Journal of Statistical Physics, Vol. 55, Nos. 1/2, 1989

On the Equilibrium Statistical Mechanics of Isothermal Classical Self-Gravitating Matter

Michael K.-H. Kiessling[1]

Received June 8, 1988; revision received November 1, 1988

The canonical ensemble is investigated for classical self-gravitating matter in a finite container $\Lambda^{[d]} \subset \mathbb{R}^d$, $d=3$ and 2. Starting with modified gravitational interactions (smoothed-out singularity), it is proven by explicit construction that, in the w^*-topology, the canonical equilibrium measure converges to a superposition of Dirac measures when the limit of exact Newtonian gravitational interactions between classical point particles is taken. The consequences of this result for more realistic classical systems are evaluated, and the existence of a gravitational phase transition is proven. The results are discussed with view toward applications in astrophysics and space science. Some attention is paid also to the problem of founding thermodynamics by means of statistical mechanics.

KEY WORDS: Canonical ensemble; classical point particles; unstable interactions; Dirac measure; mean-field limit; equilibrium states; gravitational phase transition.

1. INTRODUCTION

One of the yet unanswered questions of fundamental interest in physics is whether or not the laws of thermodynamics hold for systems controlled mainly or exclusively by gravitational interactions. As a matter of fact, this problem is rather controversially discussed in the literature. In view of this, it is interesting to observe that some basic aspects of the statistical mechanics equilibrium of classical self-gravitating matter (see the discussion below) have so far not been subject to rigorous mathematical considerations. It can be expected that careful reconsiderations of these aspects will shed new light on the above controversy.

So far, exact equilibrium statistical mechanics results in the N-particle

[1] Institut für Theoretische Physik IV, Ruhr-Universität Bochum, D-4630 Bochum 1, Federal Republic of Germany.

phase space $\Gamma^{[d]}(N)$, i.e., for finite systems, with purely gravitationally interacting classical point particles have been derived for one- and for two-dimensional systems ($d=1$ and $d=2$) only.[1,2] In ref. 1, the exact equation of state was obtained for an isothermal system in a container in one and in two spatial dimensions (see also ref. 3). In ref. 2, the one-particle distribution function of the one-dimensional system was evaluated exactly both for the canonical and for the microcanonical ensemble in all $\mathbb{R}$, but with fixed center of mass, and then compared with the results obtained from the mean-field limit. The exact evaluation of the one-particle equilibrium distribution function for a finite system in two spatial dimensions has not been achieved. No comparable results exist at all for the finite three-dimensional systems. On the contrary, it has been often stated that in three spatial dimensions no exact classical statistical mechanics equilibrium state exists[2,4–9] but some kind of quasiequilibrium state only (e.g., refs. 8 and 10), calculated by means of a mean-field approximation that is essentially equivalent to the continuum approximation,[11–13] and that is expected to become exact in the mean-field limit. The reason for that belief lies in the peculiar nature of the Newtonian gravitational interactions which are expected to cause the ultimate collapse of a self-gravitating system, or of parts of it, to a point singularity. It should be noted that such a point singularity is interpreted in the cited references as contradicting the principles of statistical mechanics equilibrium ("To obtain any sense from statistical mechanics we must consider systems in which these highly desirable states with infinite weight are unattainable"[8]).

Nonexistence of thermodynamic equilibrium for three-dimensional classical self-gravitating systems is also regarded to be manifest from the divergence of the various partition functions (e.g., ref. 2), which is due to the local singularity of the Newtonian pair interaction potential. In order to obtain well-defined expressions for the classical partition functions, a short-distance modification of the Newtonian interactions has to be introduced[4,7–9,14,15] which cuts off or smoothes out the local singularity, and the system has to be confined to a finite container $\Lambda^{[3]}$.[14,15]

Let us now critically inquire into the arguments presented above. Roughly speaking, the arguments for the nonexistence of a thermodynamic equilibrium state of classical self-gravitating matter can be grouped into two different kinds. On the one side there are the partially intuitively motivated arguments, rejecting the possibility of a thermodynamic equilibrium state of classical self-gravitating matter because it is tacitly assumed that a thermodynamic equilibrium state must somehow be smooth and extended, as is familiar from usual (laboratory) thermodynamics. On the other side there are the related but more technically oriented arguments, rejecting the possibility of a statistical mechanics equilibrium state of classical self-gravitating matter because of the nonexistence of the partition functions, which play a fundamental and well-known role in the calculation of the thermodynamic quantities for systems with short-range stable interactions.(16) It is obvious that both lines of thought are closely oriented to "classical thermodynamics," which was developed for systems of laboratory size. However, it is tempting to argue that the above difficulties with the concept of a thermodynamic equilibrium state of classical self-gravitating matter arise because of a biased standpoint regarding what must be the properties of a thermodynamic equilibrium state. Let us therefore try to take a less biased standpoint and consider a generalized meaning of equilibrium thermodynamics simply as a tool that describes the average fate of physical systems possessing an overwhelmingly large number of degrees of freedom, without making too restrictive assumptions about the properties of the final state. However, such a generalized concept of a thermodynamic equilibrium state must incorporate the usual notion and thus has to be formulated in terms of a natural extension of the established concepts of statistical mechanics equilibrium. Having said that, let us now first consider a *Gedankenexperiment* and try to guess the final fate of self-gravitating classical matter in an idealized special situation.

We consider a large system of N purely gravitating classical point particles. The system is confined to an energetically open finite container Λ[3] in order to prevent particle escape. Furthermore, due to the interaction with the boundary, the angular momentum that is possibly stored in close binaries will finally be transported out of the system. Such a system should show a strong tendency to shrink, giving up the liberated gravitational energy to the outside world. Obviously this scenario is reminiscent of the situation depicted when discussing the canonical ensemble. Now, if we assume that the system actually collapses to a single material point, it is further clear that it cannot collapse any further. In other words, the whole system will settle down in a single point. This means that the material point should indeed be the ultimate equilibrium state.

Following our reasoning, that equilibrium state should be describable in terms of, say, the canonical ensemble. It should be noted that the argument that the configurational integral Q does not exist for particles with purely Newtonian interactions(2) does not pose a severe problem for our reasoning. Of course, when Q and, hence, the partition function Z exist, then Z is a convenient tool for the calculation of the thermodynamic quantities. However, the physical concept of a statistical mechanics equilibrium state is independent of concepts such as thermodynamic relations, from the beginning. Here, the basic quantity is the phase space probability measure μ, and that quantity has a meaning independently of whether Q exists or not.

In order to incorporate such exotic states like material points as equilibrium states in the framework of statistical mechanics we have to weaken somewhat the postulates which so far govern equilibrium statistical mechanics (see, for instance, ref. 16). Let us therefore introduce the notions of "strong formulation" and of "weak formulation of thermodynamics in terms of statistical mechanics." (We consider here only continuous systems,[16] i.e., the phase space will not be a lattice). In both cases the thermodynamic equilibrium state is identified, as usual, by a corresponding equilibrium probability measure on the phase space, given as the usual expression for the microcanonical, or the canonical, etc., probability measure, involving the Hamiltonian of the system. We stipulate that by the strong formulation we understand the following: It is required that the corresponding expressions for the thermodynamic potentials of the various ensembles (i.e., the entropy, the free energy, etc.) exist for finite N as bounded functions of their natural variables, which restricts the allowed forms of the Hamiltonian. For instance, the canonical equilibrium measure is then necessarily absolutely continuous with respect to Lebesgue measure. In the strong formulation the partition functions exist *par force*, and it is this formulation that is usually understood in the literature[16] as yielding phenomenological thermodynamics in an appropriate $N \to \infty$ limit. It should be noted, however, that the strong formulation does not mean that the equilibrium state is homogeneous. Complementary to that now, we propose to introduce also a weak formulation and to relax somewhat the constraints on the probability measures. We propose to allow as (weak) thermodynamic equilibrium states also those weak* limit points (meaning the limit in the space of distributions, roughly) of an infinite sequence of thermodynamic equilibrium states within the strong formulation for which the thermodynamic potentials cease to exist as bounded functions, provided the measures behave reasonably in an appropriate $N \to \infty$ limit. By the latter statement we mean that the result must not depend on whether we take a w^*-limit for $N < \infty$ and then let N go to infinity or whether we take that limit after we let $N \to \infty$. In the weak formulation the canonical equilibrium measure, for instance, is allowed to be a Dirac measure.

Having introduced the above distinction, we return to the consideration of classical self-gravitating matter consisting of Newtonian point particles with Newtonian gravitational interactions. The nonexistence of the partition functions for these systems immediately implies that these systems do not have a thermodynamic equilibrium state in the strong sense. However, it is one of the aims of this paper to prove the existence of a statistical mechanics equilibrium state within the weak formulation. Obviously, the weak formulation thus offers a possibility of investigating equilibrium statistical mechanics for systems for which the strong formulation does not hold. In this sense the weak formulation might be viewed as being complementary, not alternative, to the strong one.

A few words of caution: The above concept of the weak formulation means that we shall treat the Newtonian expression for the gravitational interactions as if it were valid down to arbitrarily small interparticle distances. Although that is unphysical, from our results we will later be able to draw conclusions about the thermodynamics of physically more reasonable approximations to real systems.

In Section 2 the weak*-limit of the canonical equilibrium measure μ will be calculated for a finite system. More precisely, we will consider N classical point particles in a container $\Lambda^{[3]}$ with finite volume $\mathscr{V}$, which interact through modified gravitational pair interactions (smoothed out singularity) and with an externally generated gravitational potential ϕ. It will then be proven that, as the interaction potential converges pointwise to the exact Newtonian interaction potential, the configurational equilibrium density $g_N(\mathbf{r}_1,..., \mathbf{r}_N)$ of the canonical ensemble converges, in the w^*-topology, to a superposition of Dirac distributions, i.e.,

$$g_N(\mathbf{r}_1,..., \mathbf{r}_N) \rightharpoonup \int_{\Lambda^{[3]}} d^3r\, \mathscr{N} \exp[-\beta N m \phi(\mathbf{r})] \prod_{i=1}^{N} \delta(\mathbf{r}-\mathbf{r}_i) \qquad (1.1\text{a})$$

with

$$\mathscr{N}^{-1} = \int_{\Lambda^{[3]}} d^3r \exp[-\beta N m \phi(\mathbf{r})] \qquad (1.1\text{b})$$

and m the particle mass. As usual, β^{-1} denotes the product of the Boltzmann constant k_B with the temperature. The limit (1.1) is proven for a very large class of paths in the space of regularized gravitational interactions. The limit is likely to be completely independent of that path; however, that is not shown here.

Result (1.1) in fact proves that in the limit of Newtonian gravitational interactions the canonical ensemble describes systems that have collapsed to a single material point. The probability density of finding a collapsed system at $\mathbf{r} \in \Lambda^{[3]}$ depends on the external potential ϕ and is given by a Boltzmann-like factor (strictly, it is not the Boltzmann factor). For $\phi \equiv 0$, the localization in $\Lambda^{[3]}$ is completely indeterminate, and the Boltzmann-like probability density reduces to $\mathscr{V}^{-1}$.

We will also briefly address, in Section 2, the corresponding problem in two-dimensional physical space. (The one-dimensional problem has been solved in ref. 2. In one spatial dimension, no collapse occurs.) In ref. 1 it has been shown that, for finite isothermal systems in two spatial dimensions the configurational integral Q exists, for given particle number N, particle mass m, and given container $\Lambda^{[2]}$ with two-dimensional volume $\mathscr{A}$, only if the temperature T exceeds a critical value T_c. For $T < T_c$, Q diverges to $+\infty$. It has been conjectured[1] that this has to be interpreted as evidence for the collapse of such systems to a single (2D) material point for $T < T_c$. The technique presented in this work to prove (1.1) for three-dimensional systems may equally well be applied (after some minor modifications) to two-dimensional systems. Thereby we will see that an analogue of (1.1) holds if the temperature lies below a critical value T_0, with $T_0 \ll T_c$. This result verifies, and extends, the corresponding conjecture given in ref. 1; however, only for $T < T_0$. In this sense, the results presented here for the two-dimensional systems are complementary to those given in ref. 1. A rigorous treatment of the regime $T_0 < T < T_c$ has to be left for future investigations.

Result (1.1) is valid in the N-particle phase space $\Gamma^{[d]}(N)$ for $d = 3$, and also for $d = 2$ if $T < T_0$. On the other hand, equilibrium statistical mechanics is expected to become exact (i.e., equivalent to phenomenological thermodynamics) in an appropriate $N \to \infty$ limit only. For stable interactions[16] the usual thermodynamic limit (infinite volume) is the appropriate one. For unstable interactions, as are the gravitational forces, the appropriate limit is the mean-field limit.[15,17] The weak formulation requires proving the pendant of (1.1) in the mean-field limit. One can conceive of the following two procedures: (1) One can take the singular-interactions limit (denoted by w^*-lim$\cdot$) first and then the appropriate infinitely many-particle limit; (2) one can first take the mean-field limit and then go over to singular interactions. Both limiting procedures must yield the same result. The first procedure can immediately be carried out. We obtain from (1.1)

$$\lim_{N \to \infty} [w^*\text{-lim } g_N(\mathbf{r}_1, \dots, \mathbf{r}_N)] = \int_{\Lambda^{[3]}} \nu(d^3 r) \prod_{i=1}^{\infty} \delta(\mathbf{r} - \mathbf{r}_i) \tag{1.2}$$

where $\nu(d^3 r)$ is a probability measure that replaces the Boltzmann-like measure in (1.1). The measure $\nu(d^3 r)$ that comes out in the limit $N \to \infty$ depends on the chosen scaling of ϕ and m with respect to N.

In Section 3 it is shown that (1.2) also holds if the mean-field limit, as established in ref. 15, is taken before letting the interactions become Newtonian. The various possibilities for $\nu(d^3 r)$ are explored there. Use will be made of a representation theorem for permutation-invariant measures on infinite Cartesian products of measurable spaces.[18]

It is of interest to explore some of the implications of the above results for more realistic model systems. By "more realistic model systems" we understand mathematical models of self-gravitating point particle systems which take into account that the concept of Newtonian point particles with purely Newtonian classical gravitational interactions means an idealized mathematical approximation. This approximation is physically sensible only if the typical interparticle distances are considerably larger than the sizes of the real particles. For instance, more realistic model systems in the above sense are self-gravitating fermionic systems. But so are classical models that take the size of the particles into account heuristically through the introduction of a short-distance modification of the Newtonian potential, which extents over the size of a particle. The latter models may be viewed as a classical approximation to quantum mechanical systems when the size of the particles is assumed to be the classical Bohr radius or to be the nucleon radius. However, as classical particles one can also take dust particles, stars, or perhaps even galaxies. Gravitating fermions have been treated extensively in the literature (e.g., refs. 12, 17, 19–21, and references therein). However, comparatively few explicit yet rigorous results exist[14,15] for the classical interaction approximation, i.e., the smoothed-out interactions. Therefore, in Section 4 some attention is paid to the problem of (strong) isothermal equilibrium for systems of particles with modified Newtonian interactions.

Among other questions, we will address the following problem: Given an isothermal system confined to a hollow sphere, the particle interactions being given through only slightly modified Newtonian interactions (in the sense that the local singularity is smoothed out), what is the overall structure of the exact statistical mechanics equilibrium state in the mean-field limit? Our goal will be to clarify the role that is played for this problem by the well-known isothermal Emden gas spheres. Rather recently[4,22] it has been stated that, for finite systems ($N < \infty$), in the limit of vanishing modification of the interactions a (locally) stable isothermal Emden gas sphere, in the parameter regime where these objects exist, will give the best mean-field approximation to the exact statistical mechanics equilibrium state, up to errors of order N^{-1}. This implies that with slight modification of the interactions something close to an isothermal Emden gas sphere should give the best approximation. With the aid of the above-mentioned results of Sections 2 and 3, however, we will be able to infer that the isothermal statistical mechanics equilibrium state (in the strong sense) converges in the mean-field limit generally not near to an isothermal Emden gas sphere unless the temperature is extremely high. Complementary to this, we shall rigorously show that, when decreasing the temperature formally from infinity down to smaller and smaller values a phase transition occurs at a temperature located well inside the temperature regime where there exist isothermal Emden gas spheres. From the proofs there is evidence that the transition is from a self-gravitating gas phase of nearly uniform density to a phase consisting most likely of a highly condensed nucleus and a dilute atmosphere (a "planet").

The results derived in Sections 2–4 will be briefly summarized and then discussed in Section 5. Comparison will be made with existing results of the related quantum mechanical problem. Some emphasis will also be given to the important question of the dynamical accessibility of the derived equilibrium structures.

2. THE NEWTONIAN-INTERACTION LIMIT FOR FINITE SYSTEMS

We start with the consideration of the three-dimensional systems ($d=3$). The treatment of the two-dimensional ($d=2$) systems requires only moderate changes as compared to the case $d=3$. For that reason the results for $d=2$ will be stated without proof at the end of this section; however, it will be briefly outlined where the differences in the treatment of $d=2$ and $d=3$ come in.

In nearly all of what follows we assume that all quantities are measured in suitable dimensionless units such that $\hbar=1$, $G=1$, and $k_B=1$, where $\hbar$ is Planck's constant divided by 2π, G is Newton's gravitational constant, and k_B is Boltzmann's constant.

