharmonic functions). But, by assumption, $\int_X u_\mu \mu = E(\mu) < \infty$, and hence, $-u_\mu$ cannot be identically $-\infty$. Since $\mu_0 = e^{-\Psi_0} d\lambda$ 6.1 thus follows directly in the case when Ψ_0 is bounded. In the general case we can use that by Cor 8.1, there exists $q > 1$ such that $\int_X e^{-q\psi} d\lambda < \infty$ for any psh function ψ (not identically $-\infty$) and apply Hölder's inequality to conclude. □

Lemma 6.3. *Assume that the Main Assumptions hold. Then the corresponding energy approximation property is satisfied.*

Proof. First consider the case when the support of μ is contained in the interior of X. Set

$$\mu_\epsilon := \int_{a\in B_\epsilon} \sigma_\epsilon (T_a)_* \mu, \tag{6.2}$$

where, for a given $a \in \mathbb{R}^d$, T_a is the map $x \mapsto x + a$ and B_ϵ denotes the ball of radius ϵ centered at the origin. For ϵ sufficiently small μ_ϵ is also supported in X. It is a standard fact that μ_ϵ is absolutely continuous wrt Lebesgue measure and $\mu_\epsilon \to \mu$ weakly as $\epsilon \to 0$. Moreover,

$$\lim_{\epsilon\to 0} E(\mu_\epsilon) = E(\mu). \tag{6.3}$$

Indeed, setting $\Psi := -W(x, y)+V(x)+V(y)$ and changing the order of integration gives

$$\begin{aligned} -E(\mu_\epsilon) &= \int_{B_\epsilon\times B_\epsilon} \sigma_\epsilon \otimes \sigma_\epsilon \int_{X\times X} \Psi (T_a)_*\mu \otimes (T_b)_*\mu \\ &= \int_{X\times X} \mu(x) \otimes \mu(y) \int_{B_\epsilon\times B_\epsilon} \sigma_\epsilon \otimes \sigma_\epsilon \Psi(x + a, x + b). \end{aligned}$$

Recall that, in general, if $\psi(x)$ is a subharmonic function, then $\int_{B_\epsilon} \sigma_\epsilon \psi(x + a)$ decreases to $\psi(x)$, as ϵ decreases to 0. Hence, the convergence 6.3 follows from the monotone convergence theorem. Finally, for a general $\mu \in \mathcal{P}(X)$ we consider for $\tau \in \mathbb{C}$ the holomorphic action $\tau \circledcirc z$ defined by formula 6.6. If $z \in X$ and $\Re\tau < 0$, then it follows readily from the definitions that $\tau \circledcirc z$ is contained in the interior of X (compare the proof of Theorem 6.7). Hence, fixing $t < 0$ and setting $F_t(z) := t \circledcirc z$ the probability measure $\mu^t := (F_t)_*\mu$ is supported in the interior of X. Moreover, μ^t converges weakly toward μ when $t \to 0$ and

$$\lim_{t\to 0} E(\mu^t) = E(\mu). \tag{6.4}$$

Indeed, proceeding as above

$$-E(\mu^t) = \int \Psi(e^t x, e^t y)\mu(x) \otimes \mu(y)$$

where $t \mapsto \Psi(e^t x, e^t y)$ is increasing (as shown in the course of the proof of Theorem 6.7). Hence, the convergence 6.4 follows from the monotone convergence theorem. We can thus conclude the proof by combining the convergence in 6.3 and 6.3, using a standard diagonal argument.

Combining the previous two lemmas with Lemma 3.6, we arrive at the following □

Proposition 6.4. *Under the Main Assumptions $S(e)$ is finite on $]e_{min}, e_{max}[$.*

6.2. *Concavity of the Microcanonical Entropy* $S_+^{(N)}(e)$

The following result is a slight generalization of a result shown in the course of the proof of [10, Theorem 2.3], which, in turn, is based on the main result in [11].

Proposition 6.5. *Let Y be a pseudoconvex domain, Ψ a psh function on Y and μ_0 a measure on Y such that $\mu_0 = e^{-\psi_0} d\lambda$ for a psh function ψ_0 on Y. Assume that Y is endowed with a holomorphic action of by compact group G such that If Ψ and μ_0 are G-invariant and assume also that for any $t \in \mathbb{R}$ any G-invariant holomorphic function on $\{\Psi < t\}$ is constant. Then either the function*

$$t \mapsto \log \mu_0 \{\Psi < t\}$$

is identically equal to $+\infty$ or concave. In particular, if μ_0 is moreover assumed to have finite total mass, then the concavity in question holds.

Proof. First recall the main result in [11]. Consider $\mathbb{C}^{n+1}$ with holomorphic coordinates $(z, t) \in \mathbb{C}^n \times \mathbb{C}$. Let $\mathcal{D}$ be a pseudoconvex domain in $\mathbb{C}^{n+1}$ endowed with a psh function $\psi(z, t)$. Denote by D_t the subset of $\mathcal{D}$ obtained by fixing the t-coordinate. According to the main result of [11] the function $B_t(z)$ on $\mathcal{D}$ defined by

$$B_t(z) := \sup \left\{ |f(z)|^2 : \ f \text{ holomorphic on } D_t \ \text{ and } \int_{D_t} |f|^2 e^{-\psi(\cdot,t)} d\lambda \leq 1 \right\} \tag{6.5}$$

has the property that either $\log B_t(z)$ is subharmonic in t or identically equal to $-\infty$. In the present case, we take

$$\mathcal{D} := \{\Psi(z) - \Re(t) < 0\} \subset Y \times \mathbb{C}$$

and $\psi(z, t) = \psi_0(z)$. Note that $\mathcal{D}$ is a pseudoconvex domain in $\mathbb{C}^{n+1}$. Indeed, this follows from Lemma 2.3, using that $\Psi(z) - \Re(t)$ is psh in $Y \times \mathbb{C}$ and $Y \times \mathbb{C}$ is pseudoconvex (since Y is). Now, by assumption, the group G acts holomorphically on D_t. In particular, if dG denotes a G-invariant measure on G (i.e., Haar measures) and f is holomorphic on D_t then the function

$$f_G(z) := \int_{g \in G} f(g \cdot z) dG$$

is holomorphic and G-invariant. Hence, replacing f with f_G and using the "triangle inequality" the sup in formula 6.5 may as well be restricted to all G-invariant holomorphic functions f. But, by assumption, any such functions is constant, and hence, we may as well take f to be identically equal to 1. But this means that $B_t(z) = 1/\int_{Y \cap \{\Psi(z) < t\}} e^{-\psi_0} d\lambda$. The theorem thus follows from the main result of [11], recalled above (also using that if $\phi(t)$ is subharmonic in t and only depends on $\Re(t)$ then $\phi(t)$ is convex wrt $t \in \mathbb{R}$). □

Remark 6.6. (Brunn–Minkowski inequality) This proposition can be viewed as a generalization of the classical fact that the logarithm of the volume $\mu_0(\{\phi \leq t\})$ is concave if ϕ is a convex function on $\mathbb{R}^n$ and μ_0 is a log concave measures, i.e., $\mu_0 = 1_C e^{-\phi_0(x)} d\lambda$ for $C \subset \mathbb{R}^n$ a convex body and ϕ_0 a convex function. This is a consequence of the Brunn–Minkowski inequality, but it also follows from the previous proposition by considering the map $L(z) = (\log|z_1|, ..., \log|z_n|)$ from $(\mathbb{C} - \{0\})^n$ onto $\mathbb{R}^n$ which has the property that $\phi(x)$ is convex iff $\psi := L^*\phi$ is psh, and hence, C is convex iff $Y := L^{-1}(C)$ is pseudoconvex in $\mathbb{C}^n$ (using that $L^*\phi$ is bounded from above and this extends to a psh function on $\mathbb{C}^n$). The classical fact in question then follows by taking G as the n-dimensional compact torus, acting on $\mathbb{C}^n$ in the standard way (and thus preserving the fibers of the map L).

It should be stressed that plurisubharmonicity alone is not enough to ensure the concavity in the previous proposition, as illustrated by the case when Y is the unit disk in $\mathbb{C}$ and $\Psi(z)$ is the Green function for the Laplacian with a pole at $w \in Y$, for a given nonzero w in the interior of Y, i.e., $\Psi(z) = \log|(z-w)/(1-\bar{w}z)|$ Indeed, as shown in the proof of [10, Thm 2.3] the concavity in question is then equivalent to the subharmonicity of the Schwartz symmetrization of $\Psi(z)$, which only holds when a is zero, i.e., when Ψ is S^1-invariant, as pointed out in the introduction of [10].

We next apply the previous proposition to the case when the N-particle Hamiltonian $H^{(N)}$ on X^N comes from a pair interaction W and exterior potential V such that $-W$ and $-V$ and also ϕ and Ψ_0 satisfy the Main Assumptions (but there is no need to assume that $W(x, y)$ is symmetric or that the indices i, j range over all of $\{1, 2, ..., N\}$):

Theorem 6.7. *If the Main Assumptions hold and $H^{(N)}$ is the function on X^N defined by*

$$H^{(N)}(x_1, ..., x_N) := \sum_{(i,j)\in\mathcal{I}} a_{ij} W(x_i, x_j) + \sum_{j\in\mathcal{J}} b_i V(x_j)$$

for some subsets $\mathcal{I}$ of $\{1, .., N\}^2$ and $\mathcal{J}$ of $\{1, ..., N]$ and nonnegative constants a_{ij} and b_i. Then

$$S_+^{(N)}(e) := \log \mu_0^{\otimes N}\left\{H^{(N)} > e\right\}$$

is concave in e and finite when $e > \sup_{X^N} H^{(N)}$. In particular, $S_+^{(N)}(e)$ is concave when $H^{(N)}$ is the mean field Hamiltonian 1.1.