Let $\Lambda^{[3]} \subset \mathbb{R}^3$ denote the interior of a three-dimensional, simply connected box with volume $\mathscr{V}$, which is bounded by a smooth but reflecting boundary $\partial\Lambda^{[3]}$. The largest possible Euclidian distance between two points on $\partial\Lambda^{[3]}$ is denoted by $\not{A}$. The closure $\Lambda^{[3]} \cup \partial\Lambda^{[3]}$ is denoted by $\overline{\Lambda^{[3]}}$. We consider N classical Newtonian point particles in $\Lambda^{[3]}$, with equal mass m, the dynamics in the corresponding phase space $\Gamma^{[3]}(N)$ being determined by the Hamiltonian

$$\begin{aligned} H_{\varepsilon,\gamma}(N) &= \sum_{i=1}^{N} \left[\frac{\mathbf{p}_i^2}{2m} + \frac{1}{2} \sum_{\substack{j=1 \\ i \neq j}}^{N} V_{\varepsilon,\gamma}(|\mathbf{r}_i - \mathbf{r}_j|) + m\phi(\mathbf{r}_i) + W(\partial\Lambda^{[3]}, \mathbf{r}_i) \right] \\ &\equiv \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m} + H'_{\varepsilon,\gamma}(N) \end{aligned} \tag{2.1}$$

with $\mathbf{r}_i \in \Lambda^{[3]}$, and with the $\mathbf{p}_i \in \mathbb{R}^3$ being the particle momenta.

The meaning of W is that of the wall potential energy; it is to reflect the particles from the boundary of $\Lambda^{[3]}$ in such a way as to provide a thermal contact with the outside world. In principle, we should allow W to depend on the particle momenta also, but for simplicity we assume that this will not be the case. That seems to be reasonable at least for systems which have reached thermodynamic equilibrium, in which we are interested here. In thermodynamic equilibrium the influence of W on the dynamics inside most of $\overline{\Lambda^{[3]}}$ should be negligible. In total this means that W has to be $+\infty$ on $\partial\Lambda^{[3]}$, but has to decrease rapidly to zero away from $\partial\Lambda^{[3]}$. To obtain a well-defined thermodynamic mean-field limit we shall need these properties of W. For the calculations in the present section, however, we are allowed to set W equal to zero inside $\overline{\Lambda^{[3]}}$, for simplicity, and which we hereby do.

The contributions from all external sources to the gravitational potential are contained in ϕ. Thus, inside $\overline{\Lambda^{[3]}}$, ϕ is a solution of Laplace's equation. We assume $\phi \in C^{\infty}(\overline{\Lambda^{[3]}})$, for simplicity.

The interactions between the particles are described by V. This interparticle potential energy is split into two parts:

$$V_{\varepsilon,\gamma}(|\mathbf{r}_i - \mathbf{r}_j|) = V^{(1)}(|\mathbf{r}_i - \mathbf{r}_j|, \varepsilon) + \gamma V^{(2)}(|\mathbf{r}_i - \mathbf{r}_j|) \tag{2.2}$$

with $\varepsilon, \gamma \in \overline{\mathbb{R}^+}$, where $\overline{\mathbb{R}^+} := \mathbb{R}^+ \cup \{0\}$. The two interaction energies have the following meaning: $V^{(1)}$ is a member of a convenient class of modified Newtonian interaction energies. It describes purely attractive but soft (without singularity) interactions. It converges pointwise to the classical Newtonian potential, denoted by V_{cl}, as $\varepsilon \to 0$. Moreover, the modifications to the Newtonian interactions will be important (in a suitable sense to be explained below) at small interparticle distances only. It is the interaction energy $V^{(1)}$ that concerns us in this work. The second part, $V^{(2)}$, is included in (2.2) only for the sake of broader generality. It will be chosen such that it does not influence any of the principal results to be derived here. It accounts for possible further central forces of short range, e.g., finite repulsive forces. By "short range" we mean that with increasing distance between the particles the contributions from $V^{(2)}$ can be neglected against $V^{(1)}$.

Explicitly, we postulate that

$$V^{(1)}(\cdot, \cdot)\colon\ \overline{\mathbb{R}^+} \times \mathbb{R}^+ \to \mathbb{R}^- \tag{2.3}$$

is continuous in both variables and strictly increasing a.e. in the first variable, which implies that for given ε the function $V^{(1)}(\cdot, \varepsilon)$ might have a horizontal tangent at $\xi = 0$. It is required that for any $\xi \in \mathbb{R}^+$

$$\lim_{\varepsilon \searrow 0} V^{(1)}(\xi, \varepsilon) = -m^2 \xi^{-1} \tag{2.4}$$

We define

$$\Delta(\xi, \varepsilon) \equiv V^{(1)}(\xi, \varepsilon) + m^2 \xi^{-1} \tag{2.5}$$

and require $|\Delta(\xi, \varepsilon)|$ to have a convex upper envelope, denoted by $\mathscr{C}\mathscr{E}_\Delta(\xi)$, which is strictly monotonically decreasing in $\xi \in \mathbb{R}^+$. Furthermore, we observe the following nice homogeneity property of $V_{\rm cl}$: $V_{\rm cl}(t\xi) = t^{-1} V_{\rm cl}(\xi)$ for any $t \in \mathbb{R}^+$. Therefore, in addition to the above postulates we require $V^{(1)}(\cdot, \cdot)$ to be homogeneous of degree -1, i.e.,

$$V^{(1)}(t\xi, t\varepsilon) = t^{-1} V^{(1)}(\xi, \varepsilon) \tag{2.6}$$

For $V^{(2)}$ we choose a bounded and smooth function, which decreases rapidly to zero with increasing argument (e.g., a sufficiently differentiable function multiplied by a function of rapid decrease; see, for instance, ref. 23). The typical range r_0 of $V^{(2)}$ has to fulfill $r_0 \Lambda^{-1} \ll 1$.

The above-proposed properties of $V^{(1)}$ guarantee that for any small positive ξ^* and any small positive v we can find ε^* such that for $\varepsilon < \varepsilon^*$ we have $\mathscr{C}\mathscr{E}_\Delta(\xi) < v$ if $\xi > \xi^*$. That means that the modifications to the Newtonian interactions become important for $\xi < \xi^*$ only. The combination (2.2) is chosen such that both for very large and for very small interparticle distances the attractive part dominates if ε is small enough. The limit of purely Newtonian gravitational interactions is given by taking both $\varepsilon \to 0$ and $\gamma \to 0$. It will become clear from the proofs given below that $V^{(2)}$ will be important only for dynamical details (which we are not interested in here primarily), and for the equilibria as long as $\varepsilon \neq 0$. Hence, from now on we set $\gamma = 0$, for simplicity, unless otherwise stated. We drop the suffix γ, and we omit the superscript (1) at (2.3).

Remark. By construction, V converges for $\xi \in \mathbb{R}^+$ pointwise to the Newtonian expression $V_{\rm cl}$ for the gravitational interactions. The convergence is, however, not uniform in ξ. The bearing of this on the physics is readily seen by writing $V(\cdot, \varepsilon)$ explicitly as the sum of $V_{\rm cl}$ and $\Delta(\cdot, \varepsilon)$ given in (2.5). The quantity $\Delta(\cdot, \varepsilon)$ can be interpreted physically as an interaction energy to which there corresponds a force that is strictly repulsive at small enough interparticle distances, and that is singular for zero interparticle distance, just to stabilize the singular attractive Newtonian gravitational forces. This holds for any $\varepsilon > 0$. When $\varepsilon \searrow 0$ the range of the stabilizing force is reduced to zero but its maximum strength is always $+\infty$.

Let us assume that the particles are distributed in $\Lambda^{[3]}$ according to the canonical ensemble, with equilibrium density $f_N^{\rm can}$ given by

$$f_N^{\rm can} = (N!\, Z_\varepsilon)^{-1} \exp(-\beta H_\varepsilon) \tag{2.7}$$

where $Z_\varepsilon(\Lambda^{[3]}, N, \beta)$ is the canonical partition function

$$Z_\varepsilon(\Lambda^{[3]}, N, \beta) = (N!)^{-1} \int_{\Gamma^{[3]}(N)} \exp(-\beta H_\varepsilon)\, d\tau \tag{2.8}$$

The measure $d\tau$ is the usual Cartesian product measure on $\Gamma^{[3]}(N)$, and β is the inverse temperature, essentially. Since (2.7) factorizes into the momentum part and the configurational part, and since (2.7) yields a (trivial) Gaussian measure over the momentum subspace $\mathbb{R}^{3N}$ of $\Gamma^{[3]}(N)$ which is independent of ε, it suffices to restrict the considerations to the configurational subspace $\Omega(N)\subset\Gamma^{[3]}(N)$, with $\Omega(N)=\mathsf{X}_{k=1}^{N}\Lambda_k^{[3]}$, where $\Lambda^{[3]}$ with subscript k denotes the domain of particle k. The canonical probability measure on $\Omega(N)$ is given by

$$\mu_\varepsilon^{(N)}(d\omega)=Q_\varepsilon^{-1}\exp\left[-\frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{N}\beta V(|\mathbf{r}_i-\mathbf{r}_j|,\varepsilon)\right]\lambda(d\omega^{(N)}) \tag{2.9}$$

where the configurational integral Q_ε reads

$$Q_\varepsilon(\Lambda^{[3]},N,\beta)=\int_{\Omega(N)}\exp\left[-\frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{N}\beta V(|\mathbf{r}_i-\mathbf{r}_j|,\varepsilon)\right]\lambda(d\omega^{(N)}) \tag{2.10}$$

with

$$d\omega^{(N)}=\prod_{k=1}^{N}d^3r_k \tag{2.11}$$

and

$$\lambda(d\omega^{(N)})=\prod_{k=1}^{N}\exp[-\beta m\phi(\mathbf{r}_k)]\,d^3r_k \tag{2.12}$$

We would like to know (2.9) in the limit $\varepsilon\to 0$ for $\beta>0$.

First of all, since Q_ε diverges to $+\infty$ as $\varepsilon\to 0$, it is clear that the density of (2.9) tends to zero wherever $|\mathbf{r}_i-\mathbf{r}_j|\neq 0$ for all i and j. Thus, the interesting points ω in $\Omega(N)$ are those where $|\mathbf{r}_i-\mathbf{r}_j|=0$ for at least one pair (i, j), since $\exp(-\beta V)$ diverges to $+\infty$ then, too, in the limit $\varepsilon\to 0$.

For the following investigations it is convenient to single out the coordinates of any two particles, here the Nth and the $(N-1)$th. We shall use the abbreviations $\mathbf{r}_N=\mathbf{r}$ and $\mathbf{r}_{N-1}=\mathbf{r}'$.

It suffices to consider the density of (2.9) with respect to $d\omega^{(N)}$, denoted by $g_N(\mathbf{r},\mathbf{r}',\tilde{\omega};\varepsilon)$, where $\tilde{\omega}\in\Omega(N-2)$ [we shall sometimes simply write $g_N(\omega,\varepsilon)$ when it is not necessary to distinguish between the particle coordinates]. Explicitly, the density reads

$$\begin{aligned}g_N(\mathbf{r},\mathbf{r}',\tilde{\omega};\varepsilon)={}&Q_\varepsilon^{-1}(\Lambda^{[3]},N,\beta)\\ &\times\exp\{-\beta[V(|\mathbf{r}-\mathbf{r}'|,\varepsilon)+m\phi(\mathbf{r})+m\phi(\mathbf{r}')]\}\\ &\times\exp\left\{-\sum_{i=1}^{N-2}\beta[V(|\mathbf{r}-\mathbf{r}_i|,\varepsilon)+V(|\mathbf{r}'-\mathbf{r}_i|,\varepsilon)+m\phi(\mathbf{r}_i)]\right\}\\ &\times\exp\left[-\frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{N-2}\beta V(|\mathbf{r}_i-\mathbf{r}_j|,\varepsilon)\right]\end{aligned} \tag{2.13}$$

The right-hand side of (2.13) is abbreviated by introducing a positive function h such that

$$g_N(\mathbf{r}, \mathbf{r}', \tilde{\omega}; \varepsilon) = Q_\varepsilon^{-1} \exp[-\beta V(|\mathbf{r}-\mathbf{r}'|, \varepsilon)]\, h(\mathbf{r}, \mathbf{r}', \tilde{\omega}; \varepsilon) \tag{2.14}$$

We concentrate first on the case $\mathbf{r} \neq \mathbf{r}'$. We shall now prove

$$\lim_{\varepsilon \to 0} g_N(\mathbf{r}, \mathbf{r}', \tilde{\omega}; \varepsilon) = 0; \qquad \mathbf{r} \neq \mathbf{r}' \tag{2.15}$$

pointwise.

Proof. For finite ε we have the estimate

$$g_N(\mathbf{r}, \mathbf{r}', \tilde{\omega}; \varepsilon) \leqslant Q_\varepsilon^{-1} \exp[-\beta V(|\mathbf{r}-\mathbf{r}'|, \varepsilon)]\, \hat{h}_\varepsilon(\mathbf{r}, \mathbf{r}') \tag{2.16}$$

where

$$\hat{h}_\varepsilon(\mathbf{r}, \mathbf{r}') \equiv \sup_{\tilde{\omega}} [h(\mathbf{r}, \mathbf{r}', \tilde{\omega}; \varepsilon)] \tag{2.17}$$

Since $\lim_{\varepsilon \to 0} \exp[-\beta V(|\mathbf{r}-\mathbf{r}'|, \varepsilon)]$ is finite for $|\mathbf{r}-\mathbf{r}'| \neq 0$, to prove (2.15) it suffices to show

$$\lim_{\varepsilon \to 0} Q_\varepsilon^{-1} \hat{h}_\varepsilon(\mathbf{r}, \mathbf{r}') = 0 \tag{2.18}$$

for $\mathbf{r} \neq \mathbf{r}'$. We estimate $\hat{h}_\varepsilon(\mathbf{r}, \mathbf{r}')$ from above. By construction, $V(\xi, \varepsilon)$ is bounded from below, with the infimum given by

$$\inf_{\xi} V(\xi, \varepsilon) = V(0, \varepsilon) < 0 \tag{2.19}$$

Furthermore, since ϕ is $C^\infty(\overline{\Lambda^{[3]}})$, there exist real numbers ϕ_+ and ϕ_- such that

$$\phi_- \leqslant \phi \leqslant \phi_+ \tag{2.20}$$

in $\Lambda^{[3]}$. We then obtain the strict inequality

$$\hat{h}_\varepsilon(\mathbf{r}, \mathbf{r}') < \exp\{[(N+1)(N-2)/2]\, \beta\, |V(0, \varepsilon)| - \beta N m \phi_-\} \tag{2.21}$$

for $\mathbf{r} \neq \mathbf{r}'$. The right-hand side of (2.21) is independent of $\mathbf{r}$, $\mathbf{r}'$, and will be abbreviated as

$$\text{r.h.s.}(2.21) \equiv \tilde{h}_\varepsilon \tag{2.22}$$

Inequality (2.21) implies that (2.18) is fulfilled if

$$\lim_{\varepsilon \to 0} Q_\varepsilon^{-1} \tilde{h}_\varepsilon = 0 \tag{2.23}$$

We now construct a lower bound of Q_ε which is sufficiently strong to prove (2.23).

Let μ denote any permutation-invariant probability measure on $\Omega(N)$, with density f_N (with respect to Lebesgue measure). We define the entropy of μ, denoted by $S(\mu)$, by

$$S(\mu) \equiv -\int_{\Omega(N)} f_N \ln f_N \, d\omega^{(N)} \tag{2.24}$$

and the corresponding free energy $F(\mu)$ by

$$F(\mu) \equiv \mu(H'_\varepsilon) - \beta^{-1} S(\mu) \tag{2.25}$$

where H'_ε is defined in (2.1), and $\mu(H'_\varepsilon)$ is the energy of μ. Our starting point is then the well-known fact that, for given β, the configurational free energy $F^{\text{conf}} = -\beta^{-1} \ln Q_\varepsilon$ pertaining to (2.10) obeys the inequality

$$F^{\text{conf}} \leqslant F(\mu) \tag{2.26}$$

equality holding only for μ given by (2.9). Obviously, (2.26) can immediately be translated into a lower bound for Q_ε upon choosing a convenient μ and then evaluating the right-hand side of (2.26).

Let μ be a product measure, with the density given by

$$f_N(\omega) = \prod_i^N \rho(\mathbf{r}_i) \tag{2.27}$$

Then $S(\mu)$ can be written as an entropy functional $\mathscr{S}(\rho)$, given by

$$\mathscr{S}(\rho) = -N \int_{\Lambda^{[3]}} \rho(\mathbf{r}) \ln \rho(\mathbf{r}) \, d^3r \tag{2.28}$$

Similarly, the energy $\mu(H'_\varepsilon)$ becomes the energy functional

$$\mathscr{E}(\rho) = [N(N-1)/2] \int_{\Lambda^{[3]} \times \Lambda^{[3]}} \rho(\mathbf{r})\, \rho(\mathbf{r}')\, V(|\mathbf{r}-\mathbf{r}'|, \varepsilon)\, d^3r\, d^3r'$$

$$+ N \int_{\Lambda^{[3]}} \rho(\mathbf{r})\, m\phi(\mathbf{r})\, d^3r \tag{2.29}$$

which leads to the free-energy functional

$$\mathscr{F}(\rho) = \mathscr{E}(\rho) - \beta^{-1} \mathscr{S}(\rho) \tag{2.30}$$

We now choose

$$\rho = \rho_0 \chi_{[B_{s,\mathbf{r}_0}]} \tag{2.31}$$

where $\rho_0 \in \mathbb{R}^+$ is a constant, and $B_{s,\mathbf{r}_0}$ is the three-dimensional ball with center $\mathbf{r}_0 \in \Lambda^{[3]}$ and radius s. For simplicity we require $B_{s,\mathbf{r}_0} \subset\subset \Lambda^{[3]}$ (proper subset). Then χ is the usual characteristic function of set theory. Thus,

$$\rho_0 = [(4/3)\,\pi s^3]^{-1} \tag{2.32}$$

[In order to evaluate $\mathscr{S}(\rho)$ for ρ given by (2.31), we have to consider a sequence of strictly positive probability densities with limit given by the right-hand side of (2.31).] Abbreviating the right-hand side of (2.31) by ρ^B, we then obtain

$$\begin{aligned}\mathscr{F}(\rho^B) &= [N(N-1)/2] \int_{B_{s,\mathbf{r}_0}} \int_{B_{s,\mathbf{r}_0}} V(|\mathbf{r}-\mathbf{r}'|, \varepsilon)\, \rho_0^2\, d^3r\, d^3r' \\ &\quad + N \int_{B_{s,\mathbf{r}_0}} \rho_0 m\phi(\mathbf{r})\, d^3r + N\beta^{-1} \ln \rho_0 \\ &\leqslant [N(N-1)/2]\; V(2s, \varepsilon) + Nm\phi_+ + N\beta^{-1} \ln \rho_0 \end{aligned} \tag{2.33}$$

Upon multiplying (2.33) by $-\beta$ and taking the exponential, we obtain an explicit lower bound for Q_ε. We multiply this bound for Q_ε by $\tilde{h}_\varepsilon^{-1}$. Furthermore, since (2.33) is valid for any choice of s, we now choose

$$s = \varepsilon s_0 \tag{2.34}$$

By Eq. (2.6) we have

$$V(0, \varepsilon) = \varepsilon^{-1} V(0, 1) \tag{2.35a}$$

and

$$V(2\varepsilon s_0, \varepsilon) = \varepsilon^{-1} V(2s_0, 1) \tag{2.35b}$$

In total we obtain

$$\begin{aligned} Q_\varepsilon \tilde{h}_\varepsilon^{-1} &> [(4/3)\,\pi s_0^3]^N \exp[-\beta N m(\phi_+ - \phi_-)]\, \varepsilon^{3N} \\ &\quad \times \exp\{\varepsilon^{-1}\beta[N(N-1)\,|V(2s_0, 1)| - (N+1)(N-2)\,|V(0, 1)|]/2\} \end{aligned} \tag{2.36}$$

Now, it follows from the proposed properties of $V(\xi, \varepsilon)$ that $|V(2s_0, 1)|$ is a positive, continuous, and decreasing (a.e.) function of s_0, with maximum value given by $|V(0, 1)|$. That guarantees the existence of a positive s^* such that

$$|V(2s_0, 1)| > [1 - 2/N(N-1)]\,|V(0, 1)| \qquad \text{for} \quad s_0 < s^* \tag{2.37}$$

Therefore, choose s_0 such that (2.37) is fulfilled. Then

$$N(N-1)\,|V(2s_0, 1)| - (N+1)(N-2)\,|V(0, 1)| \equiv 2Y > 0 \tag{2.38}$$

This implies

$$\varepsilon^{3N} \exp(\varepsilon^{-1}\beta Y) \xrightarrow{\varepsilon \searrow 0} +\infty \tag{2.39}$$

for $\beta > 0$, which proves (2.23) and, hence, (2.15). ∎

Remark. A corollary of (2.39) is that $\mathscr{F}(\rho^B)$ as given by (2.33) tends to minus infinity as $\varepsilon \to 0$. We shall come back to this point in the next two sections.

Result (2.15) means that $w^*\text{-}\lim_{\varepsilon \to 0} g_N(\mathbf{r}, \mathbf{r}', \tilde{\omega}; \varepsilon)$ is proportional to the Dirac distribution $\delta(\mathbf{r} - \mathbf{r}')$. This follows (e.g., ref. 24) immediately from the fact that $g_N(\mathbf{r}, \mathbf{r}', \tilde{\omega}; \varepsilon)$ is a probability density, i.e., a distribution of order zero that is normalized to one,

$$\int_{\Omega(N)} g_N(\omega; \varepsilon)\, d\omega^{(N)} = 1 \tag{2.40}$$

independently of ε. The general solution to [(2.15), (2.40)] in the space of positive measures is

$$\lim_{\varepsilon \to 0} g_N(\mathbf{r}, \mathbf{r}', \tilde{\omega}; \varepsilon) = \hat{G}(\mathbf{r}, \tilde{\omega})\, \delta(\mathbf{r} - \mathbf{r}') \tag{2.41}$$

(in the w^*-topology), where $\hat{G}(\mathbf{r}, \tilde{\omega})$ is a probability density which is independent of $\mathbf{r}'$. The probability density $\hat{G}(\mathbf{r}, \tilde{\omega})$ can immediately be evaluated further by noting that the choice $\mathbf{r} = \mathbf{r}_N$, $\mathbf{r}' = \mathbf{r}_{N-1}$ is ambiguous, since (2.9) is invariant against permutations of the particle indices, which holds also in the limit $\varepsilon \to 0$. Hence

$$\hat{G}(\mathbf{r}, \tilde{\omega}) = G(\mathbf{r}) \prod_{i=1}^{N-2} \delta(\mathbf{r} - \mathbf{r}_i) \tag{2.42}$$

where $G(\mathbf{r})$ is a probability density depending only on $\mathbf{r}$. Note that [(2.41), (2.42)] is permutation invariant, although not written in a form that displays this symmetry explicitly. A manifestly invariant form is readily obtained through switching back to the previous notation, i.e., $\mathbf{r} \to \mathbf{r}_N$, $\mathbf{r}' \to \mathbf{r}_{N-1}$, and rewriting [(2.41), (2.42)] in the equivalent form

$$w^*\text{-}\lim_{\varepsilon \to 0} g_N(\omega; \varepsilon) = \int_{\Lambda^{[3]}} d^3 r_c \, G(\mathbf{r}_c) \prod_{i=1}^{N} \delta(\mathbf{r}_c - \mathbf{r}_i) \tag{2.43}$$

The right-hand side of (2.43) describes a Gibbs ensemble of N-particle systems, each of which has collapsed to a single material point, as argued in the introduction. We go one step further and determine $G(\mathbf{r}_c)$.

The distribution $G(\mathbf{r}_c)$ has the obvious interpretation of being the probability density that a system has collapsed at the point $\mathbf{r}_c \in \Lambda^{[3]}$. Clearly, $G(\mathbf{r}_c)$ will depend on the geometry of $\Lambda^{[3]}$ and on external conditions, e.g., on the details of the superimposed external gravitational field. We shall now prove

$$G(\mathbf{r}_c) = \frac{\exp[-\beta N m \phi(\mathbf{r}_c)]}{\int_{\Lambda^{[3]}} \exp[-\beta N m \phi(\mathbf{r})] \, d^3 r} \tag{2.44}$$

which, together with (2.43), verifies (1.1).

Proof. Because of the equivalence between (2.43) and [(2.41), (2.42)] (note that $\mathbf{r} = \mathbf{r}_N$ and $\mathbf{r}' = \mathbf{r}_{N-1}$ in [(2.41), (2.42)]), it suffices to consider now $g_N(\omega; \varepsilon)$, as $\varepsilon \to 0$, in the case that all the particle coordinates coincide. We set $\mathbf{r}_i = \mathbf{r}$, $i = 1,..., N$, in the numerator of (2.13). The probability density $g_N(\omega; \varepsilon)$ in this case takes the value

$$g_N(\omega; \varepsilon) = Q_\varepsilon^{-1} \exp[-\beta N m \phi(\mathbf{r})] \exp\{\varepsilon^{-1} [N(N-1)/2] \, \beta \, |V(0, 1)|\} \tag{2.45}$$

which already has the desired explicit dependence on $\mathbf{r}$. Furthermore, the product of Q_ε^{-1} with $\exp\{[N(N-1)/2] \beta \, |V(0, \varepsilon)|\}$, which is independent of $\mathbf{r}$, blows up as $\varepsilon \to 0$, as should be the case. This can be shown by assuming that it does not blow up and then proving a contradiction with (2.40) upon using (2.15). Thus, the claim (2.44) follows if we can show that

$$Q_\varepsilon \to \int_{\Lambda^{[3]}} \exp[-\beta N m \phi(\mathbf{r})] \, d^3 r \, f(\varepsilon)[1 + O(0^+)] \tag{2.46}$$

as $\varepsilon \to 0^+$, where $f(\varepsilon)$ does not depend functionally on ϕ.

For the evaluation of Q_ε as $\varepsilon \to 0^+$ we single out again the position of the Nth particle and identify $\mathbf{r}_N \to \mathbf{r}$. In the following we shall also write $\Omega(N) = \Lambda^{[3]} \times \Omega(N-1)$, in order to emphasize $\mathbf{r} \in \Lambda^{[3]}$ and $(\mathbf{r}_1,..., \mathbf{r}_{N-1}) \equiv \hat{\omega} \in \Omega(N-1)$. Let $\mathscr{K}_\delta \subset\subset \Lambda^{[3]}$ be a compact and simply connected three-dimensional strict subdomain of $\Lambda^{[3]}$, with volume $|\mathscr{K}_\delta|$. We require

$$\operatorname{dist}(\mathscr{K}_\delta, \partial \Lambda^{[3]}) = \delta \ll \Lambda \tag{2.47}$$

uniformly. Assume $\mathbf{r} \in \mathscr{K}_\delta$ and let $B_{\delta,\mathbf{r}}$ be the three-dimensional open ball with radius δ, center $\mathbf{r}$, and closure $\overline{B_{\delta,\mathbf{r}}}$. Then $B_{\delta,\mathbf{r}} \subset\subset \Lambda^{[3]}$. For all particles $i = 1, \ldots, N-1$ we can decompose their domain of integration, i.e., $\Lambda_i^{[3]}$, as

$$\Lambda_i^{[3]} = B_{\delta,\mathbf{r};i} \cup \Lambda_i^{[3]} \setminus B_{\delta,\mathbf{r};i} \tag{2.48}$$

where the additional subscript i on $\Lambda^{[3]}$ and $B_{\delta,\mathbf{r}}$ denotes that it is the domain of particle i. By means of (2.48) we then obtain the decomposition

$$\Omega(N-1) = \bigcup_{k=0}^{N-1} \binom{N-1}{k} \Xi^{(k)} \tag{2.49}$$

with

$$\Xi^{(k)} = \left(\mathop{\times}_{i=1}^{k} B_{\delta,\mathbf{r};i} \right) \times \left(\mathop{\times}_{j=k+1}^{N-1} \Lambda_j^{[3]} \setminus B_{\delta,\mathbf{r};j} \right) \tag{2.50}$$

This allows us to write

$$Q_\varepsilon = \lim_{\mathscr{K}_\delta \nearrow \overline{\Lambda^{[3]}}} \sum_{k=0}^{N-1} \binom{N-1}{k} \int_{\mathscr{K}_\delta \times \Xi^{(k)}} q(\omega; \varepsilon)\, d\omega^{(N-1)}\, d^3 r \tag{2.51}$$

where $q(\omega; \varepsilon)$ is the integrand of Q_ε. The right-hand side of (2.51) has a simple interpretation, which is readily illustrated by means of the following scheme (the term "o-particle" refers here to particles *other* than the Nth one, i.e., belonging to the subset $i = 1, \ldots, N-1$):

$$\begin{aligned}
\text{r.h.s.}(2.51) \sim \Bigg[& \int \{\text{no} \quad o\text{-particles in } B_{\delta,\mathbf{r}}\} \\
& + (\#\ \text{realizations}) \times \int \begin{Bmatrix} 1 \quad o\text{-particle in } B_{\delta,\mathbf{r}} \\ N-2 \quad o\text{-particles outside} \end{Bmatrix} \\
& \vdots \\
& + (\#\ \text{realizations}) \times \int \begin{Bmatrix} k \quad o\text{-particles in } B_{\delta,\mathbf{r}} \\ N-1-k \quad o\text{-particles outside} \end{Bmatrix} \\
& \vdots \\
& + \int \{\text{all} \quad o\text{-particles in } B_{\delta,\mathbf{r}}\} \Bigg]
\end{aligned}$$

We now show that for any given δ the term $\int \{\text{all } o\text{-particles in } B_{\delta,\mathbf{r}}\}$ gives the leading contribution as $\varepsilon \to 0^+$.

To prove this assertion, we estimate $q(\omega;\varepsilon)$ for $\hat{\omega}\in\Xi^{(k)}$ from above, for all $k\leqslant N-2$. It suffices to take into account that $N-1-k$ particles are located outside $B_{\delta,\mathbf{r}}$, which means that the minimum value of the interaction energy of each of these particles with the Nth one, located at $\mathbf{r}$, is $V(\delta,\varepsilon)$. Replacing the other interaction energies by $V(0,\varepsilon)$ and estimating the external potential with the aid of (2.20), we find that for $\hat{\omega}\in\Xi^{(k)}$

$$q(\omega;\varepsilon)<\exp(-\beta\{Nm\phi_{-}+(N-1-k)\,V(\delta,\varepsilon)\\ +[k+(N-1)(N-2)/2]\,V(0,\varepsilon)\}) \tag{2.52}$$

Abbreviating the right-hand side of (2.52) by $K^{(k)}_{\varepsilon,\delta}$, we get

$$\int_{\mathscr{K}_\delta\times\Xi^{(k)}} q(\omega;\varepsilon)\,d\omega^{(N-1)}\,d^3r<|\mathscr{K}_\delta|\,(\tfrac{4}{3}\pi\delta^3)^k\,(\mathscr{V}-\tfrac{4}{3}\pi\delta^3)^{N-1-k}\,K^{(k)}_{\varepsilon,\delta}\\ <\mathscr{V}^N K^{(k)}_{\varepsilon,\delta} \tag{2.53}$$

As next step we seek the maximum of $K^{(k)}_{\varepsilon,\delta}$ with respect to $k\in\{0,\ldots,N-2\}$. Since $V(\cdot,\varepsilon)$ is negative and strictly increasing a.e., $K^{(k)}_{\varepsilon,\delta}$ is strictly increasing with k. Thus,

$$\sup_{k\in\{0,\ldots,N-2\}}[K^{(k)}_{\varepsilon,\delta}]=K^{(N-2)}_{\varepsilon,\delta}\\ =\exp(-\beta\{Nm\phi_{-}+V(\delta,\varepsilon)\\ +[(N+1)(N-2)/2]\,V(0,\varepsilon)\}) \tag{2.54}$$

We now have to estimate the contribution from $\Xi^{(N-1)}$ in (2.51) from below. Let $s<\delta$. Then $B_{s,\mathbf{r}}\subset\subset B_{\delta,\mathbf{r}}$. Let

$$\Theta=\mathop{\times}_{i=1}^{N-1}B_{s,\mathbf{r};i} \tag{2.55}$$

Then

$$\int_{\mathscr{K}_\delta\times\Xi^{(N-1)}} q(\omega;\varepsilon)\,d\omega^{(N-1)}\,d^3r\\ >\int_{\mathscr{K}_\delta\times\Theta} q(\omega;\varepsilon)\,d\omega^{(N-1)}\,d^3r\\ >|\mathscr{K}_\delta|\,(\tfrac{4}{3}\pi s^3)^{N-1}\exp[-\beta Nm\phi_{+}-\tfrac{1}{2}N(N-1)\,\beta V(2s,\varepsilon)] \tag{2.56}$$

where the last inequality again follows from (2.20) and from the proposed properties of V.

We now factorize out the contribution from $k=N-1$ in the sum (2.51). We write

$$Q_\varepsilon = \lim_{\mathscr{K}_\delta \nearrow \overline{\Lambda^{[3]}}} \int_{\mathscr{K}_\delta \times \Xi^{(N-1)}} q(\omega; \varepsilon)\, d\omega^{(N-1)}\, d^3r\, (1 + \mathscr{R}_{\varepsilon,\delta}) \tag{2.57}$$

Upon using the estimates (2.53), (2.54), and (2.56), we see that $\mathscr{R}_{\varepsilon,\delta}$ is bounded from above by a $B_{\varepsilon,\delta,s}$, given by

$$B_{\varepsilon,\delta,s} = b \exp\left[\frac{N(N-1)}{2} \beta V(2s, \varepsilon) - \beta V(\delta, \varepsilon) - \frac{(N+1)(N-2)}{2} \beta V(0, \varepsilon)\right] \tag{2.58}$$

where b is a harmless factor that depends on ε at most through a power law. Next we consider the sequence $\mathscr{K}_\delta \nearrow \overline{\Lambda^{[3]}}$ as $\varepsilon \to 0$, which due to (2.47) means $\delta \to 0$ as $\varepsilon \to 0$. We choose $\delta = \varepsilon\delta_0$ with $\delta_0 > 0$ fixed. We then have to show that there exists an $s_0 > 0$ such that with $s = \varepsilon s_0$ we have $\mathscr{R}_{\varepsilon,\delta} \to 0$ as $\varepsilon \to 0$.