Proof. It is a standard fact that the closure of the orbits of the vector field $\mathcal{V}_a$ 2.5 coincides with the orbits of a compact torus G acting holomorphically on $\mathbb{C}^n$. The assumptions imply that ϕ, V, W and μ_0 are invariant under the action of G (using the diagonal action on $X \times X$). By assumption, X admits a continuous psh exhaustion function ρ. By the construction in Lemma 2.3, ρ may as well be assumed to be G-invariant. Since the maximum of a finite number of psh functions is still psh, the function ρ_N on X^N defined by

$$\rho(z_1, ..., z_N) := \max_{i=1,...,N} \rho(z_i)$$

is psh and G-invariant wrt the diagonal action of G on X^N and thus defines a G-invariant continuous psh exhaustion function of X^N. The corollary will thus follow from the previous theorem applied to X^N, $\Psi_N := -H^{(N)}$ and the measure $\mu_0^{\otimes N}$ if X^N, once the assumptions on G have been verified. To this end, consider the holomorphic action of the additive group $\mathbb{C}$ on $\mathbb{C}^n$ defined as follows: given $\tau \in \mathbb{C}$ and $z \in \mathbb{C}^n$

$$\tau \circledcirc z := (e^{a_1\tau} z_1, ..., a^{a_n\tau} z_n). \tag{6.6}$$

Note that, for z fixed, $\tau \circledcirc z \to 0$ as the real part $\Re\tau \to -\infty$. Moreover, if ϕ is a psh function on $\mathbb{C}^n$ then

$$\phi(\tau \circledcirc z) \leq \phi(z) \text{ if } \Re\tau \leq 0 \tag{6.7}$$

To see this, first observe that since ϕ is psh and the orbits of the $\mathbb{C}$-action define holomorphic curves the function $\phi(\tau) := \phi(\tau \circledcirc z)$ is subharmonic for a fixed z. Moreover, since $\phi \in PSH_{\boldsymbol{a}}$ the function $\phi(\tau)$ is independent of the imaginary part $\Im\tau$, and hence, $\phi(\tau)$ is convex wrt the real part t of $\Re\tau$. Since ϕ is bounded from above close to the origin it follows that there exists a constant C such that $\phi(\tau) \leq C$ as $\Re\tau \to -\infty$. But then the convexity of $\phi(t)$ implies that $d\phi(t)/dt \to 0$ as $\tau \to -\infty$ in $\mathbb{R}$. Hence, by convexity, $\phi(\tau)$ is increasing in the real part of τ, proving the inequality. As a consequence, X is G-invariant, connected and the origin 0 is contained in the interior of X (using that the action by $\mathbb{C}$ is locally free). Moreover, the same thing goes for the sublevel sets $\{\Psi_N(z_1, ..., z_N) < t\}$. Indeed, by the previous argument $\Psi_N(\tau \circledcirc z_N, ..., \tau \circledcirc z_N)$ is increasing wrt the real part of τ. In particular, the minimum of Ψ is attained at the origin in X^N, which implies that 0 is an interior point of $\{\Psi_N(z_1, ..., z_N) < t\}$, as long as $t > \inf_{X^N} \Psi_N$. Thus, if f is a holomorphic function on $\{\Psi_N(z_1, ..., z_N) < t\}$, then in order to verify that f is constant on $\{\Psi_N(z_1, ..., z_N) < t\}$ it is enough to verify that its Taylor expansion at the origin 0 in $\mathbb{C}^{nN}$ is a constant. To simplify the notation, we will prove this when $N = 1$ (but the general case if the same up to a change of notation). Using multinomial notation the action of $\nu_{\boldsymbol{a}}$ on $f(z)$ close to the origin in $\mathbb{C}^n$ gives, by Taylor expansion of f,

$$\nu_{\boldsymbol{a}}(f) = \nu_{\boldsymbol{a}}\left(\sum_{\alpha_i \geq 0} c_{\boldsymbol{\alpha}} z_1^{\alpha_1} \cdots z_n^{\alpha_n}\right) = \sum_{\alpha_i \geq 0} i\boldsymbol{a} \cdot \boldsymbol{\alpha} c_{\boldsymbol{\alpha}} z_1^{\alpha_1} \cdots z_n^{\alpha_n}.$$

Since $a_i > 0$ the scalar product $\boldsymbol{a} \cdot \boldsymbol{\alpha}$ is nonvanishing for $\boldsymbol{\alpha} \neq \mathbf{0}$.Hence, $\nu_{\boldsymbol{a}} f = 0$ can only hold if the Taylor coefficients $c_{\boldsymbol{\alpha}}$ vanish for $\boldsymbol{\alpha} \neq \mathbf{0}$. Since X is connected it follows that f is identically constant (by the identity principle for holomorphic functions).

As for the finiteness of $S_N^+(e)$, for $e > \sup_{X^N} H^{(N)}$, it follows directly from the fact that any psh function is usc, hence the subset where $H^{(N)} > e$ is open. □

When $H^{(N)}$ is replaced by the "attractive" Hamiltonian $-H^{(N)}$ the previous theorem also shows that $S_-^{(N)}(e)$ (formula 1.9) is concave. The following simple example illustrates the relevance of the plurisubharmonicity assumption in the previous theorem. Consider the "attractive" Hamiltonian obtained by taking $W = 0$ in formula

1.1 and assume that V is S^1-invariant. Then V is psh iff $V = \phi(\log|z|)$ for a convex increasing function ϕ on $\mathbb{R}$. Assuming that $\phi(x)$ is *strictly* increasing and $N = 1$ we get

$$\mu_0^{\otimes N}\left\{H^{(N)} \leq e\right\} = C_n e^{2nf(e)}, \tag{6.8}$$

where $f(e)$ is the function defined, on the image of ϕ, as the inverse of $\phi(x)$ and C_{nN} is the volume of the unit-ball in $\mathbb{R}^{2n}$. Hence, the logarithm of $\mu_0^{\otimes N}\left\{H^{(N)} \leq e\right\}$ is concave iff f is concave iff its inverse ϕ is convex iff V is psh. In the case when $N \geq 1$ an illustrative class of "attractive" Hamiltonians is given by the case when V is a power law, $V = |z|^\alpha$ for $\alpha > 0$ (i.e., $\phi(x)$ is the convex function $= e^{\alpha x}$). Then a simple scaling argument reveals that the volume $\mu_0^{\otimes N}\left\{H^{(N)} \leq e\right\}$ is of the form 6.8 for $f(e) = \log e$ when n is replaced by nN/α, if $e \geq 0$. Hence, the logarithm of $\mu_0^{\otimes N}\left\{H^{(N)} \leq e\right\}$ is indeed concave. Note that in this example the logarithm of the surface area of $\{H^{(N)} = e\}$, i.e., of the derivative of $\mu_0^{\otimes N}\left\{H^{(N)} \leq e\right\}$, is *not* concave unless N is taken sufficiently large; $N \geq \alpha/2n$. (Otherwise, it is convex.)

Remark 6.8. In the case of the mean field Hamiltonian $H^{(N)}$, the concavity of $S_+^{(N)}(e)$ holds more generally if the assumptions on W and V are replaced by the weaker assumption that the negative of $W(x, y) + \frac{N}{N-1}(V(x) + V(y))$ is in $PSH_{\boldsymbol{a},\boldsymbol{a}}(X \times X)$. Indeed, rewriting

$$H^{(N)}(x_1, ..., x_N) = \frac{1}{2}\frac{1}{N}\sum_{i \neq j \leq N}\left(W(x_i, x_j) + \frac{N}{N-1}\left(V(x_i) + V(x_j)\right)\right),$$

we can then apply the previous theorem to the mean field Hamiltonian corresponding to the pair interaction $W(x, y) + \frac{N}{N-1}(V(x) + V(y))$.

6.2.1. Incorporating Constraints Prop 6.5 may be generalized by replacing Ψ with a finite number of functions $\psi_1, ..., \psi_r$ on Y satisfying the same assumptions as Ψ and replacing the sublevel $\{\Psi(< t\}$ with the intersection of the sublevel sets $\{\psi_1 < t_1\}, ..., \{\psi_r < t_r\}$ for a given $\boldsymbol{t} = (t_1, ..., t_r) \in \mathbb{R}^r$. Then the logarithm of the corresponding volume defines a concave function of $\boldsymbol{t} \in \mathbb{R}^r$. Indeed, one simply replaced the domain $\mathcal{D}$ in the proof with the intersection of the pseudo-convex domains $\{\psi_1(z) - \Re(t_1) < 0\}$ in $\mathbb{C}^n \times \mathbb{C}^r$. Since the intersection of pseudo-convex domains is pseudo-convex, the main result in [11] then implies that the corresponding function $\log B(\boldsymbol{t})$ is a psh function of $\boldsymbol{t} \in \mathbb{C}^r$, and thus, by translational invariance in the imaginary arguments, it defines a convex function on $\mathbb{R}^r$. As a consequence, if one assumes given $\psi_1, ..., \psi_r$ as above then Theorem 5.4 may be generalized to the statement that the "constrained microscopic entropy"

$$\log \mu_0^{\otimes N}\left\{H^{(N)}(z_1, ..., z_N) > e, \sum_{i=1}^N \psi_1(z_i) \leq l_1, ..., \sum_{i=1}^N \psi_r(z_i) \leq l_r\right\}$$

is a concave function of $(e, l_1, ..., l_r) \in \mathbb{R}^{1+r}$. In particular, when $X \subset \mathbb{C}^n$ this applies to $\psi_i(z_1, .., z_n) = \lambda_i|z_i|^2$ for given positive numbers $\lambda_1, ..., \lambda_n$, as in the Gaussian case discussed in Sect. 3.3. Anyhow, in this paper we will, for simplicity, stick to the non-constrained setup. On the other hand, as pointed out in Sect. 3.3, the constraints may be incorporated in the prior measure.

6.3. Concavity of the Entropy $S(e)$ when $e_0 \leq e$

Now assume that $X \Subset \mathbb{R}^{2n}$ that we identify with $\mathbb{C}^n$, as usual.