Making use of (2.6) in (2.58), we see that the right-hand side of (2.58) and, hence, $\mathscr{R}_{\varepsilon,\delta}$ tend to zero if there exists $s_0 > 0$ such that

$$N(N-1)\, V(2s_0, 1) - 2V(\delta_0, 1) - (N+1)(N-2)\, V(0, 1) < 0 \tag{2.59}$$

Such an s_0 always exists for every positive δ_0. To show this, for given $\delta_0 > 0$ we replace s_0 by 0 in the left-hand side of (2.59) and find

$$\text{l.h.s.}(2.59)|_{s_0=0} = -2[V(\delta_0, 1) - V(0, 1)] < 0 \tag{2.60}$$

by (2.19). Since $V(\cdot, \varepsilon)$ is C^0 and strictly increasing a.e., there exists a finite right neighborhood $]0, s_*(\delta_0)[$ of 0 such that (2.59) holds for any $s_0 \in\,]0, s_*(\delta_0)[$. Hence,

$$\lim_{\varepsilon \to 0} \mathscr{R}_{\varepsilon, \varepsilon\delta_0} = 0 \tag{2.61}$$

The remaining step in the proof of (2.44) is now to evaluate the leading functional dependence on ϕ of

$$\int_{\mathscr{K}_\delta \times \Xi^{(N-1)}} q(\omega; \varepsilon)\, d\omega^{(N-1)}\, d^3r$$

as $\mathscr{K}_\delta \nearrow \overline{\Lambda^{[3]}}$ with $\varepsilon \to 0^+$. Since ϕ is C^∞, for $\mathbf{r}_i \in B_{\delta,\mathbf{r};i}$, $i = 1,\ldots, N-1$, we have

$$\phi(\mathbf{r}_i) = \phi(\mathbf{r}) + O(\delta) \tag{2.62}$$

Thus, noting that $\delta \to 0$ as $\varepsilon \to 0$, we obtain

$$\lim_{\mathscr{K}_\delta \nearrow \Lambda^{[3]}} \int_{\mathscr{K}_\delta \times \Xi^{(N-1)}} q(\omega; \varepsilon)\, d\omega^{(N-1)}\, d^3r$$

$$= \int_{\Lambda^{[3]}} \exp[-\beta N m\phi(\mathbf{r})]\, d^3r \lim_{\varepsilon \searrow 0} \int_{\Xi^{(N-1)}} q_0(\omega; \varepsilon)\, d\omega^{(N-1)}[1 + O(\varepsilon)] \quad (2.63)$$

where q_0 is obtained from q through setting $\phi \to 0$. Obviously, the integral $\int_{\Xi^{(N-1)}} q_0(\omega; \varepsilon)\, d\omega^{(N-1)}$ does not contain ϕ. We are also allowed to write that integral outside the integral over d^3r because it is independent of $\mathbf{r}$ also, which is easily seen to hold by means of the coordinate transformation $\mathbf{r}_i \to \mathbf{r}_i - \mathbf{r}$ for $i = 1,..., N-1$. We see that (2.63) is just (2.46). This proves the claim. ▌

Remark. Results (2.15) and, hence, (2.43) remain valid if we take the full Hamiltonian (2.1) into account. Result (2.44) has to be modified however, if we consider $W(\partial\Lambda^{[3]}, \mathbf{r}_i)$ not as being identically zero on all $\Lambda_i^{[3]}$ but only on a compact subdomain of $\Lambda_i^{[3]}$, increasing rapidly and continuously to infinity as $\mathbf{r}_i$ approaches $\partial\Lambda_i^{[3]}$. We then have to replace $m\phi(\mathbf{r}_c)$ by $m\phi(\mathbf{r}_c) + W(\partial\Lambda^{[3]}, \mathbf{r}_c)$ in (2.44).

We have thus verified our intuitively motivated expectation. The statistical mechanics equilibrium state of a classical self-gravitating system, computed from the canonical ensemble for $\beta > 0$, is a system that has collapsed to a single material point. The probability density of finding that point located at $\mathbf{r}_c$ in the considered domain $\Lambda^{[3]}$ is given by the Boltzmann-like factor (2.44). However, (2.44) is not the usual Boltzmann factor of a system of N point particles in an external field, which is apparent from the occurrence of N in the argument of the exponential functions.

For the sake of completeness, before considering the infinitely many-particle limit, we wish here to discuss briefly also the two-dimensional analogue of the finite systems. As mentioned in the introduction, the analogous one-dimensional problem was solved in ref. 2.

Remark. In the literature, the terminology as to what is meant by two-dimensional is not unique. There exist two convenient interpretations which are mathematically equivalent: (1) One thinks of these systems as consisting of infinitely long parallel wires in three-dimensional physical space. These wires take over the role of the particles. Since then the extensive quantities like mass and energy diverge trivially, one considers such quantities per unit length. The physical constants are the usual ones. (2) One considers a strictly two-dimensional model world, and assumes that some "two-dimensional" physicists had developed physics much along the same lines as we did. The concepts of length, time, mass force, etc., are then the same as in the three-dimensional world; however, some of the physical constants in this model world have to be redefined. Depending on the circumstances, it might be favorable to pick one or the other of these concepts. Here we like to make contact with the work presented in ref. 1, which uses the second concept, and which will be used here, too. To facilitate the comparison with ref. 1, the following is written in a dimensional manner.

We have to replace $\Lambda^{[3]}$ by a simply connected finite domain $\Lambda^{[2]} \subset \mathbb{R}^2$ with two-dimensional volume $\mathscr{A}$. We assume, for the moment, that the 2D point particles (of equal mass m) interact via exact classical 2D gravitational forces derived from the interaction energy

$$\tilde{V}(|\mathbf{r}-\mathbf{r}'|) = G^* m^2 \ln(|\mathbf{r}-\mathbf{r}'|/L) \tag{2.64}$$

where $\mathbf{r}$ and $\mathbf{r}'$ are in $\Lambda^{[2]}$, and L is a normalizing length. The constant G^* is the gravitational constant for the two-dimensional model world. Let $[\cdot]$ denote the physical dimension of the quantity to be inserted between the brackets. Then G^* is connected with the usual Newtonian gravitational constant via $[G^*] = [G]/[L]$ and $G^*/[G^*] = 2G/[G]$.[1]

We are interested in the configurational canonical ensemble on $\mathsf{X}_{i=1}^{N} \Lambda_i^{[2]}$, pertaining to the interaction Hamiltonian

$$\tilde{H}(N) = \frac{1}{2} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} \tilde{V}(|\mathbf{r}_i - \mathbf{r}_j|) \tag{2.65}$$

where the external gravitational field has been neglected, for simplicity. As mentioned in the introduction, it has been shown[1] that the configurational integral Q exists, for given particle number N, particle mass m, and given container $\Lambda^{[2]}$, only if the temperature T exceeds a critical value $T_c = (k_B \beta_c)^{-1}$, given by

$$T_c = (N-1) G^* m^2 / 4k_B \tag{2.66}$$

where k_B is the usual Boltzmann constant. For $T < T_c$, Q diverges and it has been conjectured[1] that this has to be interpreted as meaning the collapse of (each of) the systems in the ensemble to a single (2D) material point. Clearly, the inclusion of an external gravitational field would not alter this result.

Complementary to the results derived in ref. 1, we might wish to verify that collapse conjecture[1] by calculating the canonical equilibrium measure for $\beta > \beta_c$. The technique presented above to treat the three-dimensional systems is applicable, after minor modifications, to the two-dimensional systems as well. We have to replace (2.64) by

$$\tilde{V}(|\mathbf{r}-\mathbf{r}'|, \varepsilon) = G^* m^2 \ln(|\mathbf{r}-\mathbf{r}'|/L + \varepsilon) \tag{2.67}$$

for example, and then again to consider the limit $\varepsilon \searrow 0$. We might add also the contributions from external sources to the gravitational field. We then have to prove the analogue of (2.15). Following closely the steps of the proof of (2.15), we find, however, that we can prove the analogous result of (2.15) and, finally, of (1.1) only for $\beta > \beta_0$, with

$$\beta_0 = 2N/G^*m^2 \tag{2.68}$$

This means a verification of the collapse conjecture of ref. 1, but only for small enough temperatures. To investigate the intermediate temperature regime, we need sharper estimates. We have to leave open here the question of whether in the intermediate range $\beta_0 > \beta > \beta_c$ a system coalesces to a single point, too, or whether configurations with some particles located at a single point but others not are of finite measure, then, also.

3. THE LIMIT $N \to \infty$

Our goal is now to investigate the canonical probability measure in the limit of both $N \to \infty$ and $V \to V_{\text{cl}}$. For convenience we will confine ourselves to the discussion of the three-dimensional systems. We would like to show that in this double limit the statistical mechanics equilibrium state is the material point, independently of whether we take first the Newtonian interactions limit for finite systems and then let N go to infinity, or whether we take first $N \to \infty$ for smoothed-out interactions and then take the limit of Newtonian gravitational interactions.

In order to obtain a well-defined limit $N \to \infty$ we have to bear in mind that the particles interact via unstable interactions. In our case this means that the usual thermodynamic limit does not exist. The appropriate limit to investigate the canonical ensemble for $N \to \infty$ is a mean-field limit, which is obtained by means of suitably rescaling the interactions with N. For details of this limit for unstable but Lipschitz continuous interactions, we refer to ref. 15. For the related mean-field problem for the van der Waals gas see, e.g., refs. 25 and 26.

Remark. A rescaling of the interactions with N does not alter any of the finite-N results obtained in Section 2.

We shall drop the superscript [3] on $\Lambda^{[3]}$ for simplicity. We have $\lim_{N \to \infty} \Omega(N) = \Lambda^{\mathbb{N}}$, and we understand now $\omega \in \Lambda^{\mathbb{N}}$. We consider the probability measures of the finite systems as probability measures on $\Lambda^{\mathbb{N}}$.

We investigate first the limit $N \to \infty$ of (2.43). The upper limit N in the product of the delta distributions in (2.43) can simply be replaced by ∞, for $\omega \in \Lambda^{\mathbb{N}}$. The result is just (1.2). This means that the limiting measure again describes an ensemble of systems which have collapsed to a single material point each. Our problem here reduces to the study of $G(\mathbf{r}_c)$ as $N \to \infty$.

When studying this limit, setting W identically to zero inside $\bar{\Lambda}$ but to $+\infty$ on $\partial\Lambda$ no longer means a simplification. The reason is that ϕ, being a solution of Laplace's equation, takes its minimum value on $\partial\Lambda$. We shall need an external potential, however, which takes its global minimum on Λ. Therefore, we let W be zero only on a compact subdomain of Λ. It then increases monotonically and Lipschitz continuously, with Lipschitz constant L_W, when approaching $\partial\Lambda$. On $\partial\Lambda$ it jumps again to infinity. Instead of (2.44), for the finite systems we have to take

$$G_N(\mathbf{r}_c) = \frac{\exp[-\beta N \psi(\mathbf{r}_c)]}{\int_\Lambda \exp[-\beta N \psi(\mathbf{r})]\, d^3 r} \tag{3.1}$$

with

$$\psi(\mathbf{r}) = m\phi(\mathbf{r}) + W(\partial\Lambda, \mathbf{r}) \tag{3.2}$$

(We have added the subscript N to G in order to emphasize that it is a measure for finite systems.) The joint potential energy ψ is Lipschitz continuous on Λ. Moreover, it is always possible now to choose W such that ψ takes its global minimum value on a strict subset $\Upsilon \subset\subset \Lambda$. In the following we always assume that this is the case.

Remark. We shall need the Lipschitz continuity in order to take over directly some results of ref. 15. However, some of the results to be proved below remain valid if W satisfies weaker smoothness conditions.

It remains to specify the scaling of ψ with N. We would like to allow for various possibilities and assume

$$\psi = N^{a-1} \psi_0 \tag{3.3}$$

where a is a real number and ψ_0 is independent of N. As will become clear below, (3.3) covers the relevant cases. In the following we shall always understand ψ in (3.1) to be given by (3.3), unless otherwise stated. Accordingly, we shall write

$$\lim_{N \to \infty} \frac{\exp[-\beta N^a \psi_0(\mathbf{r}_c)]}{\int_\Lambda \exp[-\beta N^a \psi_0(\mathbf{r})]\, d^3 r} = G^{(a)}(\mathbf{r}_c) \tag{3.4}$$

There are then essentially three different limiting distributions $G^{(a)}(\mathbf{r}_c)$, depending on whether $a > 0$, $a < 0$, or $a = 0$. Let $\delta_{[\Upsilon]}$ be either the Dirac measure concentrated on Υ (in the case that Υ has zero Lebesgue measure) or (in the case that Υ has finite Lebesgue measure $|\Upsilon|$) be given by

$|\Upsilon|^{-1}\chi_{[\Upsilon]}$, where $\chi_{[\Upsilon]}$ is the characteristic function of Υ (combinations will be possible, too; they are not considered explicitly). We then have

$$G^{(a)}(\mathbf{r}_c)=\begin{cases}\mathscr{V}^{-1}, & a<0 \quad (3.5a)\\ \dfrac{\exp[-\beta\psi_0(\mathbf{r}_c)]}{\int_\Lambda \exp[-\beta\psi_0(\mathbf{r})]\,d^3r}, & a=0 \quad (3.5b)\\ \delta_{[\Upsilon]}, & a>0 \quad (3.5c)\end{cases}$$

Proof. The proofs of (3.5a) and (3.5b) are trivial. To prove (3.5c) we abbreviate N^a by p, and $\exp[-\beta\psi_0(\mathbf{r})]$ by $f_\psi(\mathbf{r})$. Then

$$\int_\Lambda \exp[-\beta N^a\psi_0(\mathbf{r})]\,d^3r=\|f_\psi\|_p^p \tag{3.6}$$

and (3.1), with (3.3), can be rewritten as

$$G_N(\mathbf{r}_c)=[f_\psi(\mathbf{r}_c)/\|f_\psi\|_p]^p \tag{3.7}$$

Consider now the case $\mathbf{r}_c\notin\Upsilon$. Then

$$f_\psi(\mathbf{r}_c)<\sup_{\mathbf{r}\in\Lambda} f_\psi(\mathbf{r}) \tag{3.8}$$

where strict inequality holds indeed. On the other hand, since $a>0$, p grows monotonically to infinity with N. Hence

$$\|f_\psi\|_p\to\|f_\psi\|_\infty=\text{ess sup}\, f_\psi=\sup f_\psi \tag{3.9}$$

as $N\to\infty$. Consequently, there exists a $p^*>0$ such that

$$f_\psi/\|f_\psi\|_p<1; \qquad p>p^* \tag{3.10}$$

Even stronger,

$$\lim_{p\to\infty} f_\psi/\|f_\psi\|_p<1 \tag{3.11}$$

strictly. By means of (3.7) this implies

$$G_N(\mathbf{r}_c)\xrightarrow{N}0; \qquad \mathbf{r}_c\notin\Upsilon \tag{3.12}$$

Now, since G_N is a probability density on Λ for each N, and since G_N does not vary with $\mathbf{r}_c$ for $\mathbf{r}_c\in\Upsilon$, our claim follows from standard theorems[24] of the theory of distributions. ∎

In total, when taking first the limit $V \to V_{\text{cl}}$ and then the limit $N \to \infty$, the density of the canonical probability measure converges to a limit density $g_\infty(\omega)$, given by

$$g_\infty(\omega) = \int_\Lambda d^3 r_c G^{(a)}(\mathbf{r}_c) \prod_{i=1}^{\infty} \delta(\mathbf{r}_c - \mathbf{r}_i) \tag{3.13}$$

where the three possibilities for the probability density $G^{(a)}$ are listed in (3.5). The meaning of the various scalings becomes clear from (3.5) and its proof: As $N \to \infty$ (always), the case $a < 0$ means that the whole system feels no external influences inside $\bar{\Lambda}$. This limit is equivalent to the situation where no external field is present from the beginning. The case $a = 0$ means a finite potential energy, in the external field, of the system as a whole. Finally, $a > 0$ corresponds to an infinite potential energy of the whole system in the field of the external sources; however, for $a = 1$ we have still a finite energy per particle due to ψ. (Clearly, the potential energy due to the pair interactions is always minus infinity, both for the system as a whole and when counted per particle.) The case $a = 1$ will be of primary interest below.

We now have to verify that (3.13) comes out also when we let $V \to V_{\text{cl}}$ after having performed the mean-field limit for the smoothed-out gravitational interactions. We take the interaction Hamiltonian $H'_\varepsilon(N)$ as given in (2.1) and consider the weak limit points, as $N \to \infty$, of the configurational canonical ensemble for fixed ε. We again set $\gamma = 0$, for simplicity. Since we are facing systems with unstable interactions but nevertheless would like to obtain thermodynamic behavior, we have to keep the total potential energy proportional to the particle number. As a consequence, we rescale the interaction energy V by a factor N^{-1}, meaning a weak ($\sim 1/N$) pair interaction. In addition, we would like the external sources to be of equal importance. Thus we have to keep $m\phi$ and W independent of N. This guarantees that as $N \nearrow \infty$ a particle feels mainly the self-consistent field generated by all the particles together, plus the contributions from the external sources. This mean-field scaling implies $a = 1$ in (3.3). Consequently, we have to verify (3.13) with $G^{(a)}$ given by (3.5c).

It may be desirable to have an explicit example. Consider the class of indefinitely differentiable interaction energies $V^{(n)}$, $n \in \mathbb{N}$, given by

$$V^{(n)}(\xi, \varepsilon) = -m^2(\xi^n + \varepsilon^n)^{-1/n} \tag{3.14}$$

Every member of this class fulfills all the requirements proposed for V in Section 2. From (3.14) we obtain a $1/N$ interaction rather naturally in the following way: Along the sequence $N \nearrow \infty$ we take particles of smaller and smaller mass such that m in (3.14) is to be replaced as $m \to N^{-1/2} m$. This implies also a rescaling of the external gravitational potential ϕ as $\phi \to N^{1/2}\phi$. We shall come back to the example (3.14) at some places below. However, we remark that all proofs given below do not require that V is differentiable.