Theorem 6.9. *If the Main Assumptions holds, then the entropy $S(e)$ is a decreasing concave continuous function on $[e_0, e_{max}[$*

In order to prove this, we first assume that X is compact and invoke the following

Proposition 6.10. *[29] Assume that W and V are continuous and X is compact. Then*

$$\lim_{N\to\infty} S_+^{(N)}(e) = S_+(e) := \sup_{E(\mu)\geq e} S(\mu)$$

Proof. Consider the open interval $\Delta := \{t > e\}$ in $\mathbb{R}$. By [29, Thm 2.1], the limsup and liminf of $S_+^{(N)}(e)$ is equal to the sup of $S(\mu)$ over all $\mu \in \mathcal{P}(X)$ such that $E(\mu) \in \overline{\Delta}$ and $E(\mu) \in \Delta$, respectively. But by Lemma 3.2, both these quantities are equal to $S_+(e)$. □

When W and V are continuous and X is compact Theorem6.7 thus shows that $S_+(e)$ is a limit of concave functions on $\mathbb{R}$ and thus concave on $\mathbb{R}$. Next, we invoke the monotonicity properties shown in Lemma 3.2 which show that

$$S_+(e) = \max\{e_0, S(e)\}$$

and hence, $\max\{e_0, S(e)\}$is is concave. But by Lemmas 3.6, 6.3 $S(e)$ is finite for any $e \in]e_{min}, e_{max}[$. Hence, $\max\{e_0, S(e)\}$ is concave and finite on $]e_{min}, e_{max}[$ and thus concave and continuous on $]e_{min}, e_{max}[$. This proves the theorem in the case when W and V are continuous and X is compact.

6.3.1. Conclusion of the Proof of Theorem 6.9 Still assuming that X is compact we will next show that $S(e)$ is decreasing, concave and continuous when $e \in [e_0, e_{max}[$, i.e., when

$$E(\mu_0) \leq e < \sup_{\mathcal{P}(X)} E(\mu) \tag{6.9}$$

We will proceed by an approximation argument and exploit that $S(e) > -\infty$ (Prop 6.4). Take a sequence $W_\delta(x, y)$ of continuous pair interactions increasing to W satisfying the Main Assumptions. For example, W_δ may be defined as a convolution of W with a compactly supported smooth density ρ_δ. First observe that since, by assumption, $E(\mu_0) < e$, we get $E_\delta(\mu_0) < e$ for δ sufficiently small (by the monotone convergence theorem). Thus, by the concavity and continuity of $S_{+,\delta}(e)$ on $\mathbb{R}$ established in the previous section, we just have to verify that

$$\lim_{\delta\to 0} S_\delta(e) = S(e) \tag{6.10}$$

for any fixed e satisfying the inequalities in formula 6.9. To this end, first note that

$$S_\delta(e) \leq S(e). \tag{6.11}$$

Indeed, by Lemma 3.2, it is enough to prove the corresponding inequality for the *upper* entropies, where it follows directly from the assumption that $E_\delta(\mu) \leq E_0(\mu)$. Now fix a candidate μ for the sup defining $S_0(e)$ and set $e_\delta := E_\delta(\mu)$. Then

$$S_0(\mu) \leq S_\delta(e_\delta) = S(\mu_\delta),$$

where μ_δ realizes the sup defining $S_\delta(e_\delta)$. Moreover, fixing a positive number ϵ we have

$$e_\delta \geq e - \epsilon$$

for δ sufficiently small ($\delta < \delta_\epsilon$). Hence, since S_δ is decreasing when $e > E_\delta(\mu_0)$ (by Lemma 3.2), we get

$$S_0(\mu) \leq S_\delta(e - \epsilon)$$

for $\delta < \delta_\epsilon$. Using that $S_\delta(e)$ is concave, we thus deduce that

$$S_\delta(e_\delta) \leq S_\delta(e) + \epsilon |\frac{d}{de} S_\delta(e)|.$$

Combining the latter inequality with the inequality 6.11 reveals that all that remains, in order to prove the convergence 6.10, is to verify that

$$|\frac{d}{de} S_\delta(e)| \leq C \tag{6.12}$$

as $\delta \to 0$. To this end first observe that, using again that $S_\delta(e)$ is decreasing and concave yields for any fixed $e' > e$

$$|\frac{d}{de} S_\delta(e)| = -\frac{d}{de} S_\delta(e) \leq \frac{S_\delta(e) - S_\delta(e')}{e' - e} \leq \frac{-S_\delta(e')}{e' - e}.$$

In particular, if e' is a fixed number satisfying $e < e' < \sup_{\mathcal{P}(X)} E(\mu)$ we get (by Lemma 3.2) that

$$S_\delta(e') \geq S(\mu')$$

for any $\mu' \in \mathcal{P}(X)$ such that $E_\delta(\mu') \geq e'$. Now, by Lemmas 3.6, 6.3 μ' can be chosen, independently of δ, so that $E(\mu') \geq e' + \epsilon$ and $S(\mu') > -\infty$. We then get, for any δ sufficiently small, that $E_\delta(\mu') \geq e'$ (by the monotone convergence theorem), and thus, the uniform bound 6.12 follows.

This concludes the proof of Theorem 6.9 in the case when X is compact. In the general case we fix $R > 0$ and denote by X_R the intersection of X with a ball B_R of radius R centered at the origin. Then X_R is also pseudoconvex (as follows from Lemma2.3 applied to $\phi(z) = |z|^2 - R$). Thus, as shown in the previous section, the entropy $S_R(e)$ associated to the restrictions to B_R of W, V and μ_0 is concave in e for $e > E(\mu_0)$. Hence, all that remains is to verify that

$$\lim_{R \to \infty} S_R(e) = S(e) \tag{6.13}$$

for any fixed e satisfying the inequalities 6.9. To this end, first note that, since $X_R \subset X$, it follows immediately that $S_R(e) \leq S(e)$. Now assume that $e > E(\mu_0)$ and fix a candidate μ for the sup defining $S(e)$. Set

$$\mu_R := 1_{B_R}\mu/\mu(B_R).$$

Then

$$S(\mu) \leq S_R(e_R), \quad e_R := E(\mu_R).$$

By the monotone convergence theorem, $E(\mu_R) \to E(\mu)$. Hence, using that $S_R(e)$ is decreasing and concave for R sufficiently large (by the previous step) we can proceed essentially as when approximating W with W_δ above, to get

$$S(\mu) \leq \limsup_{R\to\infty} S_R(e),$$

which concludes the proof of the convergence 6.13 and thus the concavity in Theorem 6.9.

7. Global Concavity of $S(e)$ and Examples

Recall that, in classical terminology, a symmetric function $W(x, y)$ is a *weakly positive definite kernel,* i.e., that for any positive integer N

$$\sum_{i,j\leq N} W(x_i, x_j)a_i a_j \geq 0 \ \ \forall (a_i) \in \mathbb{R}^N : \sum_{i=1}^N a_i = 0$$

If the first inequality holds for *any* sequence $(a_i)_{i=1}^N$, then $W(x, y)$ is called a *positive definite kernel.* [2]

Now assume that $W(x, y)$ is weakly positive definite and satisfies the Main Assumptions in a neighborhood of $X \times X$. Then W can be expressed as increasing limit of *continuous* (and even smooth) such functions $W_\delta(x, y)$. Indeed, if ρ is a smooth compactly supported probability density on $\mathbb{R}^{2n} \times \mathbb{R}^{2n}$ we can take

$$W_\delta := (W * \rho_\delta) := \int W(\cdot + a, \cdot + b)\rho_\delta(a, b)d\lambda(a)d\lambda(b), \quad \rho_\delta(x, y) = \rho(\delta^{-1}x, \delta^{-1}x)\delta^{4n}. \tag{7.1}$$

Since $-W$ is psh, W_δ indeed increases to W. Moreover, since $W(\cdot + a, \cdot + b)$ is weakly positive definite and satisfies the Main assumptions for any (a, b) so does W_δ.

Theorem 7.1. *If the Main Assumptions hold and moreover $W(x, y)$ is assumed weakly positive definite, then the entropy $S(e)$ is globally concave, and hence, thermodynamic equivalence of ensembles holds for any $e \in]e_{min}, e_{max}[$.*

[2] $W(x, y)$ is a weakly positive kernel iff $-W(x, y)$ is a *negative definite kernel* in the terminology of [5].

Proof. By Theorem 6.9 $S(e)$ is concave and continuous on $[e_0, e_{max}[$. Next, recall the classical fact that a weakly positive definite kernel defines a convex functional $E_W(\mu)$ on $\mathcal{P}(X)$ (and vice versa). Hence, by Prop 5.3 $S(e)$ is concave and continuous on $]e_{min}, e_0]$ when X is compact and W is continuous. The theorem thus follows, in the compact and continuous case, from the second and third point in Prop 3.7. Next assume that X is still compact and define W_δ to be a regularization as in formula 7.1. Then the corresponding entropy $S_\delta(e)$ is concave on $]e_{min,\delta}, e_{max,\delta}]$. Moreover, by the approximation argument used in the proof of Theorem 6.9 and the finiteness of $S(e)$ the function $S_\delta(e)$ converge point-wise to $S(e)$ on $]e_{min}, e_0[$. Thus, $S(e)$ is also concave on $]e_{min}, e_0[$. Finally, the general non-compact case is deduced from the compact case using again the approximation arguments in the proof of Theorem 6.9 and the finiteness of $S(e)$. □

Recall that, by Bochner's classical theorem, a translationally invariant kernel $W(x, y) = \mathcal{W}(x - y)$ is positive definite iff the function $\mathcal{W}$ on $\mathbb{R}^d$ is the Fourier transform of a (positive) measure on $\mathbb{R}^d$. In the case of translationally and rotationally invariant kernels, the following classical result holds [5]:

Lemma 7.2. *(Bernstein+Schoenberg). Let $w(r)$ be a continuous function on $[0, \infty[$ which is smooth on $]0, \infty[$. Then $W(x, y) := w(|x-y|)$ is a positive definite kernel iff $f(r) := w(r^{1/2})$ is* completely monotone*, i.e., $(-1)^m \partial^m f(r)/dr^m \geq 0$ for all nonnegative integers m.*

The previous lemma implies that if w is nonnegative on $[0, \infty[$ (but possibly equal to ∞ at $r = 0$) and $w(r^{1/2})$ is completely monotone for $r > 0$, then $w(|x-y|)$ is still positive definite. Indeed, one can apply the previous lemma to

$$w_\epsilon(r) := w\left((r^2 + \epsilon)^{1/2}\right)$$

and then let $\epsilon \to 0$.