We put these thoughts into explicit equations. In the following we understand

$$V(\xi, \varepsilon) = N^{-1}U(\xi, \varepsilon) \tag{3.15a}$$

in (2.1), where U is independent of N. Furthermore,

$$m = N^{-1/2}m_0 \tag{3.15b}$$

with m_0 fixed, and

$$\phi = N^{1/2}\phi_0 \tag{3.15c}$$

with ϕ_0 fixed, but with W fixed, too, such that ψ as given by (3.2) is independent of N also. Accordingly, the canonical probability measure on Λ^N for N particles is given by

$$\mu_\varepsilon^{(N)}(d\omega) = Q_\varepsilon^{-1} \exp\left[-\frac{1}{2N} \sum_{\substack{i,j=1\\ i\neq j}}^{N} \beta U(|\mathbf{r}_i - \mathbf{r}_j|, \varepsilon)\right] \lambda(d\omega^{(N)}) \tag{3.16}$$

with

$$Q_\varepsilon(\Lambda, N, \beta) = \int_{\Omega(N)} \exp\left[-\frac{1}{2N} \sum_{\substack{i,j=1\\ i\neq j}}^{N} \beta U(|\mathbf{r}_i - \mathbf{r}_j|, \varepsilon)\right] \lambda(d\omega^{(N)}) \tag{3.17}$$

where $\lambda(d\omega^{(N)})$ now reads

$$\lambda(d\omega^{(N)}) = \prod_{k=1}^{N} \exp[-\beta\psi(\mathbf{r}_k)]\, d^3 r_k \tag{3.18}$$

We understand (3.16) as a probability measure on $\Lambda^{\mathbb{N}}$. Then, as first step, for finite ε we determine the weak limit points of $\{\mu_\varepsilon^{(N)} \mid N = 1, 2,...\}$. As second step we investigate the limit $\varepsilon \to 0$ of these weak limit points.

We want to adopt here the results derived in ref. 15, where the mean-field limit has been established for unstable but Lipschitz continuous pair interactions. This requires a growth condition to be fulfilled by U, in addition to the postulates for V given in Section 2, which also hold for U when expressed in the rescaled quantities where necessary. We define the bounded positive difference function

$$U(\xi + \eta, \varepsilon) - U(\xi, \varepsilon) \equiv \Delta_U^+(\eta) \tag{3.19}$$

where the subscript U at Δ_U^+ should indicate that this function has to be evaluated for fixed ξ and ε. We postulate that for any $\xi \in \overline{\mathbb{R}^+}$

$$\Delta_U^+(\eta) \leqslant O(\eta) \tag{3.20}$$

where $\eta \in \mathbb{R}^+$.

Remark. Equations (3.19) and (3.20) are equivalent to the postulate that for given $\varepsilon > 0$, $U(\cdot, \varepsilon)$: $\overline{\mathbb{R}^+} \ni \xi \mapsto \mathbb{R}^-$ is Lipschitz continuous.

Then, for $\varepsilon \neq 0$, $U(|\mathbf{r}-\mathbf{r}'|, \varepsilon)$ is Lipschitz continuous on $\bar{\Lambda} \times \bar{\Lambda}$, i.e., there exists a positive Lipschitz constant L_U such that

$$|U(|\mathbf{r}-\mathbf{r}'|, \varepsilon) - U(|\tilde{\mathbf{r}}-\hat{\mathbf{r}}|, \varepsilon)| \leqslant L_U(|\mathbf{r}-\tilde{\mathbf{r}}| + |\mathbf{r}'-\hat{\mathbf{r}}|) \tag{3.21}$$

Proof. Since the left-hand side of (3.21) depends on the various position vectors only through $|\mathbf{r}-\mathbf{r}'|$ and $|\tilde{\mathbf{r}}-\hat{\mathbf{r}}|$, we can rearrange the four position vectors in many ways keeping both $|\mathbf{r}-\mathbf{r}'|$ and $|\tilde{\mathbf{r}}-\hat{\mathbf{r}}|$ and, thus, the left-hand side of (3.21) fixed. In particular, in that way we can minimize $|\mathbf{r}-\tilde{\mathbf{r}}| + |\mathbf{r}'-\hat{\mathbf{r}}|$ for any given fixed $|\mathbf{r}-\mathbf{r}'|$ and $|\tilde{\mathbf{r}}-\hat{\mathbf{r}}|$. Obviously, (3.21) is fulfilled for any four position vectors if it is fulfilled for any such minimizing configuration. It suffices to prove (3.21) for any convex domain $\bar{\Lambda}$, since any compact domain in $\mathbb{R}^3$ is contained in some other compact one which is also convex. It is an easy exercise to show that in a convex domain, $|\mathbf{r}-\tilde{\mathbf{r}}| + |\mathbf{r}'-\hat{\mathbf{r}}|$ is minimized with the constraint of fixed $|\mathbf{r}-\mathbf{r}'|$ and $|\tilde{\mathbf{r}}-\hat{\mathbf{r}}|$ if the four points are arranged collinearly, with the requirement that either the pair $\mathbf{r}$ and $\mathbf{r}'$ or the pair $\tilde{\mathbf{r}}$ and $\hat{\mathbf{r}}$ marks the endpoints of the collinear configuration. Clearly, the minimizing configurations include the case where $\mathbf{r} = \tilde{\mathbf{r}}$. This in turn implies that (3.21) holds for the minimizing configurations, and thus for all configurations, if it holds for any configuration for which $\mathbf{r} = \tilde{\mathbf{r}}$. So from now on we set $\mathbf{r} = \tilde{\mathbf{r}}$.

Since $\bar{\Lambda}$ is compact and U continuous, it suffices further to consider the case where the right-hand side of (3.21) tends to zero. Without loss of generality we assume

$$U(|\mathbf{r}-\mathbf{r}'|, \varepsilon) \leqslant U(|\mathbf{r}-\hat{\mathbf{r}}|, \varepsilon) \tag{3.22}$$

Now write $\hat{\mathbf{r}} = \mathbf{r}' + \mathbf{x}$. Then

$$\begin{aligned} U(|\mathbf{r}-\hat{\mathbf{r}}|, \varepsilon) &= U(|\mathbf{r}-\mathbf{r}'-\mathbf{x}|, \varepsilon) \\ &\leqslant U(|\mathbf{r}-\mathbf{r}'| + |\mathbf{x}|, \varepsilon) = U(|\mathbf{r}-\mathbf{r}'|, \varepsilon) + \Delta_U^+(|\mathbf{x}|) \end{aligned} \tag{3.23}$$

where the inequality is obtained by means of the triangle inequality and because $U(\cdot, \varepsilon)$ is monotonically increasing. The last equation in (3.23) is a consequence of (3.19). Together with (3.22) this gives

$$|U(|\mathbf{r}-\mathbf{r}'|, \varepsilon) - U(|\mathbf{r}-\hat{\mathbf{r}}|, \varepsilon)| \leqslant \Delta_U^+(|\mathbf{x}|) \tag{3.24}$$

Now the claim follows from (3.20). ∎

We are now ready to formulate the equations that determine the canonical ensemble in the mean-field limit for finite ε. These equations are obtained following closely the steps in ref. 15; however, we include here also the external field ψ. For details of the derivation see ref. 15.

Let $\mathcal{M}$ be the set of all probability measures ϱ on Λ. By μ_ϱ we denote the associated product measure $\varrho \otimes \varrho \otimes \cdots$ on Ω. Then, according to the Hewitt–Savage representation theorem,[18] any permutation-invariant probability measure μ on Ω can be written (in the notation of ref. 15) as

$$\mu = \int_{\mathcal{M}} \nu(d\varrho \mid \mu)\, \mu_\varrho \tag{3.25}$$

where $\nu(d\varrho \mid \mu)$ is a uniquely defined probability measure (ref. 18, Theorem 9.4) on $\mathcal{M}$. Equation (3.25) states that every permutation-invariant probability measure is an average over product measures. In our case we have $d\mu = g(\omega)\, d\omega^{(\infty)}$. Accordingly, we write $d\varrho = \rho(\mathbf{r})\, d^3r$. Let μ_ε be any weak limit point of $\{\mu_\varepsilon^{(N)} \mid N = 1, 2,...\}$ as given in (3.16). The canonical ensemble is then determined by the following statement about $\nu(d\varrho \mid \mu_\varepsilon)$: Let the free-energy functional

$$f(\rho) = (1/2) \int_\Lambda \int_\Lambda \rho(\mathbf{r})\, \rho(\mathbf{r}')\, U(|\mathbf{r}-\mathbf{r}'|, \varepsilon)\, d^3r\, d^3r' + \int_\Lambda \rho(\mathbf{r})\, \psi(\mathbf{r})\, d^3r + \beta^{-1} \int_\Lambda \rho(\mathbf{r}) \ln \rho(\mathbf{r})\, d^3r \tag{3.26}$$

be defined on $\mathcal{M}$ as the extension of the same functional considered on $C^\infty_{+,1}(\Lambda)$, meaning the positive C^∞ functions on Λ with integral equal to 1. The functional (3.26) is bounded from below as long as $\varepsilon \neq 0$, and it takes its infimum. By $\mathcal{M}' \subset \mathcal{M}$ we denote the subset of probabilities ϱ on Λ with density ρ for which the free-energy functional $f(\rho)$ takes its global minimum on $\mathcal{M}$. Then $\nu(d\varrho \mid \mu_\varepsilon)$ is concentrated on $\mathcal{M}'$.

Proof. The proof of the above statement is a straightforward extension of Section 2 of ref. 15. (In the following any cited lemma or theorem refers to ref. 15.) In fact, since both U and ψ are Lipschitz continuous on Λ and $\bar{\Lambda}$ is compact, U and ψ are bounded on Λ also. Then analogues of Lemmas 1 and 2 can be proved by slight modifications of the proofs given there. Lemmas 3 and 4 hold with unchanged proofs. Lemma 5 has to be modified accordingly so as to contain the external contributions to the energy due to ψ. Finally, Theorems 1 and 2 remain valid with unchanged proofs when we understand the free-energy functional to be given by our Eq. (3.26). ∎

Remark. The free energy per particle of any weak limit point of the sequence $\{\mu_\varepsilon^{(N)} \mid N = 1, 2, \ldots\}$ equals $\inf_{\varrho \in \mathcal{M}} f(\rho)$.

The subset $\mathcal{M}' \subset \mathcal{M}$ is also a subset of $\mathcal{M}^* \subset \mathcal{M}$, meaning the stationary points of (3.26) on $\mathcal{M}$. Not every member of $\mathcal{M}^*$ minimizes f. For $\varrho \in \mathcal{M}^*$ the first variation of (3.26) with respect to $\varrho \in \mathcal{M}$ has to vanish. This variational principle gives

$$\rho(\mathbf{r}) = \frac{\exp\{-\beta[\int_\Lambda \rho(\mathbf{r}')\, U(|\mathbf{r}-\mathbf{r}'|, \varepsilon)\, d^3r' + \psi(\mathbf{r})]\}}{\int_\Lambda \exp\{-\beta[\int_\Lambda \rho(\mathbf{r}')\, U(|\mathbf{r}''-\mathbf{r}'|, \varepsilon)\, d^3r' + \psi(\mathbf{r}'')]\}\, d^3r''} \tag{3.27}$$

for $\varrho \in \mathcal{M}^*$ (cf. ref. 15). To find the global minimizers of $f(\rho)$, knowledge on the type and number of solutions of (3.27) is of value. We come back to this point in the next section.

Remark. Upon inserting (3.27) into the argument of the logarithm in the entropy term of (3.26), we can bring (3.26) into an alternative form that might be more useful in certain circumstances. The free-energy functional takes then the form

$$\begin{aligned} \tilde{f}(\rho) = &-(1/2) \int_\Lambda \int_\Lambda \rho(\mathbf{r})\, \rho(\mathbf{r}')\, U(|\mathbf{r}-\mathbf{r}'|, \varepsilon)\, d^3r\, d^3r' \\ &- \beta^{-1} \ln \int_\Lambda \exp\left\{-\beta \left[\int_\Lambda \rho(\mathbf{r}')\, U(|\mathbf{r}-\mathbf{r}'|, \varepsilon)\, d^3r' + \psi(\mathbf{r})\right]\right\} d^3r \end{aligned} \tag{3.28}$$

The functional $\tilde{f}(\rho)$ is bounded from below (for finite ε) by the same bound as is $f(\rho)$; hence, it takes on the same infimum value. Also, every stationary point of $f(\rho)$ on $\mathcal{M}$ is a stationary point of $\tilde{f}(\rho)$ and, by construction, $\tilde{f}$ takes on the same value there as does f. It should be noted that, as distinct from the properties of $f(\rho)$, the functional $\tilde{f}(\rho)$ is invariant against the addition of an arbitrary element $\sigma \in \mathrm{Ker}_U$ to ρ, with $\mathrm{Ker}_U \equiv \mathrm{Ker}(\int U(|\mathbf{r}-\mathbf{r}'|, \varepsilon)(\cdot)\, d^3r')$. This holds irrespective of whether $\rho + \sigma$ is a probability density or not. Consequently, if $\mathrm{Ker}_U \neq \varnothing$, then all stationary "points" of (3.28) are degenerate. If in addition $\int \sigma\, d^3r = 0$ for a bounded element σ of Ker_U, then for some $t \in \mathbb{R}$, $\rho + t\sigma$ is a global minimizer $\in \mathcal{M}$ of $\tilde{f}$ if ρ is a global minimizer that solves (3.27). Thus, for $\mathrm{Ker}_U \neq \varnothing$ (which we do not want to exclude here, for the sake of generality) it is not possible to identify the thermodynamic equilibrium simply with the global minimizer(s) of $\tilde{f}$, since too many of them might exist. Thus, to determine the thermodynamic equilibrium state, ρ_0 say, from $\tilde{f}(\rho)$, one has to postulate that *in addition* to being a global minimizer of $\tilde{f}$, ρ_0 has to solve (3.27). In this sense (3.26) and (3.28) are not equivalent functionals, since every global minimizer of (3.26) necessarily solves (3.27).

We now have to show the following: As $\varepsilon \to 0$, which means $U \to U_{\mathrm{cl}} \equiv -m_0^2 |\mathbf{r}-\mathbf{r}'|^{-1}$, the free-energy functional given by (3.26) takes its infimum value if and only if $\rho_\varepsilon(\mathbf{r}) \rightharpoonup \delta(\mathbf{r}-\mathbf{r}_c)$, where ρ_ε means a global minimizer of $f(\rho)$ for $\varepsilon \neq 0$. We drop the subscript ε at μ_ε in the limit $\varepsilon = 0$. We have then further to show that $\nu(d\varrho \mid \mu)$ is concentrated on those Dirac point measures with support in Υ, with Υ as defined below (3.2). We prove this by means of a sequence of proofs of partial results.

It follows from the remark below (2.39) that $f(\rho)$ is unbounded below if $\varepsilon = 0$. Thus, to prove our assertion, we have first of all to show that when ε is zero, $f(\rho)$ goes to minus infinity along any sequence of C^∞ probability densities which converge to a single-point Dirac measure. To keep the proof short, we shall postulate a mild growth condition for the probability densities such that the sequence is not completely arbitrary.

Proof. Let $\{C_0^\infty(\Lambda) \ni \rho_{\mathbf{r}_0}^{(n)}(\mathbf{r}) \mid n = 1, 2,...\}$ be a sequence of positive C_0^∞ functions, normalized to 1, which converge to the Dirac point measure concentrated on $\mathbf{r}_0 \in \Lambda$. Without loss of generality we assume that the support of $\rho_{\mathbf{r}_0}^{(n)}$ is a convex neighborhood of $\mathbf{r}_0$ which is strictly contained in Λ. For simplicity let $\operatorname{supp} \rho_{\mathbf{r}_0}^{(n+1)} \subset\subset \operatorname{supp} \rho_{\mathbf{r}_0}^{(n)} \searrow \mathbf{r}_0$ monotonically. Let $\mathscr{V}_n$ be the Lebesgue measure of $\operatorname{supp} \rho^{(n)}$ and let $2l_n$ denote the diameter of the smallest ball $\subset \Lambda$ containing $\operatorname{supp} \rho^{(n)}$. Clearly, $\mathscr{V}_n \leqslant (4\pi/3)\, l_n^3$. Then, from the normalizing condition $\int \rho \, d^3r = 1$ we obtain the inequality

$$\|\rho^{(n)}\|_\infty \geqslant 1/\mathscr{V}_n \geqslant 3/(4\pi l_n^3) \tag{3.29}$$

Equality holds for $\rho^{(n)} = \rho^B$ as given by (2.31), (2.32), with $s = l_n$. In addition and consistent with (3.29), we now postulate that there exist positive constants M_1, and $M_2 > 3$, independently of n such that $\rho^{(n)}$ fulfills the growth restriction

$$\|\rho^{(n)}\|_\infty \leqslant M_1/l_n^{M_2} \tag{3.30}$$

Consider now the case $\varepsilon = 0$ for the free-energy functional (3.26), i.e., $U = U_{\mathrm{cl}}$. Since W was chosen to be Lipschitz continuous on Λ, ψ is bounded from above on Λ by some constant $\psi^+ \equiv \sup_{\mathbf{r} \in \Lambda} \psi$. (By construction, $\sup_{\mathbf{r} \in \Lambda \cup \partial\Lambda} \psi = +\infty$, but ψ^+ is finite.) It is then readily verified that

$$f(\rho^{(n)}) \leqslant -(1/2)\, m_0^2 l_n^{-1} + \psi^+ + \beta^{-1} \ln(\|\rho^{(n)}\|_\infty) \tag{3.31}$$

By means of (3.30) the right-hand side of (3.31) tends to minus infinity as $n \to \infty$, which proves the claim. ▌

Were the sequences which converge (weakly) to the single-point Dirac measure the only ones which lead to a diverging of $f(\rho)$ to minus infinity, we would have almost completed the proof of our assertion. There exist, however, other sequences which are not equivalent to the above one but which still result in a diverging of the free energy to minus infinity. For example, one might consider a sequence of probability densities which approaches the Dirac measure of a short, straight line in Λ. Then, for small enough temperature, and only then, the free energy becomes minus infinity as well. Thus, as next step we have to show that in a certain well-defined and reasonable sense the Dirac single-point measure gives a "more negative infinite value" of the free energy than does any other probability measure. We do this by showing that the amount of negative gravitational energy lost by the system upon separating two pieces of matter located initially at the same point in space overcompensates the increase in β^{-1} times the entropy.