Corollary 7.3. *Under the Homogeneous Assumptions together with the assumption that $w(r^{1/2})$ is completely monotone for $r > 0$, the entropy $S(e)$ is globally concave and thermodynamic equivalence of ensembles holds at all energies.*

It should be pointed out that assumptions in the previous corollary are preserved if w is replaced by $w_\epsilon(r)$ above (using that $\log(|z|^2 + \epsilon)$ is psh) and similarly for v and ψ_0. This gives a convenient explicit regularization procedure preserving the property that $S(e)$ is globally concave.

7.1. Examples where $S(e)$ is Globally Concave

We next provide some examples where Theorem 7.1 applies, and thus, $S(e)$ is globally concave. More examples may, for example, be obtained by taking convolutions (as in formula 7.1). Note also that if the entropy $S_{W,V}(e)$ corresponding to the interactions W and V is globally concave, then so is $S_{-W,-V}(e)$, since $S_{-W,-V}(e) = S_{W,V}(-e)$. In this way, one may thus go from a situation of repulsive interactions to attractive ones.

Theorem 7.1 applies to the case when $W(x, y) = -\log|z - w|$ when, for example, μ_0 is Lebesgue measure on for example a ball in $\mathbb{R}^{2n}$ or a centered (possibly nonstandard) Gaussian measure in $\mathbb{R}^{2n}$ (as in formula 1.12). Indeed, then W satisfies the Homogeneous Assumptions and the positive definiteness follows, for example, from the fact that W is the Green kernel on $\mathbb{R}^{2n}$ of the n th power of the Laplacian, which is positive definite as a formally self-adjoint operator. More generally, the ball may be replaced with any domain X satisfying the Main Assumptions, for example,

$$X = \{z \in \mathbb{R}^{2n} : \sum_{i=1}^{r} |P_i(z)|^{\alpha_i} \leq 1\},$$

for a quasi-homogeneous polynomials $P_1, ..., P_r$ and $\alpha_i > 0$ (see Example 2.5).

Theorem 7.1 also applies to the continuous repulsive *power laws* with exponent in $]0, 2]$

$$W(x, y) = -|x - y|^a, \ \ a \in]0, 2],$$

as well as to

$$W(x, y) = e^{-\alpha|x-y|^a}, \ \ a \in]0, 2]$$

when X is taken to be a disk centered at the origin with radius at most $(1/2\alpha)^{1/a}$. Indeed, a direct computation reveals that $w(r)$ satisfies the Homogeneous Assumptions for any $a, \alpha > 0$ (by a scaling it is enough to verify the case when $a = \alpha = 1$) Moreover, by [5, Cor 3.3] (and its proof) the kernels in question are weakly positive definite when $a \in]0, 2]$. Note that in the case of the repulsive logarithmic interaction, as well as for repulsive power laws with $a \in]0, 2[$, Prop 4.8 ensures the existence of maximum entropy measures μ^e, when μ_0 is a centered Gaussian measure (by taking $\psi_0 = |x|^2$).

7.1.1. The Point Vortex Model Consider the point vortex model (for vortices with identical circulations) on a domain X in $\mathbb{R}^2$. In the case when $X = \mathbb{R}^2$,

$$W(x, y) = -\log|x - y|, \quad V(x) = 0$$

(with our normalizations). As discussed in the previous section, $S_+^{(N)}(e)$ and $S(e)$ are both globally concave (and thermodynamic equivalence of ensemble holds) if μ_0 is taken to be a centered Gaussian measure. As indicated in [15, Section 5], the concavity of $S(e)$ also follows from the results in [15], using completely different techniques (see also[42] where the concavity of the corresponding multivariable entropy $S(e, l)$, discussed in Sect. 3.3, is shown). But, as discussed in the introduction of the paper, the main point of the present technique is that it also applies to regularizations of W.

In the case of when X is a compact domain with smooth boundary, $W(x, y)$ is defined as the negative of Green function $G_X(x, y)$ for the Laplacian on X with Dirichlet boundary conditions and $V(x) = \gamma(x)/N$ where γ is the restriction to

the diagonal of $G_X(x, y) + \log|x - y|$ [14,15,46]. In particular, when X is the unit disk

$$W(z, w) = -\log \frac{|z - w|}{|1 - z\bar{w}|}, \quad V(x) = \frac{1}{N} \log|1 - |z|^2| \tag{7.2}$$

In this case, Theorem 6.7 implies that $S_+^{(N)}(e)$ is globally concave when $N \leq 3$, as follows from combining Remark 6.8with the following lemma, proved in appendix.

Lemma 7.4. *Denote by D the interior of the unit disk in $\mathbb{C}$ and set*

$$\psi(z, w) := \log\left(|z - w|^2/|1 - z\bar{w}|^2\right), \quad \phi(z) = -\log\left(\left|1 - |z|^2\right|^2\right)$$

The function $\psi(z, w) + \lambda(\phi(z) + \phi(w))$ is psh in $D \times D$ iff $\lambda \geq 1/2$.

We leave open the question whether $S_+^{(N)}(e)$ is concave also when $N > 3$. As for $S(e)$, it was shown to be concave in [15], using a completely different method. In the case when a rotationally invariant exterior potential V_e is added to $V(x)$ in formula 7.2, the previous lemma shows that $S_+^{(N)}(e)$ is concave for any N (and hence also $S(e)$) if $-\partial\bar{\partial}V_e \geq \partial\bar{\partial}\phi/2$ in D, i.e., if the Laplacian of V_e is sufficiently negative:

$$\frac{1}{4}\Delta V_e(z) \leq -\frac{1}{(1 - |z|^2)^2}.$$

This should be contrasted with the fact that the global concavity of $S(e)$ may fail if the Laplacian is positive, e.g., in the case when $V_e(z) = \omega|z|^2$, for $\omega > 0$, studied in [54] and [15, Lemma 8.2].

7.1.2. Insulated Plasmas and Self-Gravitating Matter in 2D The point vortex model on a compact domain X is physically equivalent to a one-component Coulomb plasma if inertial effects are ignored (i.e., the limit of infinite damping is considered) and the boundary of X is assumed to be conductive [54]. On the other hand, the case when the boundary of X is non-conducting, i.e., X is insulated, corresponds to the mean field Hamiltonian on X with Coulomb pair interaction $-\log|x - y|$ (and $V \equiv 0$) [32]. In this case the Main Assumptions apply when μ_0 is the uniform measure on the X unit disk X, as discussed in the beginning of Sect. 7.1. More generally, the Main assumptions apply when the exterior potential V is radial and $\Delta V \leq 0$, i.e., V is the potential induced by a distribution of fixed particles with the same charge as the plasma. Switching the sign of the Coulomb interaction yields a system of self-gravitating matter, studied in [1] with inertial effects included.

8. Critical Inverse Temperatures and Existence of Maximum Entropy Measures

In the Very General Setup, the *macroscopic inverse temperatures* is defined by

$$\beta_c := \inf\left\{\beta \in \mathbb{R} : \inf_{\mu} F_\beta(\mu) > -\infty\right\}. \tag{8.1}$$

The *microscopic inverse temperature* $\beta_{c,N}$ is, in the General Setup, defined by

$$\beta_{c,N} := \left\{ \beta \in \mathbb{R} : \; Z_{N,\beta} := \int_{X^N} e^{-\beta H^{(N)}} (e^{-\Psi_0} dx)^{\otimes N} < \infty \right\},$$

and, respectively, where $H^{(N)}$ denotes the mean field Hamiltonian 1.1 corresponding to W and V.

8.1. Dual Expressions for the Critical Inverse Temperatures

We start with the following dual "slope formula" for $\beta_{N,c}$, under the Main Assumptions, which also shows that $\beta_{N,c} < 0$.

Corollary 8.1. *Under the same assumptions as in Prop 6.5, the following holds if μ_0 has finite mass on Y and Ψ is not identically constant:*

$$\begin{aligned} c_{(Y,\mu_0)}(\Psi) &:= -\inf \left\{ \beta \in]-\infty, 0] : \int_Y e^{\beta\Psi} \mu_0 < \infty \right\} \\ &= \lim_{e \to -\inf_Y \Psi} \frac{d}{de} \log \left(\mu_0 \{ \Psi < -e \} \right), \end{aligned}$$

using either right or left derivatives in the right-hand side. As a consequence, the set of all negative β such that $\int_Y e^{\beta\Psi}\mu_0 < \infty$ is open. In particular, under the Main Assumptions

$$\beta_{N,c} = \lim_{e \to \sup_{X^N} E_N} \frac{dS^{(N)}(e)}{de} < 0, \quad Z_{N,\beta_{N_c}} = \infty$$

Proof. By Prop 6.5 (and Theorem6.7) the function

$$\phi(t) := -\log \mu(t), \quad \mu(t) := (\mu_0 \{ \Psi < -t \})$$

is convex wrt $t \in \mathbb{R}$. Consider first the case when $t_0 := \inf_Y \Psi > -\infty$. Then, trivially, $\beta_c = -\infty$. Moreover, $\phi(t)$ is convex and finite for $t > t_0$ and $\phi(t) \to \infty$ as t decreases to t_0. But this forces $d\phi(t)/dt \to -\infty$ as t decreases to t_0. Indeed, by the convexity of ϕ the limit of $d\phi(t)/dt$ decreases to $M_0 \in [-\infty, \infty[$ as t decreases to t_0. Assume, to get a contradiction, that $M_0 > -\infty$. Then, fixing $t_1 > t_0$ gives $\phi(t) \leq \phi(t_1) + |M||t_1 - t_0| < \infty$ as $t \to t_0$, which contradicts that $\phi(t) \to \infty$ as t decreases to t_0.

Next, assume that $\inf_Y \Psi = -\infty$. Since $\beta \leq 0$ we have

$$\int_Y e^{\beta\Psi}\mu_0 \leq \int_{\{\Psi<0\}} e^{\beta\Psi}\mu_0 + \mu_0(Y),$$

where, by assumption, the second term is finite. Pushing forward the measure μ_0 on Y to $\mathbb{R}$ under the map $z \mapsto \Psi(z)$ gives

$$\int_{\{\Psi<0\}} e^{\beta\Psi}\mu_0 = \int_{-\infty}^0 e^{\beta t} \frac{dV(t)}{dt} dt = -\beta \mathcal{Z}(\beta) + V(0), \; \mathcal{Z}(\beta) := \int_{-\infty}^0 e^{\beta t} V(t) dt,$$

where the second equality follows from integrating by parts. We may then conclude the proof of the first formula in the corollary by expressing

$$\mathcal{Z}(\beta) := \int_{-\infty}^{0} e^{\beta t - \phi(t)} dt$$

and applying Lemma 8.3 to the convex function $\Phi = \beta t - \phi(t)$, which implies that

$$\int_Y e^{\beta\Psi}\mu_0 < \infty \iff -\beta < \lim_{t\to\infty} \frac{d\phi}{dt}, \tag{8.2}$$

concluding the proof of formula in question. To prove that $\beta_{N,c} < 0$, note that $\phi(t) \to \infty$ as $t \to -\infty$ and $\phi(t) \to 0$ as $t \to \infty$. Since $\phi(t)$ is convex if follows that, using either left or right derivatives, $\lim_{t\to-\infty} d\phi(t)/dt \leq 0$ and $\lim_{t\to-\infty} d\phi(t)/dt = 0$. But if $\beta_{N,c} = 0$, then, by the previous step, $\lim_{t\to-\infty} d\phi(t)/dt = 0$, and hence, by convexity, $\phi(t)$ is constant. But this can only happen if Ψ is constant, which is excluded by the assumptions. Thus, $\beta_{N,c} < 0$, as desired. Finally, to prove the last openness statement we just have to verify that if $\int_Y e^{\beta\Psi}\mu_0 < \infty$, then there exists $\delta > 0$ such that $\int_Y e^{(\beta-\delta)\Psi}\mu_0 < \infty$. But this follows directly from the strict inequality in the right-hand side of formula 8.2. □

Remark 8.2. In the case when Y is compact and $\Psi_0 = 0$ (or, equivalently, bounded), the number $c_Y(\Psi)$ is called the *integrability threshold* of Ψ on Y (or the *complex singularity exponent*) in the complex geometry literature (whose inverse is the *Arnold multiplicity*) [24]. It follows from Skoda's local integrability inequality that $c_Y(\Psi) > 0$ for any function which is psh on a neighborhood of Y and not identically $-\infty$. Moreover, $\int_Y e^{\beta\Psi} d\lambda = \infty$ in the critical case $\beta = -c_Y(\Psi)$, by the resolution of the openness conjecture in [12] (see also [34] for the resolution of the strong openness conjecture). The proof above yields a simplification of the proof in [12] under the symmetry assumption that $\Psi \in PSH(Y)_{\boldsymbol{a}}$ (anyhow, just like [12], it is based on [11]).

In the above proof, the following elementary fact was used:

Lemma 8.3. *Let $\Phi(t)$ be a convex function on $]-\infty, 0[$ such that $\Phi(t)$ is bounded as $t \to 0$. Then*

$$\int_{-\infty}^{0} e^{-\Phi(t)} dt < \infty$$

iff $\lim_{t\to-\infty} d\Phi(t)/dt < 0$, *using either left or right derivatives.*

Corollary 8.4. *Consider the Main Assumptions and assume also that $S(e_{max}) = -\infty$, if $e_{max} < \infty$. Then, as e increases strictly toward e_{max}*

$$\beta_c = \lim_{e\to e_{max}} \frac{dS(e\pm)}{de}, \tag{8.3}$$

where $dS(e\pm)/ds$ denotes either the left or the right derivative of the concave function $S(e_\pm)$.

Proof. By Theorem 6.9, $S(e)$ is concave and continuous on $[e_0, e_{max}[$. Denote by $\tilde{F}$ the usc concave function defined as F when $\beta \leq 0$ and as $-\infty$ when $\beta > 0$. By Prop 3.7 $S = (\tilde{F})^*$ on $[e_0, e_{max}[$. Set $g := (\tilde{F})^*$. Thus, g is constant for $e \leq e_0$ and on $[e_0, e_{max}[$ it coincides with $S(e)$ (by 3.7). Moreover, $g^* = \tilde{F}$, and hence, $\overline{\{g^* < \infty\}} = [\beta_c, \infty[$. Thus, by formula 2.3,

$$[\beta_c, \infty[= \overline{\partial g(\{g > -\infty\})} = \overline{\partial S(]e_0, e_{max}[)},$$

which proves formula 8.3, using that $dS(e+)/ds \leq dS(e-)/ds$ and $dS(e+)/ds$ and $dS(e-)/ds$ are both decreasing (by concavity). □

8.2. Concrete Expressions in the Homogeneous Case

It seems natural to expect that, under rather general assumptions, $\beta_{N,c} \to \beta_c$ as $N \to \infty$. Here we will show that this is the case under the Homogeneous Assumptions; in fact, $\beta_{N,c} = \beta_c$ for any N. The starting point is the following essentially well-known consequence of the Gibbs variational principle (compare [6,14,41]):

Lemma 8.5. *Let $H^{(N)}$ be a mean field Hamiltonian of the form 1.1. Then*

$$Z_{N,\beta} \leq \int_X e^{-\beta V(x)} \mu_0(x) \left(\int e^{-\beta\left(\frac{1}{2}W(x,y)+V(y)\right)} \mu_0(y) \right)^{N-1}$$

and

$$-\frac{1}{N\beta} \log Z_{N,\beta(N-1)N} \leq \inf_{\mu \in \mathcal{P}_0(X)} F(\beta) =: F(\beta) \tag{8.4}$$

As a consequence, $\beta_c \leq \limsup_{N\to\infty} \beta_{N,c}$ and if there exists $\beta_0 < 0$ such that

$$\sup_{x\in X} \int e^{-\beta_0\left(\frac{1}{2}W(x,y)+V(y)\right)} \mu_0(y) < \infty, \quad \int_X e^{-\beta_0 V} \mu_0 < \infty \tag{8.5}$$

then $\beta_{N,c} < \beta_0$ and $\beta_c < \beta_0$.

Proof. First observe that it will be enough to consider the case when $V = 0$. (Otherwise, we just replace μ_0 with $e^{-\beta V}\mu_0$.) Decompose $-\beta H^{(N)} = \frac{1}{N}\sum_{i=1}^N f_i$, where f_i is the sum of $\frac{1}{2}W(x_i, x_j)$ over all j such that $j \neq i$. The arithmetic–geometric means inequality gives

$$\int_{X^N} e^{-\beta H^{(N)}} \mu_0^{\otimes N} \leq \sum_{i=1}^N \frac{1}{N} \int_{X^N} e^{f_i} \mu_0^{\otimes N} = \int_X \mu_0 \left(\int e^{-\beta\frac{1}{2}W(x,y)} \mu_0(y) \right)^{N-1}.$$

Hence, estimating the latter integral over X with the sup over X proves the first inequality in the proposition. To prove the second one first note Gibbs variational principle (Jensen's inequality) gives: for any given $\mu \in \mathcal{P}(X)$

$$-\frac{1}{N\beta} \log Z_{N,\beta} := \int_{X^N} e^{-\beta N E^{(N)}} \mu_0^{\otimes N} \leq \beta \int_{X^N} E^{(N)} \mu^{\otimes N} - S(\mu), \quad E^{(N)} := H^{(N)}/N$$

as long as the right-hand side is well defined. In the case when $H^{(N)}$ is of the form in the lemma

$$\int_{X^N} E^{(N)}\mu^{\otimes N} = \frac{1}{N}\frac{1}{(N-1)}N(N-1)E(\mu) = \frac{N-1}{N}E(\mu),$$

which proves 8.4, by taking the infimum over μ. □

The following result generalizes the case of the logarithmic interaction considered in [14,41].

Proposition 8.6. *Under the Homogeneous Assumptions in* $\mathbb{R}^d$ *(but allowing* d *to be odd)*

$$\beta_c = \beta_{c,N} = \frac{2d}{\dot{w}}, \quad \dot{w} := \lim_{t\to-\infty}\frac{dw(e^t)}{dt} = \lim_{t\to-\infty}\frac{w(e^t)}{t}$$

if v *and* ψ_0 *are assumed bounded in a neighborhood of* 0. *Moreover,* $Z_{N,\beta} = \infty$ *when* $\beta = \frac{4n}{\dot{w}}$.

Proof. To simplify the notation, we will prove the proposition in the case when $V = 0$ (but the proof in the general case is essentially the same). First observe that

$$\sup_X \int_X e^{-\frac{\beta}{2}W(x,y)}\mu_0(y) < \infty \iff \int_0^1 e^{-\frac{\beta}{2}w(r)}r^d\frac{dr}{r} < \infty \iff \beta > \frac{2d}{\dot{w}} \tag{8.6}$$

Indeed, since w is decreasing, $w(r) \leq C$ if $r \geq 1$, and hence, using that μ_0 is a probability measure,

$$\int_X e^{-\frac{\beta}{2}W(x,y)}\mu_0(y) = \int_X e^{-\frac{\beta}{2}w(|x-y|)}\mu_0(y) \leq \int_{X\cap\{|x-y|\leq 1} e^{-\frac{\beta}{2}w(|x-y|)}\mu_0(y) + e^{-\frac{\beta}{2}C}$$

Changing variables in the integral above and setting $\gamma := -\beta$ yields

$$\int_{X\cap\{|x-y|\leq 1} e^{\frac{\gamma}{2}w(|x-y|)}\mu_0(y) = \int_{\{|z|\leq 1\}} e^{\frac{\gamma}{2}W(|z|)}e^{-\psi_0(x+z)}d\lambda(z) \leq C'\int_{\{|z|\leq 1\}} e^{\frac{\gamma}{2}W(|z|)}d\lambda(z)$$

using that ψ_0 is bounded from below. This proves 8.6, using Lemma 8.3 in the last equivalence (by setting $t := \log r$). Hence, applying the previous lemma gives

$$\beta_{N,c} < \frac{2d}{\dot{w}} \tag{8.7}$$

To prove that $\beta_{N,c} \geq 2d/\dot{w}$, we restrict the integration over X^N to a ball B_R of radius R centered at the origin and use that w is decreasing to get

$$Z_{N,\beta} \geq \int_{B_R^N} e^{-\frac{N(N-1)}{2N}w(R)}\mu_0^{\otimes N} \geq Ce^{-\frac{\beta N}{2}w(R)}(R^d)^N \geq C'$$

Setting $R = e^t$ thus gives

$$(Z_{N,\beta})^{1/N} \geq C^{1/N}e^{-t\left(\frac{\beta}{2}\frac{1}{2t}w(e^t)-d\right)}$$

Hence, if $\beta < 2d/\dot{w}$, then as $R \to 0$, i.e., $t \to -\infty$ we get $(Z_{N,\beta})^{1/N} \geq C^{1/N} e^{-t\delta}$ for some $\delta > 0$. This means that $Z_{N,\beta} = \infty$, which proves $\beta_{N,c} = 2d/\dot{w}$. Moreover, if $\beta = 2d/\dot{w}$ then the argument shows that the integral of $e^{-\beta N E^{(N)}} \mu_0^{\otimes N}$ over B_R^N does not tend to zero as $R \to 0$. Since μ_0 does not charge single points, it follows that $Z_{N,\beta} = \infty$ (for d even this is a special case of the last statement in Cor 8.1).