Proof. We compare the behavior of the free energy along any of the above-introduced sequences of C^{∞} probability densities which converge to a single-point Dirac measure with that along a sequence of weighted linear superpositions of the *same* densities, now defined, however, on two different supports, and which converge, hence, to a superposition of two such point measures concentrated on different points. Let the sequences of normalized positive C_0^{∞} functions $\{\rho_{\mathbf{r}_i}^{(n)} \mid n = 1, 2, \ldots\}$ be defined as above, with, however, $\mathbf{r}_0$ replaced by $\mathbf{r}_i \in \Lambda$; $i = 1, 2, 3$. Choose three locations $\mathbf{r}_1, \mathbf{r}_2 \neq \mathbf{r}_3$ in Λ. We allow that $\mathbf{r}_1$ coincides with one of the other two locations. We consider then

$$\Delta_{\alpha}^{(n)} \equiv f(\rho_{\mathbf{r}_1}^{(n)}) - f([1-\alpha]\,\rho_{\mathbf{r}_2}^{(n)} + \alpha\rho_{\mathbf{r}_3}^{(n)}) \tag{3.32}$$

with $\alpha \in \,]0, 1[$ real, as $n \to \infty$. By construction, $\operatorname{supp} \rho_{\mathbf{r}_i}^{(n)} \subset\subset \Lambda$ for all i. Then the following integrals contained in $\Delta^{(n)}$ are independent of the subscript i:

$$w^{(n)} \equiv \int_{\Lambda}\int_{\Lambda} \rho_{\mathbf{r}_i}^{(n)}(\mathbf{r})\,\rho_{\mathbf{r}_i}^{(n)}(\mathbf{r}')\,U_{\mathrm{cl}}(|\mathbf{r}-\mathbf{r}'|, \varepsilon)\,d^3r\,d^3r' \tag{3.33}$$

and

$$s^{(n)} \equiv -\int_{\Lambda} \rho_{\mathbf{r}_i}^{(n)}(\mathbf{r}) \ln \rho_{\mathbf{r}_i}^{(n)}(\mathbf{r})\,d^3r \tag{3.34}$$

These integrals diverge to minus infinity as $n \to \infty$. In addition, the following integral is finite as $n \to \infty$:

$$\hat{w}^{(n)} \equiv \int_\Lambda \int_\Lambda \rho_{\mathbf{r}_2}^{(n)}(\mathbf{r})\, \rho_{\mathbf{r}_3}^{(n)}(\mathbf{r}')\, U_{\text{cl}}(|\mathbf{r}-\mathbf{r}'|, \varepsilon)\, d^3r\, d^3r' \tag{3.35}$$

which is due to supp $\rho_{\mathbf{r}_2}^{(n)} \cap$ supp $\rho_{\mathbf{r}_3}^{(n)} = \varnothing$ for large enough n, and also the integrals

$$\varphi_i^{(n)} \equiv \int_\Lambda \rho_{\mathbf{r}_i}^{(n)}(\mathbf{r})\, \psi(\mathbf{r})\, d^3r \tag{3.36}$$

are finite. Noting further that due to the compactness of $\bar{\Lambda}$ the entropy (per particle) functional $-\int \rho \ln \rho\, d^3r$ is bounded from above by $\mathscr{V}/e$, we obtain the estimate

$$\Delta_\alpha^{(n)} \leqslant \alpha(1-\alpha)\, w^{(n)} - \beta^{-1} s^{(n)} + C \tag{3.37}$$

where C is finite and independent of n. Then, by an estimate that is essentially equivalent to (3.31), and noting again (3.30), we obtain $\Delta_\alpha^{(n)} \searrow -\infty$ as $n \to \infty$ for $\alpha \neq 0, 1$. In this sense, the single-point Dirac measure gives a "more negative infinite" free energy than does the weighted double-point Dirac measure. ∎

The above proof means that separating self-gravitating matter takes free energy. Taking into account now the geometrical argument that any probability measure not concentrated on a single point means matter being separated, the above proof suffices to show that $f(\rho)$ approaches its infimum for $\rho(\mathbf{r}) \rightharpoonup \delta(\mathbf{r}-\mathbf{r}_c)$ if $\varepsilon = 0$.

The above proofs give no information of where to choose $\mathbf{r}_c$, i.e., we have not gained any information about the decomposition measure $\nu(d\varrho \mid \mu)$ [see (3.25)] from $f(\rho)$. It is possible to establish by means of a kind of renormalization technique that Υ carries the minimizing support in the sense that $\nu(d\varrho \mid \mu)$ is concentrated on the Dirac point measures with support in Υ. In this sense, by essentially the same technique as used above, we now show that $f(\rho)$ approaches a "more negative infinite value in the renormalized sense" if the Dirac measure is concentrated on a point in Υ than it does if the Dirac measure is concentrated on any other point $\in \Lambda \setminus \Upsilon$.

Proof. Let $\alpha = 0$ (equivalently, $\alpha = 1$) in (3.32) and let $\mathbf{r}_1 \in \Upsilon$, $\mathbf{r}_2 \notin \Upsilon$. In this case

$$\Delta_0^{(n)} = \varphi_1^{(n)} - \varphi_2^{(n)} \tag{3.38}$$

by means of (3.33) and (3.34). Obviously, $\Delta_0^{(n)}$ is finite for all n, and by definition of Υ we obtain

$$\lim_{n \to \infty} \Delta_0^{(n)} < 0 \tag{3.39}$$

which was to be shown. ▌

So far, we have shown that for $\varepsilon = 0$ the free-energy functional $f(\rho)$ approaches its infimum value $-\infty$ for $\rho(\mathbf{r}) \rightharpoonup \delta(\mathbf{r} - \mathbf{r}_c)$ and, upon using a kind of simple renormalization argument, we have shown that sense can be given to the statement that $\nu(d\varrho \mid \mu)$ singles out those Dirac measures with $\mathbf{r}_c \in \Upsilon$. The remaining step is to show that $f(\rho)$ takes that infimum with $\delta(\mathbf{r} - \mathbf{r}_c)$ as limit (as $\varepsilon \searrow 0$) of the global minimizers for $\varepsilon \neq 0$. We do this by proving the following: For $\varepsilon = 0$, the Dirac delta distribution $\delta(\mathbf{r} - \mathbf{r}_c)$ solves (3.27) in the sense that it is a weak* limit point of (3.27) for any $\mathbf{r}_c \in \Lambda$.

Proof. Let the sequence of normalized, positive C_0^∞ functions $\{\rho_{\mathbf{r}_c}^{(n)} \mid n = 1, 2, \ldots\}$ again be defined as in the proofs above. Set $\varepsilon = 0$ in (3.27) and insert $\rho_{\mathbf{r}_c}^{(n)}$ for ρ. Consider the w^*-limit of both the left- and the right-hand side of (3.27). The left-hand side converges (weakly) to $\delta(\mathbf{r} - \mathbf{r}_c)$, by construction. To facilitate the discussion of the right-hand side, we introduce the notation $(\rho * U)(\mathbf{r})$, meaning the convolution product of ρ with U (see, for instance, ref. 24), such that (3.27) then reads

$$\rho = \frac{\exp\{-\beta[(\rho * U) + \psi]\}}{\int_\Lambda \exp\{-\beta[(\rho * U) + \psi]\}\, d^3r} \tag{3.40}$$

for all ε. For $\rho = \rho^{(n)}$ and $\varepsilon = 0$ we get

$$(\rho_{\mathbf{r}_c}^{(n)} * U_{\text{cl}}) \to -m_0^2\, |\mathbf{r} - \mathbf{r}_c|^{-1} \tag{3.41}$$

strongly as $n \to \infty$. Thus, $\int \exp\{-\beta[(\rho^{(n)} * U_{\text{cl}}) + \psi]\}\, d^3r$ diverges to $+\infty$ with $n \to \infty$. On the other hand, since $U_{\text{cl}}(\xi)$ is C^∞ in the complement of the origin and ψ Lipschitz continuous, the right-hand side of (3.27) converges to zero pointwise with n for $\varepsilon = 0$ and $\mathbf{r} \neq \mathbf{r}_c$. The fact that normalization of the right-hand side of (3.27) or (3.40) is independent of n then guarantees the validity of the claim. ▌

One might wonder that $\delta(\mathbf{r} - \mathbf{r}_c)$ solves (3.27) for $\varepsilon = 0$ independently of whether $\mathbf{r}_c$ is in Υ or not. However, we must not overlook that in the proof that $f(\rho)$ approaches its infimum for $\mathbf{r}_c \in \Upsilon$ we employed a kind of renormalization technique that allowed us to speak of "more negative infinite values of f in the renormalized sense." One should remember that this notion was introduced to obtain information on $\nu(d\varrho \mid \mu)$. The counterpart to this renormalization argument might be considered to be the following fact: For given sequence $\{\rho^{(n)}\}$, the convergence of the right-hand side of (3.27) to the Dirac distribution is fastest for $\mathbf{r}_c \in \Upsilon$, which can readily be proved.

4. SYSTEMS WITH BOUNDED, UNSTABLE PAIR INTERACTIONS

In this section we would like to see how the results derived in the preceding two sections bear on the conclusions that can be drawn for the thermodynamic behavior of classical systems with only slightly smoothed-out gravitational interactions, i.e., $\varepsilon \neq 0$ but still very small. Our final goal in this section will be to show the existence of a gravitational-type phase transition (for the spherical systems) at temperatures far above the critical value beyond which the stable isothermal Emden gas spheres cease to exist.

For the systems that we are interested in here the exact evaluation of the configurational integral Q and, hence, of the configurational free energy F^{conf} in terms of closed expressions in β, N, and $\mathscr{V}$, as is familiar from the statistical mechanics of ideal homogeneous systems (e.g., ref. 16), seems to be completely unfeasible, at least by present calculational techniques. In this situation the thermodynamic limit in the mean-field scaling (3.15) offers the possibility to investigate the canonical ensemble in a rather manageable way.

An exact evaluation of this limit requires the solution of the nonlinear integral equation (3.27). Hence, some knowledge of the general solution properties of (3.27) is of value. [That (3.27) always has a solution follows from the fact that $f(\rho)$ takes its global minimum,[15] which exists and is finite for $\varepsilon \neq 0$.] Some attention was paid to this problem in ref. 15, where two uniqueness results were proved (depending on the assumptions made for the potential energy U) and an explicitly soluble one-dimensional model (cosine interaction) was discussed. This section here will center around a discussion of (3.27) with $U(|\mathbf{r}-\mathbf{r}'|, \varepsilon)$ being a gravitational-type interaction energy for three-dimensional systems. We point out that we do not specialize to spherical symmetry until Eq. (4.16).

To facilitate the discussion, we assume that the external gravitational field ϕ is identically zero, and we shall again set the wall energy W equal to zero in the calculations below. Hence, in the following, whenever we refer to (3.40) we understand that equation with $\psi \equiv 0$ without mentioning this explicitly, unless otherwise stated.

We start with the discussion of the two extreme cases, i.e., the zero- and the infinite-temperature limits, meaning $\beta \to \infty$ and $\beta \to 0$, respectively.

Both limits can be evaluated exactly, and in both cases there exists a unique (up to, possibly, translation) globally minimizing equilibrium density ρ which solves (3.27) [equivalently, (3.40)]. We denote these limiting solutions of [(3.27), (3.40)] as $\lim_{\beta \to 0} \rho = \rho_0$ and $\lim_{\beta \to \infty} \rho = \rho_\infty$. Then

$$\rho_0 = \mathscr{V}^{-1} \tag{4.1}$$

and

$$\rho_\infty = \delta(\mathbf{r} - \mathbf{r}_c) \tag{4.2}$$

with $\mathbf{r}_c \in \Lambda$ arbitrary.

Proof. By construction, $-U(\xi, \varepsilon)$ is absolutely bounded and strictly positive for $\xi \leqslant \Lambda$. Thus, since ρ is a probability density, necessarily $-\beta U * \rho \in L_+^\infty(\Lambda)$, meaning the equivalence classes of positive bounded functions. Obviously, then,

$$\lim_{\beta \to 0} \exp(-\beta\rho * U) \Big/ \int_\Lambda \exp(-\beta\rho * U)\, d^3r = \mathscr{V}^{-1} \tag{4.3}$$

independently of the choice of ρ, which proves (4.1).

On the other hand, noting that

$$\int_\Lambda \exp(-\beta\rho * U)\, d^3r = \|\exp(-\rho * U)\|_\beta^\beta \tag{4.4}$$

for $\beta \geqslant 1$, and postulating that $\rho * U$ has a unique global minimum at $\tilde{\mathbf{r}}$, we can repeat the analogous steps as in the proof of (3.5c) [Eqs. (3.6)–(3.12)] and find for any such given ρ

$$w^*\text{-}\lim_{\beta \to \infty} \exp(-\beta\rho * U) \Big/ \int_\Lambda \exp(-\beta\rho * U)\, d^3r = \delta(\mathbf{r} - \tilde{\mathbf{r}}) \tag{4.5}$$

Now, $\delta_{\tilde{\mathbf{r}}} * U = U(|\mathbf{r} - \tilde{\mathbf{r}}|, \varepsilon)$ has a unique global minimum at $\tilde{\mathbf{r}}$, which follows from the proposed properties of U. Hence, we have $\tilde{\mathbf{r}} = \mathbf{r}_c$, such that in the limit $\beta \to \infty$ the Dirac distribution $\delta(\mathbf{r} - \mathbf{r}_c)$ solves (3.40) for any $\mathbf{r}_c \in \Lambda$. This is also the only minimizing solution, up to translation, in the limit $\beta \to \infty$, which follows from the fact that for $T = 0$ the free energy per particle is just the gravitational (potential) energy per particle, here for regularized interactions. It is readily shown that separating matter takes gravitational energy [the analogue to the proof that separating matter takes free energy, Eqs. (3.32)–(3.37)]. That way one finds that there is no other minimal point of $f(\rho)$ for $T = 0$. Hence the minimizing solution is unique up to translation inside Λ. This proves (4.2). ∎

Equations (4.1) and (4.2) are in accord with intuitive expectations. In the limit of infinite temperature the particle interactions become unimportant and matter is distributed homogeneously over the domain Λ. At zero temperature matter is distributed such that the total potential energy due to U is minimal, which, in our case, implies that matter is concentrated on a single point in Λ. The location of that point is completely indeterminate, which means that the equilibrium measure is unique only up to translation inside Λ. The physical reason for this is that a completely collapsed system feels no influence of the confining walls, except when it hits them.

There exists also a finite right neighborhood $\Xi =]0, \check{\beta}[$ of $\beta = 0$, with $\check{\beta} = 1/(2\,\|U\|_{\infty})$, such that for $\beta \in \Xi$ the solution of (3.40) is unique (ref. 15, Theorem 3). It is not known yet whether there exists an analogous and complementary regime $\beta \nearrow \infty$ with a unique solution of (3.40), or whether such a regime can exist at all.

To discuss the solution properties of (3.40) further, the following is of value: For $0 \leqslant \beta < \infty$ there is a one-to-one correspondence between the solutions ρ of the nonlinear integral equation (3.40) and the (positive) solutions Ψ of

$$\Psi = \eta(-U) * (e^{\Psi}) \tag{4.6}$$

where $\eta \in \overline{\mathbb{R}^{+}}$.

Proof. For every solution pair (β, ρ) of (3.40) we define a unique generalized Newtonian potential Ψ via $\Psi \equiv -\beta U * \rho$, with $\Psi \in L_{+}^{\infty}(\Lambda)$ (see above). That potential Ψ must not be confused with ψ. We derive an equation for Ψ starting from (3.40).

For finite β the integral $\int_{\Lambda} \exp(-\beta U * \rho)\, d^3 r$ exists, and therefore to every solution pair (β, ρ) of (3.40) we can define a unique parameter $\eta(\beta; \rho) \equiv \beta / \int_{\Lambda} \exp(-\beta U * \rho)\, d^3 r$, with $\eta < \infty$. Upon taking now the convolution product of $-\beta U$ with (3.40), and noting the definitions of Ψ and η, we obtain (4.6). Hence, to every solution pair (β, ρ) of (3.40) there corresponds uniquely a bounded positive solution pair (η, Ψ) of (4.6).