Next, thanks to the second inequality in Lemma 8.5 the inequality 8.7 implies that

$$\beta_c \leq \frac{2d}{\dot{w}}$$

All that remains is thus to verify the reversed inequality. To this end, fix β such that $F(\beta) > -\infty$, i.e., such that there exists a constant C such that

$$\beta E(\mu) - S(\mu) \geq -C \tag{8.8}$$

For $\epsilon > 0$ set $\nu_\epsilon = (T_\epsilon)_* \nu_0$ where ν_0 is any fixed probability measure such that $S(\nu_0) > -\infty$. Then, on the one hand, as $t := (\log \epsilon)$ tends to $-\infty$

$$\frac{1}{t} E(\nu_{e^t}) = \frac{1}{2} \int_{X^2} \frac{1}{2t} w(e^t |x - y|) \nu_0(x) \nu_0(y) \to \frac{1}{2} \dot{w}$$

by the monotone convergence theorem (using that the integrand is monotone in t, by concavity). On the other hand,

$$S(\nu_\epsilon) = S(\nu_0) + d \log \epsilon$$

Hence, applying the inequality 8.8 to ν_ϵ and dividing both sides with t implies, by letting $t \to -\infty$, that

$$\frac{\beta}{2} \dot{w} - d \leq 0.$$

This shows that $\beta_c \geq \frac{2d}{\dot{w}}$, as desired. □

Remark 8.7. Remarkably, it is always the case that $F(\beta_c) < \infty$ when X is a compact domain in $\mathbb{R}^d$ and $W(x, y) = -\log(|x - y|$. Indeed, this follows from Adam's generalization of the Moser–Trudinger inequality in $\mathbb{R}^2$, as discussed in [8] (see also [14,41]). This finiteness should be contrasted with the general divergence $Z_{N,\beta_c} = -\infty$ for any N (see Cor 8.1).

Note that if X is compact and W is finite, then $\beta_c = \beta_{c,N} = -\infty$, but the converse does not hold, as illustrated by an application of the previous proposition to the case when

$$W(x, y) = \log(\log 1/|x - y|).$$

8.3. The Anisotropic Case

Consider now the case when the Main Assumptions hold and W is translationally invariant

$$W(z,w) = -\Psi(z-w), \quad \Psi \in PSH_{\boldsymbol{a}}(\mathbb{C}^n), \tag{8.9}$$

but not necessarily isotropic. More generally, since we will only be concerned with integrability properties we allow that 4.7 only holds up to a bounded term.

Proposition 8.8. *Consider the Main Assumptions and assume moreover that W is translationally invariant (up to a bounded term). Then there exists a positive number γ such that*

$$\max\{\beta_{N,c}, \beta\} \leq -\gamma < 0$$

Proof. First assume that $V = 0$. Then the first integral appearing in the uniform integrability property 8.5 may, after making the change of variables $z = y - z$, be estimated as

$$\begin{aligned}\int e^{\frac{\beta}{2}\Psi(y-x)} e^{-\Psi_0(y)} dy &= \int e^{\frac{\beta}{2}\Psi(z)} e^{-\Psi_0(z+x)} dz \\ &\leq \left(\int e^{\frac{p\beta}{2}\Psi(z)} dz\right)^{1/p} \left(\int e^{-q\Psi_0(z+x)} dz\right)^{1/q}, \end{aligned} \tag{8.10}$$

using Hölder's inequality with conjugate exponents p and q. By the translational invariance of Lebesgue measure the integral in the second factor is given by the integral of $\int e^{-q\Psi_0(y)} dy$ and thus independent of x. Moreover, it follows from the openness statement in Cor8.1 that the integral is finite for q sufficiently close to 1. Similarly, we can then make the integral in the first factor finite by taking β negative, but sufficiently close to 0. Finally, in the case when V is not identically zero we first apply the Cauchy–Schwartz inequality to estimate

$$\left(\int e^{-\beta\left(\frac{1}{2}W(x,y)+V(y)\right)} \mu_0(y)\right)^2 \leq \int e^{-2\beta\frac{1}{2}W(x,y)} \mu_0(y) \int e^{-2\beta V(y)} \mu_0(y)$$

and then repeat the previous argument to both integrals appearing in the right-hand side. □

Next, consider the case when Ψ has an isolated singularity at the origin, i.e., Ψ is locally bounded on the complement of the origin. Then one gets the following concrete bound, expressed in terms of the integrability threshold $c_0(\Psi)$ of Ψ on a ball B_ϵ centered at the origin in $\mathbb{C}^n$ of sufficiently small radius ϵ (discussed in Remark 8.2).

Proposition 8.9. *Let X be a compact subspace of $\mathbb{C}^n$ and assume that Ψ has an isolated singularity at the origin and that V and Ψ_0 are bounded. Then, for any sufficiently small ϵ*

$$\max_{N\geq 2}\{\beta_{N,c}, \beta_c\} = -\frac{1}{2}c_0(\Psi) < 0.$$

Proof. The assumptions ensure that the bounds 8.5 in Lemma 8.5 hold iff the sup is replaced with an integral, i.e., iff $Z_{2,\beta_0} < \infty$ iff $\int_{B_\epsilon} e^{\beta_0 \frac{1}{2}\psi} d\lambda < \infty$ (as seen by changing variables as in the first equality in formula 8.10). Hence, we can conclude using the very definition of $c_0(\Psi)$. □

In fact, as discussed in Remark 8.2 it is enough to assume that $PSH(\mathbb{C}^n)$. The invariant $c_0(\Psi)$ plays a key role in current complex geometry and can be estimated from below in terms of certain multiplicities (expressed as local intersection numbers) [25]. In the "algebraic" case in formula 1.11, the integrability threshold $c_0(\Psi)$ coincides with the *log canonical threshold* at $0 \in \mathbb{C}^n$ of the ideal in the polynomial ring $\mathbb{C}[z_1, .., z_n]$ generated by the corresponding polynomials $P_j(z)$ [47].

Example 8.10. The log canonical threshold can be computed using algebro-geometric techniques. For example, when $\Psi(z) = \log\left(|z_1|^{2\alpha_1} + ... + |z_n|^{2\alpha_n}\right)$ for positive real numbers α_i one gets $c_0(\psi) = 1/\alpha_1 + ... + 1/\alpha_n$ [47, Example 1.9].

In the simplest case when Ψ is "algebraic quasi-homogeneous" of degree d (Example 2.5) with an isolated singularity at the origin (i.e., the zero locus of corresponding polynomials P_j only intersect at the origin), we have, by homogeneity, that

$$\Psi = d \log |z|^2 + \varphi(z),$$

for a positive number d and a continuous function φ, which descends to the compact quotient $(\mathbb{C}^{n+1} + \{0\})/\mathbb{C}^*_{\boldsymbol{a}}$ and is thus bounded. In this case, it thus follows from Prop 8.6 that

$$\beta_{N,c} = \beta_c = \frac{4n}{d}.$$

A wide variety of such Ψ may be obtained by taking $P_i = \partial f(z)/\partial z_i$ for given quasi-homogeneous polynomial f with an isolated degenerate zero at the origin in $\mathbb{C}^n$. Then $\Psi(z)$ can be expressed in terms of a Ginzburg–Landau-type potential:

$$\Psi(z) = \log\left(\sum_i |\frac{\partial f}{\partial z_i}(z)|^2\right),$$

so that $W(z, w)$ is the standard logarithmic interaction precisely when f is proportional $z_1^2 + ... + z_n^2$.

8.4. Existence of Maximum Entropy Measures

Combining Prop 8.8 with the results in Sect. 4.1 yields the following existence result:

Proposition 8.11. *Consider the Main Assumptions when X is compact. Then, for any $e \in]e_{min}, e_0[$ there exists a maximum entropy measure μ^e. If moreover $W(x, y)$ is assumed translationally invariant (up to a bounded term), then there exists a maximum entropy measure μ^e for any $e \in [e_0, e_{max}[$. In particular, this is the case under the Homogeneous Assumptions.*

Turning to the non-compact case we recall that, under the Main Assumptions,

$$\mu_0 = e^{-\Psi_0} d\lambda$$

for $\Psi_0 \in PSH_a(X)$. As a consequence, if Ψ_0 is also assumed to be a continuous exhaustion function (which is automatically the case if Ψ_0 is rotationally invariant), then Prop 4.8 implies the following

Proposition 8.12. *Consider the Main Assumptions and assume that Ψ_0 is continuous exhaustion function and that the growth assumption 4.2 holds for a ϕ_0 such that $\phi_0/\Psi_0 \to 0$ uniformly as $|x| \to \infty$. If $W(x, y)$ is assumed translationally invariant (up to a bounded term), then there exists a maximum entropy measure μ^e for any $e \in]e_{min}, e_{max}[$.*

Proof. According to Prop 4.8, we just have to verify that $\int e^{\delta\Psi_0}\mu_0 < \infty$for some $\delta > 0$. But this follows from openness property in Cor 8.1. □

For example, the previous proposition applies when $X = \mathbb{R}^{2n}$ endowed with a centered Gaussian measure, $V = 0$ and W is of the "algebraic quasi-homogeneous" form in Example 2.5.

9. Strict Concavity of $S(e)$

In this final section, we show how to deduce a stronger strict concavity result for $S(e)$ under the Homogeneous Assumptions, using a uniqueness result for minimizers of F_β shown in the companion paper [8]. The starting point is the following criterion for the strict concavity of $S(e)$ in the high-energy region:

Proposition 9.1. *Assume that X is compact and that F_β has a unique minimizer on $\mathcal{P}(X)$ for any $\beta \in]\beta_c, 0[$. If the energy approximation property holds and $E(\mu_\beta) \to e_{max}$ as $\beta \to \beta_c$, then $S(e)$ is strictly concave on $]e_0, e_{max}[$ (in particular, this is the case if $E(\mu)$ is continuous on $\mathcal{P}(X)$).*

Proof. As pointed out in the proof of Prop 5.3, the uniqueness assumption implies that $F(\beta)$ is differentiable. Thus, we can conclude by applying Lemma 1.1. □

In the case of the point vortex model, the uniqueness assumption in the previous proposition (and the energy approximation property) holds on any simply connected compact domain X [15]. Moreover, by the concentration/compactness alternative established in [15], the blow-up property holds iff μ_{β_j} converges weakly toward a Dirac mass (such domains X are called *domains of the first kind* in [15]). The following result is shown in [8]:

Theorem 9.2. *(Uniqueness) Let X be a ball centered at the origin in $\mathbb{R}^{2n}$ or all of $\mathbb{R}^{2n}$. Assume that W and V satisfy the Homogeneous Assumptions and that $v + \beta\psi_0$ is strictly concave wrt* $\log r$ *when $r > 0$ for a given $\beta < 0$. Then any minimizer of $F_\beta(\mu)$ is uniquely determined. If the latter assumption is replaced by the assumption that $W(x, y)$ is a weakly positive definite kernel and that $w(r)$ is strictly increasing, then minimizers are uniquely determined modulo translation when $X = \mathbb{R}^{2n}$ and unique when X is a ball.*

We finally arrive at the following

Theorem 9.3. *Under the Homogeneous Assumptions, the entropy $S(e)$ is concave for $e > E(\mu_0)$ and strictly concave if X is a ball and either v is strictly concave wrt* $\log r$ *or w is strictly increasing for $r \in]0, \infty[$. If moreover $W(x, y)$ is a weakly positive definite kernel, then $S(e)$ is strictly concave on $]e_{min}, e_{max}[$.*

Proof. First consider the case when X is compact and W, V and Ψ_0 are continuous and v is strictly concave wrt logr, when $r > 0$. Then the strict concavity of $S(e)$ follows directly from combining the previous theorem with Lemma 5.2 (and similarly if w is strictly increasing), using that the energy approximation property holds under the Main Assumptions and hence also under the Homogeneous Assumptions. Next, if v is not assumed strictly concave wrt logr, we replace v with $v + \epsilon r$. Then the corresponding entropy $S_\epsilon(e)$ is concave and letting $\epsilon \to 0$ reveals that $S(e)$ is also concave. The general case is then deduced from the previous case using the approximation arguments employed in the proof of Theorem 6.9. Finally, if $W(x, y)$ is weakly positive definite, then by Prop 5.3 $S(e)$ is also strictly concave on $]e_{min}, e_0[$ and continuous on $]e_{min}, e_0]$. This means that $S(e)$ is strictly concave on both $]e_{min}, e_0[$ and $]e_0, e_{max}[$. Since S is continuous on $]e_{min}, e_{max}[$ it follows that S is strictly concave on $]e_{min}, e_{max}[$. (Indeed, otherwise it would be affine on some open interval in $]e_{min}, e_{max}[$ which would contradict the strict concavity on $]e_{min}, e_0[$ or $]e_0, e_{max}[$.) □

10. Appendix

In this appendix, we provide, for the convenience of the reader, some proofs of essentially well-known results stated in Sect. 2 and the proof of Lemma 7.4.

10.1. Proof of Lemma 2.1

Assume that f is *not* strictly concave in the interior of $[y_0, y_1]$. Then there exists an open interval $I \subset]y_0, y_1[$ such that f^* is affine on I. In particular, there exists a number a such that $f'(y) \equiv a$ on I. Note that $a \in [x_0, x_1]$. Indeed, since f is concave $(\partial f)(]-\infty, x_0]) \Subset \{y \geq f'(x_0)\}$ and $(\partial f)([x_1, \infty[) \Subset \{y \leq f'(x_1)\}$. Hence, by 2.2, $x \in [x_0, x_1]$ and since $a \in (\partial f)(y)$ for any $y \in I$ it follows from 2.2 that $I \subset (\partial f)(a)$, showing that f is not differentiable at a.

10.2. Proof of Lemma 2.2

First observe that f^{**} is continuous on U. Indeed, f^{**} is concave (since it is an inf of affine functions), and hence, it is enough to check that f is finite on U. But $f^{**} \geq f$ and by assumption $f > -\infty$ on U. Moreover, by assumption there exists a constant C such that $f \leq C$. Since the constant function C is a contender for the inf in formula 2.4 it follows that $f^{**} \leq C$, showing that f^{**} is finite and thus continuous on U. As a consequence, $\Omega := \{f^{**} > f\} \cap U$ is open in U. Now fix a point $x_0 \in \Omega$. Since Ω is open there exists $x_{0,\pm}$ in Ω such that $x_{0,-} < x_0 < x_{0,+}$. Moreover, since f is usc we may assume that the affine function $a(x)$ on $[x_{0,-}, x_{0,+}]$

with prescribed boundary values $a(x_{0,\pm}) = f^{**}(x_{0,\pm})$ satisfies $f(x) < a(x)$ on $[x_{0,-}, x_{0,+}]$. Hence, the continuous function $\tilde{f}$ defined as f on the complement of $[x_{0,-}, x_{0,+}]$ and as $a(x)$ on $[x_{0,-}, x_{0,+}]$ is concave on $\mathbb{R}$ and satisfies $\tilde{f} \geq f$. But then it follows that $\tilde{f} = f$, by formula 2.4.

10.3. Proof of Lemma 2.3

The function $\rho := -\log(-\phi)$ on Ω is, clearly, an exhaustion function. Moreover, ρ is psh. Indeed, fix $z_0 \in \Omega$ and $z \in \mathbb{C}^n$ and consider the restriction of ρ to the complex line $w \mapsto z_0 + wz$, parametrized by $w \in \mathbb{C}$. First assume that ϕ is smooth at z_0 and factorize the Laplacian on $\mathbb{C}_w$ in the standard way, $\Delta = 4\partial_w \partial_{\bar{w}}$. We get

$$\begin{aligned}\partial_w \partial_{\bar{w}} \rho &= -\partial_w \left(\frac{\partial_{\bar{w}}(-\phi)}{-\phi} \right) = \partial_w \left(\frac{\partial_{\bar{w}} \phi}{-\phi} \right) \\ &= \frac{\partial_w \partial_{\bar{w}} \phi}{-\phi} - \partial_{\bar{w}} \phi \partial_w (\phi^{-1}) = \frac{\partial_w \partial_{\bar{w}} \phi}{-\phi} + \partial_{\bar{w}} \phi \partial_w \phi \geq 0 + 0.\end{aligned}$$

Hence, $w \mapsto \rho(z_0 + wz)$ is subharmonic close to z_0. The subharmonicity in the general case is shown in a similar way using either distributional derivatives or a regularization argument. Finally, denoting by ρ_Y a continuous psh exhaustion function of Y the maximum of ρ and ρ_Y defines a psh exhaustion function of $\{\phi < 0\} \cap Y$. Indeed, in general, the maximum of two psh functions is still psh (as follows from the corresponding standard result for subharmonic functions).

10.4. Proof of Lemma 7.4

Since $\psi(z, w)$ is locally bounded from above, it will be enough to consider the complement of the diagonal in $D \times D$ (using that the diagonal is pluripolar [23, Thm 5.24]). In this region, $\log(|z - w|^2)$ is pluriharmonic, i.e., its complex Hessian vanishes (since $\log |\xi|^2$ is harmonic when $\xi \neq 0$). Hence, the complex Hessian $\partial\bar{\partial}\psi$ coincides with $\partial\bar{\partial}$ applied to $-\log |1 - z\bar{w}|^2$, i.e., to $-\log(1 - z\bar{w}) - \log(1 - \bar{z}w)$. Accordingly, a direct computation yields

$$\begin{aligned}\partial\bar{\partial}\psi(z, w) &= \begin{pmatrix} 0 & (1 - \bar{z}w)^{-2} \\ (1 - z\bar{w})^{-2} & 0 \end{pmatrix}, \\ \partial\bar{\partial}\left(\phi(z) + \phi(w)\right) &= \begin{pmatrix} 2(1 - z\bar{z})^{-2} & 0 \\ 0 & 2(1 - w\bar{w})^{-2} \end{pmatrix}.\end{aligned}$$

In particular, when $\lambda = 1/2$ we get

$$\partial\bar{\partial}\left(\psi(z, w) + \lambda\left(\phi(z) + \phi(w)\right)\right) = \begin{pmatrix} (1 - z\bar{z})^{-2} & (1 - \bar{z}w)^{-2} \\ (1 - z\bar{w})^{-2} & (1 - w\bar{w})^{-2} \end{pmatrix}$$

Since the trace of this Hermitian matrix is manifestly nonnegative the matrix is semi-positive definite iff its determinant is nonnegative. But the determinant is nonnegative iff

$$(1-z\bar{z})^2(1-w\bar{w})^2 \leq (1-z\bar{w})^2(1-\bar{z}w)^2 \iff (1-z\bar{z})(1-w\bar{w}) \leq (1-z\bar{w})(1-\bar{z}w),$$

which in turn is equivalent to $-z\bar{z}-w\bar{w} \leq -z\bar{w}-z\bar{w}$ and hence also to the trivial inequality $0 \leq |(z-w)|^2$. The same computation also reveals that when $\lambda < 1/2$ the determinant is negative at $(z, w) = (0, 0)$ and hence also at any point in the complement of the diagonal in $D \times D$ which is sufficiently close to $(0, 0)$.