To prove the converse, we start with (4.6). Because of the boundedness properties of U, for finite η every solution of (4.6) is necessarily strictly positive and bounded. This follows from (4.6) together with the inequalities

$$-U * e^{\Psi} \geqslant |U(\Lambda, \varepsilon)|\, \|e^{\Psi}\|_1 > 0 \tag{4.7a}$$

and

$$-U * e^{\Psi} \leqslant |U(0, \varepsilon)|\; \|e^{\Psi}\|_1 \tag{4.7b}$$

which hold pointwise in Λ. In fact, the assumption that $\|e^{\Psi}\|_1$ would not exist for a solution Ψ of (4.6) would mean, by (4.6) and (4.7a), that Ψ is pointwise infinite and thus no solution, which means a contradiction. Hence the integral $\int_\Lambda e^{\Psi}\, d^3r$ exists for every solution of (4.6) pertaining to finite η. By (4.7b) and (4.6), then, $\|\Psi\|_\infty < \infty$. As a consequence, for every solution Ψ of (4.6) we can define a unique probability density ρ via

$$\rho \equiv e^{\Psi} \Big/ \int_\Lambda e^{\Psi}\, d^3r \tag{4.8}$$

We would like (4.8) to solve (3.40) for a unique choice of β. Upon comparing (4.8) with (3.40), we find that the density ρ defined by (4.8) then must also solve the integral equation $\Psi = -\beta U * \rho$, referred to as (I), given the same Ψ as in (4.8). Up to now we know nothing about β except $\beta > 0$. Inserting now (4.8) for ρ into the right-hand side of (I), and noting that Ψ solves (4.6), we see that ρ as given by (4.8) solves (I) iff we identify β uniquely with $\beta(\eta; \Psi) \equiv \eta \int_\Lambda e^{\Psi}\, d^3r$. Thereby we have shown that to every solution pair (η, Ψ) of (4.6), with η finite, there corresponds uniquely a solution pair (β, ρ) of (3.40). ▮

Since we know that (3.40) has at least one bounded solution for every $\beta < \infty$, in particular (4.1) for $\beta = 0$, it follows that (4.6) must have a bounded solution at least for all η in a small right neighborhood of $\eta = 0$. We want to conclude, however, from the solution properties of (4.6) some of the solution properties of (3.40). In the following we list some important (for our aims) solution properties (SP) of (4.6). The linear integral operator $(-\beta U * \cdot)$ will be abbreviated by $K(\cdot)$.

SP1. There exists a positive constant $\eta^* < \infty$ such that (4.6) has no positive solution for $\eta > \eta^*$, with the estimate $\eta^* \leqslant 1/e\kappa$. Here, κ is the finite, positive spectral radius of K, and e is Euler's number.

SP2. For $\eta < \eta^*$ the problem (4.6) has at least one solution. In particular, for every $\eta < \eta^*$ there exists a minimal solution $\tilde{\Psi}_\eta$. The notion of "minimal solution" means that $\tilde{\Psi}_\eta$ is pointwise smaller than any other solution of (4.6) that belongs to the same η. The minimal solution can be calculated by means of the iteration scheme

$$\Psi^{(n+1)} = \eta K(e^{\Psi^{(n)}}) \tag{4.9a}$$

with

$$\Psi^{(0)} \equiv 0 \tag{4.9b}$$

being a strict subsolution of (4.6). [Recall that $\check{\Psi}$ is a subsolution if $\check{\Psi} \leqslant \eta K(e^{\check{\Psi}})$, and a strict subsolution if strict inequality holds. Supersolutions $\hat{\Psi}$ are defined analogously by reversing the ordering.] The iterational sequence (4.9) converges monotonically increasing to $\Psi^{(\infty)} = \tilde{\Psi}_\eta$. Iff $\lim_{\eta \nearrow \eta^*} \|\tilde{\Psi}_\eta\|_\infty < \infty$ then there exists also a solution $\tilde{\Psi}^*$ for $\eta = \eta^*$.

SP3. For $0 \leqslant \eta < \eta^*$ the mapping $\eta \mapsto \tilde{\Psi}_\eta$ is pointwise strictly monotonically increasing and at least twice continuously right-differentiable with respect to η.

SP4. The minimal solutions $\tilde{\Psi}_\eta$ for $\eta < \eta^*$ are locally stable in the sense of Amann, meaning that the eigenvalues of the linearized equilibrium operator $\mathscr{T}_\eta(\cdot) \equiv \mathrm{id}(\cdot) - \eta K(e^{\tilde{\Psi}_\eta} \cdot)$ are positive for $\eta < \eta^*$.

SP5. The solution $\tilde{\Psi}^*$, if it exists, is marginally stable [the linearized operator $\mathscr{T}_{\eta^*}(\cdot)$ has the eigenvalue zero] and there exists a small left neighborhood of η^* such that (4.6) has at least two distinct solutions in the neighborhood of the pair $(\eta^*, \tilde{\Psi}^*)$. In the corresponding bifurcation diagram the solution branch containing $(\eta^*, \tilde{\Psi}^*)$ bends back at $(\eta^*, \tilde{\Psi}^*)$ with respect to η.

Proof. SP1: K has a strictly positive integral kernel $-U(|\mathbf{r} - \mathbf{r}'|, \varepsilon)$ defined on $\Lambda \times \Lambda$. Then the Krein–Rutman theorem[27] guarantees that κ is the largest eigenvalue of K, and further that the corresponding eigenspace has dimension one with a nonzero eigenfunction u_κ (a simple proof of the second statement pertaining to the situation given here is given in ref. 28, Theorem 3.3.2). This then allows us either to apply the same technique as introduced in ref. 29 to prove the existence of η^* (ref. 29, Theorem 1) for $-\Delta\Psi = \eta e^\Psi$, together with the above given bound, or equivalently, ref. 30, Proposition 20.2, applies, giving the same result. SP2: The proof is given in ref. 31, Proposition 3.1. SP3: The proof is given in ref. 30, Theorem 26.3. SP4 and SP5: The proof is again due to ref. 31, Proposition 3.1. ∎

Equation (4.6) belongs to the large class of extensively studied nonlinear fixed-point problems in ordered Banach spaces. For a detailed survey over more general results see, for instance, ref. 30.

So far we have no lower bound for η^*. The following helps in inferring such a bound if the solution properties of a suitable comparison problem are known.

SP6. Let $\hat{K}(\cdot)$ be a compact, symmetrical, linear integral operator with strictly positive integral kernel $\hat{K}(\cdot, \cdot)\colon \Lambda \times \Lambda \to \mathbb{R}^+$. Clearly, the analogues of SP1–SP5 apply also to the equation

$$\Phi = \zeta \hat{K}(e^\Phi) \equiv \zeta \int_\Lambda \hat{K}(\mathbf{r}, \mathbf{r}') \exp[\Phi(\mathbf{r}')]\, d^3r' \tag{4.10}$$

Let $\hat{K} > -U$ pointwise on $\Lambda \times \Lambda$. Then to every solution Φ_ζ of (4.10) there exists a solution Ψ_η of (4.6) with $\zeta = \eta$.

Proof. Since $\hat{K} > -U$ by assumption, for Φ_ζ a solution of (4.10) we have

$$\Phi_\zeta = \zeta \hat{K}(e^{\Phi_\zeta}) > \zeta K(e^{\Phi_\zeta}) \tag{4.11}$$

Hence, Φ_ζ is a strict supersolution of (4.6) for $\eta \equiv \zeta$. Taking Φ_ζ as starting solution for the iterational scheme (4.9a) gives a monotonically decreasing sequence, which necessarily must converge since (4.9b) is a strict subsolution which is smaller than Φ_ζ. ∎

Clearly, by the above-proved equivalence of (3.40) and (4.6), to every minimal solution $\tilde{\Psi}_\eta$ of (4.6) there belongs a unique solution $\tilde{\rho}_{\beta(\eta)}$ of (3.40), given by

$$\tilde{\rho}_{\beta(\eta)} = e^{\tilde{\Psi}_\eta} \Big/ \int_A e^{\tilde{\Psi}_\eta} \, d^3r \tag{4.12a}$$

with the corresponding inverse temperature $\beta(\eta)$ given by

$$\beta(\eta) = \eta \int_A e^{\tilde{\Psi}_\eta} \, d^3r \tag{4.12b}$$

The solutions $\tilde{\rho}$ belong to a class of solutions of (3.40) which we shall call (suggestively) "shallow solutions." The reason for that will become clear below.

Some of the more important properties (P) of the shallow solutions as defined by (4.12) are as follows.

P1. For $0 \leqslant \eta < \eta^*$ the mapping $\eta \mapsto \tilde{\rho}_{\beta(\eta)}$, given by (4.12a), is at least twice continuously right-differentiable w.r.t. η.

P2. For $0 \leqslant \eta < \eta^*$ the mapping $\eta \mapsto \beta(\eta)$, given by (4.12b), is at least twice continuously right-differentiable with respect to η and strictly monotonically increasing. We have $\beta(0) = 0$.

P3. In the parameter regime of β values, starting with 0, where the shallow solutions as defined by (4.12) exist, the mapping $\beta \mapsto \tilde{\rho}_\beta$ is at least twice continuously right-differentiable w.r.t. β.

P4. There exists $\breve{\beta} \in \Xi \cup \{\breve{\beta}\}$ such that for $\beta < \breve{\beta}$ the solutions ρ of (3.40) in the uniqueness regime $\beta \in \Xi$ are given by (4.12). Equivalently, there exists $\breve{\eta} \leqslant \eta^*$ such that for $\eta < \breve{\eta}$ the minimal solutions $\tilde{\Psi}_\eta$ correspond uniquely to solutions $\tilde{\rho}_{\beta(\eta)}$ of (3.40) with $\beta(\eta) \in \Xi$. In particular, $\breve{\beta} = \beta(\breve{\eta})$.

Proof. Property P1 follows from (4.12a) and SP3. Property P2 follows from (4.12b) and SP3. Property P3 is a direct consequence of P1 and P2. Finally, P4 follows from P2, P3, SP2, and SP3. ∎

The following property of the shallow solutions as defined by (4.12) makes them of interest for the statistical mechanics equilibrium problem: The solutions given by (4.12) are locally stable, i.e., they are local minimizers of the free-energy functional $f(\rho)$.

Proof. Let ρ be a solution of (3.40) and denote by σ a small (bounded) perturbation of the equilibrium density ρ. Clearly, in order that $\rho + \sigma$ is a probability density, σ has to fulfill $\int_\Lambda \sigma \, d^3r = 0$. Nevertheless, we shall not need this property of σ to prove our claim. We take $f(\rho)$ and consider the second variation $(\delta^2 f)_\rho (\sigma)$. It reads

$$(\delta^2 f)_\rho (\sigma) = (1/2)\langle \sigma, U * \sigma \rangle + (1/2)\, \beta^{-1} \langle \rho^{-1}, \sigma^2 \rangle \tag{4.13}$$

In (4.13), the angular brackets denote the canonical L^2 scalar product. It is readily shown that $(\delta^2 f)_\rho (\sigma)$ is absolutely bounded as $|(\delta^2 f)_\rho (\sigma)| < C_{U,\beta} \|\sigma\|_2^2$, where $C_{U,\beta}$ is a constant depending on U and β. Hence, for σ suitably normalized, $\inf \delta^2 f$ exists, and to prove the claim it suffices to show that this infimum is positive for $\rho = \tilde{\rho}$, as given by (4.12). We choose $(1/2) \|\sigma \rho^{-1/2}\|_2^2 = 1$ as normalization. Then the minimizing σ of (4.13) solves the Euler–Lagrange equation

$$U * \sigma + \beta^{-1} \rho^{-1} \sigma = \bar{\omega} \rho^{-1} \sigma \tag{4.14}$$

where $\bar{\omega}$ is the smallest eigenvalue of the operator defined by the left-hand side of (4.14). Now it follows from (4.13) or (4.14) that for $\sigma \in \operatorname{Ker}(U * \cdot)$ the second variation $(\delta^2 f)_\rho (\sigma)$ of (3.26) is strictly positive for any given ρ. Thus, we can restrict the further considerations to the case $\sigma \in (\operatorname{Ker}(U * \cdot))^\perp$, where the orthogonality is meant with respect to L^2 scalar product. We now multiply (4.14) by $\beta\rho$ and take the convolution product of (4.14) with U. This gives

$$U * \sigma + \beta U * (\rho (U * \sigma)) = \bar{\omega} \beta U * \sigma \tag{4.15}$$

Now noting the definition of $\mathscr{T}_\eta(\cdot)$ given in SP4 with ρ a shallow solution given by (4.12), we see that (4.15) is just the eigenvalue problem for $\mathscr{T}_\eta(U * \sigma)$. The eigenvalues of $\mathscr{T}_\eta(\cdot)$ are positive by SP4, for ρ a shallow solution given by (4.12). This proves the claim. ∎

By P4, for $\eta < \bar{\eta}$ the iterational sequence (4.9) converges to the thermodynamic equilibrium state, which belongs to the uniqueness regime $\beta \in \Xi$. Clearly, if $\bar{\eta} = \eta^*$, then the iteration method (4.9) does not generate solutions that cannot be computed (equally well) by iterating directly (3.40). If, however, $\bar{\eta} < \eta^*$ or even $\bar{\eta} \ll \eta^*$, then (4.9) generates locally stable solutions of (3.40) that cannot necessarily be calculated by means of an iteration scheme that uses directly the operator defined by the right-hand side of (3.40). If the latter is the case, then possibly (4.9) generates thermodynamically metastable states which are not identical with the thermodynamic equilibrium state, in the sense that $\tilde{\Psi}_\eta$ does not correspond to the global minimum of either $f(\rho)$ or $\tilde{f}(\rho)$ for $\eta > \bar{\eta}$.

So far we have not made any symmetry assumptions. For the further discussion we turn now to the more concrete situation of systems confined to a hollow sphere of radius 1 and center at the origin. For the above investigations it was also not necessary to specify ε except that $\varepsilon \neq 0$. We wish now to inquire into the thermodynamic behavior of systems which are nearly purely gravitating over the typical distances provided by the confining sphere. This means that we study here explicitly systems for which the interparticle distances where the modifications to the Newtonian potential become important are several orders of magnitude smaller than 1. This will sharply restrict the possible values of ε (see below). To simplify the discussion, we assume that $U(\xi, \varepsilon) > U_{\mathrm{cl}}(\xi)$ for all nonnegative ξ. [The explicit examples in (3.14) fulfill this requirement.] In addition, we set $m_0 = 1$ [see (2.4) and (3.15)].

For spherical domains it has been shown[10] that in the exactly self-gravitating case, i.e., $U = U_{\mathrm{cl}}$, all *local* minimizers of the relevant free-energy functional are necessarily spherically symmetric, and that then the density decreases monotonically outward from the center (see also ref. 8 and ref. 21, Appendix A). Since the only feature of U_{cl} that is used in that proof is the fact that U_{cl} is purely attractive, which holds also for U with $\varepsilon \neq 0$, we can conclude that for $\varepsilon \neq 0$ the local minimizers of (3.26) in a hollow sphere are likewise spherically symmetric and radially decreasing outward. Therefore, in what follows, we consider only densities ρ which possess these properties.

For the spherical systems we can state an important existence theorem for solutions of (4.6): To every solution Φ_ζ of

$$\Phi = \zeta r^{-1} * e^{\Phi} \tag{4.16}$$

defined in the unit sphere B_1 centered around the origin, with $r = |\mathbf{r}|$, there belongs a minimal solution $\tilde{\Psi}_\eta$ of (4.6), with $\eta = \zeta$, which is also pointwise smaller than any solution Φ_ζ for the given ζ.

Proof. We have postulated above that $U(\xi, \varepsilon) > U_{\mathrm{cl}}(\xi)$ for all $\xi \in \overline{\mathbb{R}^+}$. The claim follows from SP6. ∎

Equation (4.16) is (the integral version of) the well-known Emden equation of the isothermal gas spheres,[11,12] sometimes also called the isothermal Lane–Emden equation.[15]

Remark 1. In the limit $\varepsilon = 0$, (4.6) goes over into (4.16). Thus, all the bounded spherical solutions of (3.40) are then necessarily generated by solutions of (4.16), i.e., isothermal gas spheres, with (4.18) (see below) as limiting case. Nevertheless, in this limit, (3.40) and (4.6) (treating η as given constant) are not equivalent, because (3.40) has the Dirac measure as additional weak solution for every β.

Remark 2. For the systems with spherical symmetry, (4.16) is equivalent to the elliptic boundary value problem of second order

$$-\Delta u = \lambda e^{u} \tag{4.17a}$$

in B_1, with

$$u = 0 \tag{4.17b}$$

on ∂B_1, and $\lambda \in \overline{\mathbb{R}^+}$, which is known in the mathematical literature as Gelfand's problem.[30,32,33] For instance, from a solution pair (ζ, Φ_ζ) of (4.16) one obtains the corresponding solution pair $(\lambda(\zeta), u_{\lambda(\zeta)})$ of (4.17) via $u_{\lambda(\zeta)}(r) = \Phi_\zeta(r) - \Phi_\zeta(1)$ and $\lambda(\zeta) = 4\pi\zeta \exp[\Phi_\zeta(1)]$.