Funding Open access funding provided by Chalmers University of Technology.

References

1. Aly, J.J.: Thermodynamics of a two-dimensional self-gravitating system. *Phys. Rev. E* **49**, 3771, 1994
2. Ambrosio, L., Gigli, N., Savaree, G.: Gradient flows in metric spaces and in the space of probability measures Lectures in Mathematics ETH Zurich. Birkhauser Verlag, Basel (2005)
3. Antonov, V.A.: Original inVest. Leningrad Univ.7, 135 (1962); English translation in Dynamics of Star Clusters (IAU Symposium, Vol. 113), J. Goodman and P. Hut, ed.(Reidel,Dordrecht,1985), p.525
4. Arora, J.S., Chahande, A.I., Paeng, J.K.: Multiplier methods for engineeringoptimization. *Int. J. Numer. Methods Eng.* **32**, 1485–1525, 1991
5. Berg, C.,Christensen ,J.P.R., Ressel, P.: Harmonic Analysis on Semigroups. Graduate Texts in Mathematics100. Springer, New York. Theory of positive definite andrelated functions. (1984)
6. Berman, R.J.: On large deviations for gibbs measures, mean energy and gamma-convergence. *Construct. Approx.* **48**, 3–30, 2018
7. Berman, R.J.: Priors leading to well-behaved Coulomb and Riesz gases versus zeroth-order phase transitions - a potential-theoretic characterization. *Electron. J. Probab.* **26**(145), 1–49, 2021
8. Berman, R.J.: Uniqueness of free energy minimizers and the Q-curvature equation (in preparation)
9. Berman, R.J., Berndtsson, B.: Real Monge-Ampère equations and Kähler-Ricci solitons on toric log Fano varieties. *Ann. Fac. Sci. Toulouse Math. (6)* **22**(4), 649–711, 2013
10. Berman, R.J., Berndtsson, B.: Symmetrization of Plurisubharmonic and Convex Functions. *Indiana Univ. Math. J.* **63**(2), 345–365, 2014

11. Berndtsson, B.: Subharmonicity properties of the Bergman kernel and some other functions associated to pseudoconvex domains. *Ann Inst Fourier* **56**, 1633–1662, 2006
12. Berndtsson, B.: The openness conjecture and complex Brunn-Minkowski inequalities. Complex geometry and dynamics, 29–44, Abel Symp.,10, Springer, Cham, (2015)
13. Binney, J., Tremaine, S.: Galactic Dynamics. Princeton Series in Astrophysics (2008)
14. Caglioti, E., Lions, P.-L., Marchioro, C., Pulvirenti, M.: A special class of stationary flows for two-dimensional Euler equations: a statistical mechanics description. *Commun. Math. Phys.* **143**(3), 501–525, 1992
15. Caglioti, E., Lions, P.-L., Marchioro, C., Pulvirenti, M.: A special class of stationary flows for two-dimensional Euler equations: a statistical mechanics description. *Part II. Commun. Math. Phys.* **174**, 229–260, 1995
16. Campa, A., Dauxois, T., Ruffo, S.: Statistical mechanics and dynamics of solvable models with long-range interactions. Elsevier, Amsterdam (2009)
17. Chafaï, D., Gozlan, N., Zitt, P.-A.: First-order global asymptotics for confined particles with singular pair repulsion. *Ann. Appl. Probab.* **24**(6), 2371–2413, 2014
18. Chang, S.-Y.A.: Conformal invariants and partial differential equations. *Bull. Amer. Math. Soc.* **42**, 365–393, 2005
19. Chanillo, S., Kiessling, M.K.-H.: Rotational symmetry of solutions to some nonlinearproblems in statistical mechanics and geometry. *Comm. Math. Phys.* **160**, 217–238, 1994
20. Chavanis, P.-H.: Statistical mechanics of two-dimensional vortices and stellar systems. Pages 208-292 in [21]
21. Dauxois, T., Ruffo, S., Arimondo, E., Wilkens, M.: eds. Dynamics and Thermodynamics of Systems with Long Range Interactions, vol. 602 of Lecture Notes in Physics (Springer, New York, 2002)
22. Dembo, A., Zeitouni, O.: Large deviation techniques and applications. Jones and Bartlett Publ, Burlington (1993)
23. Demailly, J.: Complex Analytic and Differential Geometry. https://www-fourier.ujf-grenoble.fr/~demailly/manuscripts/agbook.pdf
24. Demailly, J.-P., Kollár, J.: Semi-continuity of complex singularity exponents and Kähler-Einstein metrics on Fano orbifolds. *Ann. Sci. l'École Normale Supérieure, Série 4, Tome* **34**(4), 525–55, 2001
25. Demailly, Jean-Pierre.: Hoàng Hiep Pham: a sharp lower bound for the log canonical threshold. *Acta Math.* **212**(1), 1–9, 2014
26. Di Cintio, P., Gupta, S., Casetti, L.: Dynamical origin of non-thermal states in galactic filaments. *Mon. Not. R. Astron. Soc.* **475**(1), 1137–1147, 2018
27. Driscoll, C.F., Fine, K.S.: Experiments on vortex dynamics in pure electron plasmas. *Phys. Fluids B* **2**, 1359, 1990
28. Dunkel, J., Hilbert, S.: Consistent thermostatistics forbids negative absolute temperatures. *Nat. Phys.* **10**, 67, 2014
29. Eyink, G.L., Spohn, H.: Negative-temperature states and large-scale, long-lived vortices in two-dimensional turbulence. *J. Stat. Phys.* **70**, 3/4, 1993
30. Ellis, R.S., Haven, K., Turkington, B.: Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles. *J. Statist. Phys.* **101**(5–6), 999–1064, 2000
31. Ellis, R.S., Touchette, H., Turkington, B.: Thermodynamic versus statistical nonequivalence of ensembles for the mean-field Blume-Emery-Griffiths model. *Physica A* **335**(3—-4, 15), 518–538, 2004
32. Garrett, A.J.M.: Maximum Entropy with Nonlinear Constraints: Physical Examples. In Maximum Entropy and Bayesian Methods pp 243-249, Fundamental Theories of Physics book series (FTPH, volume 39)
33. Gross, D.H.E.: Microcanonical thermodynamics: phase transitions in small systems. World scientific lecture notes in physics. Vol **66** (2001)
34. Guan, Q., Zhou, X.: A proof of Demailly's strong openness conjecture. *Ann. of Math. (2)* **182**(2), 605–616, 2015

35. Gauthier, et al.: Giant vortex clusters in a two-dimensional quantum fluid. *Science* **364**(6447), 1264–1267, 2019 arXiv:1801.06951
36. Hilbert, S., Hänggi, P., Dunkel, J.: Thermodynamic laws in isolated systems. Physical Review E, APS (2014)
37. Huang, K.: Statistical mechanics. J. Wiley, New Jersey (1987)
38. Jaynes, E.T.: Information theory and statistical mechanics. *Phys. Rev.* **106**, 620, 1957
39. Johnstone, et al.: Evolution of large-scale flow from turbulence in a two-dimensional superfluid. *Science* **365**(6447), 1267–1271, 2019 arXiv:1801.06952
40. Kiessling, M.K.H.: On the equilibrium statistical mechanics of isothermal classical self-gravitating matter. *J. Stat. Phys.* **55**, 203–257, 1989
41. Kiessling, M.K.H.: Statistical mechanics of classical particles with logarithmic interactions. *Comm. Pure Appl. Math.* **46**, 27–56, 1993
42. Kiessling, M.K.-H.: The unbounded 2D guiding center plasma. *J. Plasma Phys.* **54**, 11–29, 1995
43. Kiessling, Michael K.-H.: Statistical mechanics approach to some problems in conformal geometry. *Stat. Mech. Rigorous Results Appl. Phys. A* **279**(1), 353–368, 2000
44. Lynden-Bell, D.: Negative specific heat in astronomy, physics and chemistry. *Physica A* **263**(1–4, 1), 293–304, 1999
45. Majda, A.J., Bertozzi, A.L.: Vorticity and incompressible flow. Cambridge University Press, Cambridge (2010)
46. Marchioro, C., Pulvirenti, M.: Mathematical Theory of Incompressible Nonviscous Fluids, in: Applied Mathematical Sciences, vol. **96**, Springer-Verlag, New York, (1994)
47. Mustata, M.: IMPANGA lecture notes on log canonical thresholds. In: Contributions to Algebraic Geometry: Impanga Lecture Notes. Editor: P.Pragasz, EMS (2012)
48. Onsager, L.: Statistical hydrodynamics. *Nuovo Cim. Suppl.* **6**, 279–287, 1949
49. Padmanabhan, T.: Statistical mechanics of gravitating systems. *Phys. Rep.* **188**, 285, 1990
50. Ramsey, N.F.: Thermodynamics and statistical mechanics at negative absolute temperatures. *Phys. Rev.* **103**, 20–28, 1956
51. Rockafellar, R.T.: Convex analysis. Reprint of the 1970 original. Princeton Landmarks in Mathematics. Princeton Paperbacks. Princeton University Press, Princeton, NJ, (1997)
52. Rugh, H.H.: Microthermodynamic formalism. *Phys. Rev. E* **64**, 055101, 2001
53. Silvestrea, C.H., Rocha Filho, T.M.: Ergodicity in a two-dimensional self-gravitating many-body system. *Phys. Lett. A* **380**(3), 337–348, 2016
54. Smith, R.A., O'Neil, T.M.: Nonaxisymmetric thermalequilibria of a cylindrically bounded guiding center plasma or discrete vortex system. *Phys. Fluids B* **2**, 2961–2975, 1990
55. Touchette, H., Ellis, R.S., Turkington, B.: An introduction to the thermodynamic and macrostate levels of nonequivalent ensembles. *Physica A* **340**(1–3, 1), 138–146, 2004
56. Villani, C.: Topics in optimal transportation. Graduate Studies in Mathematics, 58. American Mathe-matical Society, Providence, RI, xvi+370 pp (2003)

Chalmers University of Technology: Chalmers tekniska hogskola,
Gothenburg
Sweden.

(*Received December 8, 2021 / Accepted September 28, 2022*)
Published online November 10, 2022