Clearly, by the analogue of SP1 there exists a positive constant ζ^* such that (4.16) has no solution for $\zeta > \zeta^*$. On the other hand, for $\zeta = 1/(2\pi e^2) \equiv \zeta_Z$, (4.16) admits the elementary solution

$$\Phi_Z = 2(1 - \ln r) \tag{4.18}$$

known as Zöllner's solution.[11,12] This implies, by noting SP2 and SP6, that in the spherically symmetric situation discussed here the existence of minimal solutions $\tilde{\Psi}_\eta$ of (4.6) is guaranteed for η values at least up to ζ_Z. That means we have ζ_Z as lower bound for η^*,

$$\eta^* \geqslant 1/(2\pi e^2) \tag{4.19}$$

[It should be clear that the minimal solution of (4.6) that belongs to $\eta = \zeta_Z$, which exists by SP6, need have nothing to do with (4.18).] We are now able to give an explicit, rather weak, lower bound β_* for the existence of shallow solutions of (3.40) in the unit sphere:

$$\beta_* = 2/(3e^2) \tag{4.20}$$

There exist locally stable shallow solutions of (3.40) with spherical symmetry at least up to values of β somewhat larger than β_* given by (4.20).

Proof. As an immediate consequence of P2 and (4.19) we have $\beta(\eta_Z)$, defined via (4.12b), as lower bound for the existence of shallow solutions of (3.40) in the unit sphere. Making use of the fact that the solutions of (4.6) are positive, we replace $\tilde{\Psi}_\eta$ in (4.12b) by 0 and obtain $\beta(\eta_Z) > \beta_*$ strictly. ∎

The interesting point now is that (4.20) is roughly 0.1 and independent of ε, whereas the bound $\breve{\beta} = 1/(2\,|U(0, \varepsilon)|)$ that guarantees[15] the uniqueness of ρ for $\beta < \breve{\beta}$ (see above) is proportional to ε. In fact, $\breve{\beta} = \varepsilon/[2\,|U(0, 1)|]$ by (2.6). [As an example, consider U to be given by (3.14) with $m_0 = 1$—see Eq. (3.15b). Then $|U(0, \varepsilon)| = 1/\varepsilon$ for all n. Thus, $\breve{\beta} = \varepsilon/2$, then.] Now ε is here a direct measure of the distance over which the Newtonian interactions are modified, which we have postulated to be very small as compared to the attainable particle separations, which are of order 1 here. Consequently, we have $\varepsilon \ll 1$ by several orders of magnitude, and $\breve{\beta}$ and β_* become well separated. More precisely, given $U(\xi, \varepsilon)$, we have to choose $\varepsilon \ll 1$ so small that $\varepsilon/[2\,|U(0, 1)|] \ll \beta_*$. Thus, there exist then shallow solutions $\tilde{\rho}_\beta$, given by (4.12), with $\beta \notin \Xi$, meaning $\bar{\beta} = \breve{\beta}$ (see P4). This opens the possibility that for some interval of β values, with $\beta > \breve{\beta}$, there are locally stable shallow solutions which nevertheless do not describe the thermodynamic equilibrium state. Our next argument shows that this is the case.

We know from the preceding two sections that as $\varepsilon \searrow 0$ the statistical mechanics equilibrium state converges to the material point. In particular, we know that for any given finite β the free-energy functional $f(\rho)$ as given by (3.26) can be made smaller than any prescribed negative value by decreasing ε to smaller but still finite values, which is a direct consequence of the analogue of inequality (2.33) for (3.26). Consequently, for ε finite but small, for small enough temperatures there will exist highly peaked but nevertheless positive (everywhere in Λ) equilibrium densities as thermodynamic equilibrium states, which may be termed to be of "core-atmosphere type." We shall use this notion rather suggestively without precise definition; however, one might think of a smeared-out approximation ($\in L^1_+$) of the delta distribution. Note that at zero temperature the equilibrium state is again the material point. We estimate how small such a "small enough" temperature has to be.

We approximate a distribution with core-atmosphere structure by the box distribution [(2.31), (2.32), (2.34)]. Here, an obviously reasonable choice for s_0 [see (2.34)] is $s_0 = 1$. Since $\varepsilon \ll 1$, the box distribution is highly peaked and concentrated on a very small sphere inside Λ. Then, in analogy to (2.33), we have

$$\inf_\rho f(\rho) \leqslant \varepsilon^{-1}(1/2)\,U(2, 1) - 3\beta^{-1} \ln \varepsilon - \beta^{-1} \ln(4\pi/3) \qquad (4.21)$$

On the other hand, for $0<\eta<\zeta^*$ the free energy of the shallow solutions pertaining to $\beta(\eta)$ is bounded below independently of ε. We find

$$f(\tilde{\rho}_{\beta(\eta)})=\tilde{f}(\tilde{\rho}_{\beta(\eta)})>-\beta^{-1}\ln\int_\Lambda e^{\Psi_\eta}\,d^3r$$

$$>-\beta^{-1}\ln\int_\Lambda e^{\tilde{\Phi}_\zeta}\,d^3r>-\beta^{-1}\ln\int_\Lambda e^{\tilde{\Phi}^*}\,d^3r \quad (4.22)$$

with $\tilde{\Phi}_\zeta$ the minimal solution of (4.16) for $\zeta=\eta$, and $\tilde{\Phi}^*$ the minimal solution of (4.16) for $\zeta=\zeta^*$.

Proof. First of all, $\tilde{\Phi}_\zeta$ and $\tilde{\Phi}^*$ exist[11,12,30,33] and are bounded by the analogues of SP2 and SP3. The equality in (4.22) is a consequence of the remark stated below (3.27). The first inequality uses the lemma stated with (4.6), but is otherwise trivial. The second inequality follows from SP6 and its proof, together with the fact that $-U(\xi,\varepsilon)<\xi^{-1}$ pointwise, which was postulated at the beginning of the discussion of the spherical systems. The last inequality is a consequence of: (1) the analogue of SP3, which, of course, holds also for (4.16), and (2) the above-mentioned fact that $\tilde{\Phi}^*$ exists and is bounded. ▌

Remark. The last inequality in (4.22) shows that it is not necessary to know the value of η which generates $\beta(\eta)$. Rather, (4.22) tells us that the free energy of the shallow solutions, interpreted as a function of β, is bounded below as $(f(\tilde{\rho}))(\beta)>-\beta^{-1}C$, where C is a positive constant which is independent of β and ε. One should note that the bound (4.22) is not uniform over the interval $[0,\beta(\zeta^*)]$.

With the aid of some elementary algebra it is now readily shown that the right-hand side of (4.21), and thus $\inf f(\rho)$, is smaller than the right-hand side of (4.22), and thus smaller than $f(\tilde{\rho}_\beta)$, if $\beta>\hat{\beta}$, with

$$\hat{\beta}=c_1\,|\varepsilon\ln\varepsilon|+c_2\varepsilon \quad (4.23a)$$

The positive constants c_1 and c_2 are given by

$$c_1=6/|U(2,1)| \quad (4.23b)$$

and

$$c_2=2\ln\left(3\int_\Lambda e^{\tilde{\Phi}^*}\,d^3r\Big/4\pi\right)\Big/|U(2,1)| \quad (4.23c)$$

Clearly, we have $\hat{\beta}\searrow 0$ with $\varepsilon\searrow 0$, with the leading order in ε of $\hat{\beta}$ being $O(|\varepsilon\ln\varepsilon|)$, as ε approaches zero from above. Obviously, we can choose $\varepsilon\ll 1$ so small that $\hat{\beta}\ll\beta_*$, which we require to be the case from now on.

Recall that β_* is a lower bound for the existence of shallow solutions. Thus we have verified our above expectation: For small enough ε there exists a temperature regime $\hat{\beta}^{-1} > T > \beta_*^{-1}$ where there exist shallow solutions (4.12); however, $\inf f(\rho) < f(\tilde{\rho}_\beta)$ in that regime such that the thermodynamic equilibrium state is then not given by a shallow solution.

Let us pull things together. The shallow solutions given by (4.12) form an at least twice continuously differentiable sequence w.r.t. β inside their existence regime (see P3). Therefore, the free energy $f(\tilde{\rho}_\beta)$ of the shallow solutions is also at least twice continuously differentiable w.r.t. β in that regime of β values. The existence regime of the shallow solutions extends beyond β_*, (4.20). For $\beta < \breve{\beta}$, with $\breve{\beta} \ll \beta_*$, the shallow solutions are the (then) unique thermodynamic equilibrium states, which means $\inf f(\rho) = f(\tilde{\rho}_\beta)$ for $\beta < \breve{\beta}$. On the other hand, for β values lying in the interval $]\hat{\beta}, \beta(\zeta^*)[$, with $\breve{\beta} < \hat{\beta} \ll \beta_* < \beta(\zeta^*)$, we have shown [inequalities (4.21) and (4.22)] that $\inf f(\rho)$ is smaller than $f(\tilde{\rho}_\beta)$. This then implies the existence of a β_{tr} with $\breve{\beta} \leqslant \beta_{\text{tr}} \leqslant \hat{\beta}$, meaning $O(\varepsilon) \leqslant \beta_{\text{tr}} \leqslant O(|\varepsilon \ln \varepsilon|)$ as $\varepsilon \searrow 0$, where the thermodynamic equilibrium state changes somehow discontinuously from the solution branch of the shallow solutions to another solution branch of (3.40) which carries most likely solutions of core-atmosphere type. Note that a continuous change of the equilibrium density ρ at β_{tr} is impossible because the shallow solutions are locally stable (proof given above), which forbids a bifurcation at β_{tr}. We shall prove below that $\inf f(\rho)$ is continuous at β_{tr}. All these facts together then imply the existence of at least two distinct solutions of (3.40) at β_{tr}, both globally minimizing $f(\rho)$. This means the existence of a gravitational-type phase transition at β_{tr}. We also prove that the derivative of $\inf f(\rho)$ w.r.t. β is discontinuous at β_{tr} provided $\inf f(\rho)$ is differentiable in a finite right neighborhood of β_{tr}. In this sense the phase transition is of first order.

Proof. (a) [Continuity of $\inf f(\rho)$.] For $\beta < \beta_{\text{tr}}$ the infimum of f is given by $f(\tilde{\rho}_\beta)$. Assume that at $\beta = \beta_{\text{tr}}$ the infimum changes discontinuously with a finite jump Δ_f to $f(\rho_c)$, where ρ_c denotes the new thermodynamic equilibrium at $\beta = \beta_{\text{tr}}$. Clearly, ρ_c must solve (3.40) and is therefore strictly positive and bounded. Hence, for fixed $\rho \equiv \rho_c$ the mapping $\beta \mapsto (f(\rho_c))(\beta)$ is arbitrarily differentiable w.r.t. $\beta \neq 0$, which implies that at least in a small left neighborhood of β_{tr} we have $(f(\rho_c))(\beta) < f(\tilde{\rho}_\beta) = \inf f(\rho)$, which is a contradiction. Thus, $\inf f(\rho)$ is continuous.

(b) [Discontinuity of $\partial(\inf f(\rho))/\partial\beta$.] Let $\beta = \beta_{\text{tr}}$. Then there exist several distinct solutions of (3.40) with the same value of $\inf f$. It suffices to consider the case where there are exactly two distinct solutions, denoted by ρ_1 and ρ_2, where $\rho_1 = \tilde{\rho}_{\beta_{\text{tr}}}$. Write $f(\rho)$ as $f(\rho) = e(\rho) - \beta^{-1} s(\rho)$, where $e(\rho)$ is the potential energy per particle and $s(\rho)$ is the entropy per particle.

Assume that $e(\rho_1)=e(\rho_2)$. Then also $s(\rho_1)=s(\rho_2)$. That implies that there exists an incompressible mapping $\rho_1 \mapsto \rho_2$. (Note that entropy is conserved for incompressible mappings.) As has been shown in ref. 10 (see also refs. 8 and 21), any given ρ_0 can be mapped incompressibly to a unique spherical minimizer ρ_M of $e(\rho)$ with $s(\rho_0)=s(\rho_M)$. By construction both ρ_1 and ρ_2 minimize $e(\rho)$ under conservation of entropy; hence, $\rho_1 \equiv \rho_2$, in contradiction to the assumption that the densities are not identical. As a consequence, we must have $e(\rho_1) \neq e(\rho_2)$. Now note that from the finite-N theory it follows that $e(\rho_T)=\partial(\beta \inf f(\rho))/\partial\beta$, wherever the derivative exists, and where ρ_T means the thermodynamic equilibrium. With the mild assumption that the derivative of $\inf f(\rho)$ exists in a finite right neighborhood of β_{tr} our finding $e(\rho_1) \neq e(\rho_2)$ implies that $\partial(\inf f(\rho))/\partial\beta$ jumps at $\beta=\beta_{\text{tr}}$. ∎

Thus, we proved the existence of a first-order gravitational-type phase transition at $\beta=\beta_{\text{tr}}$, where β_{tr} fulfills $O(\varepsilon) \leqslant \beta_{\text{tr}} \leqslant O(|\varepsilon \ln \varepsilon|)$ as $\varepsilon \searrow 0$. More precisely, $\breve{\beta} \leqslant \beta_{\text{tr}} \leqslant \hat{\beta} \ll \beta_*$ for small enough ε. This means that β_{tr} is located well inside the parameter regime of the locally stable shallow solutions, which by SP6 also means well inside the parameter regime of β values where there exist locally stable isothermal Emden gas spheres.

Remark. The notion of "gravitational-type phase transition" refers to the fact that the systems considered here have only slightly smoothed-out gravitational interactions. It should be noted that, although we have proven rigorously the existence of a phase transition, we have not rigorously shown that then $f(\rho)$ actually takes its infimum for ρ a density of core-atmosphere type. However, the fact that the global minimizer of $f(\rho)$ converges to the delta distribution as $\varepsilon \to 0$ provides a good reason to conjecture that the global minimizer for $\beta > \beta_{\text{tr}}$ will be of core-atmosphere type. Clearly, for temperatures near absolute zero $f(\rho)$ will take its infimum for ρ a density of core-atmosphere type.

We would like to go a little farther and complete, to some extent, our knowledge of the solution properties of (3.40). Since U converges pointwise to U_{cl}, for $\varepsilon \ll 1$ to every solution of (4.16) which varies on a scale much larger than ε (but only then) there will correspond a solution of (4.6) which is only slightly deformed as compared to the former one. Hence, the bifurcation diagram of (4.6) will look qualitatively similar to that of (4.16), as long as the solutions are flat enough. The corresponding solutions of (3.40) computed from (4.6) will thus be very similar to their pendants for $\varepsilon=0$, which are to be computed from (4.16). In addition, for temperatures not too high there will exist the solutions of core-atmosphere type, which in the limit $\varepsilon=0$ go over into the Dirac point measure as weak solution of (3.40), which is then the only thermodynamic equilibrium state indeed. If in addition we make use of numerical results that have been obtained for the

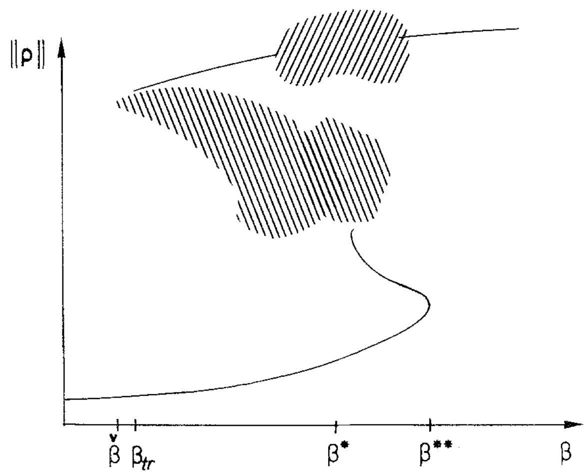

Fig. 1. Qualitative sketch of the bifurcation diagram of Eq. (3.40), showing the uniform norm of the equilibrium solutions ρ against the inverse temperature β. The lower branch represents the shallow solutions, which are locally stable, slightly inhomogeneous, self-gravitating gases. For $\beta \leqslant \check{\beta}$ they are the only solutions of (3.40), and for $\beta < \beta_{\mathrm{tr}}$ the thermodynamic equilibrium state is a shallow solution. For $\beta \leqslant \beta^*$ they are given by (4.12). They cease to exist for $\beta > \beta^{**}$. The upper branch contains the solutions which carry the thermodynamic equilibrium for $\beta > \beta_{\mathrm{tr}}$. They are most likely of core–atmosphere type. In the shaded regions the structure of the bifurcation diagram is largely unknown.

solutions of [(4.16), (4.17)] (see, for instance, refs. 11, 12, and 33), we can draw a qualitative sketch of the bifurcation diagram of (3.40) as shown in Fig. 1. (It should be noted that the existence of the gravitational-type phase transition has been shown by purely analytical techniques.) The whole class of the shallow solutions will now be defined as the solutions on the lowest branch, which exists for β values in the range $0 \leqslant \beta \leqslant \beta^{**} \approx 3$. The shallow solutions are locally stable and sharply bounded in norm, and connect differentiably to the homogeneous infinite-temperature state. We remark that from the local stability of the shallow solutions together with fact that $\inf f(\rho)$ is smaller than the free energy of the shallow solutions for $\beta > \beta_{\mathrm{tr}}$ there should follow the existence of at least one further (unstable) solution of (3.40) for $\beta_{\mathrm{tr}} < \beta < \beta^{**}$, to be proved by means of a mountain-pass lemma.

5. SUMMARY AND OUTLOOK

In the preceding sections we have shown that, in the weak sense defined in the introduction, it is meaningful to speak of a thermodynamic equilibrium state, in the canonical ensemble, of classical self-gravitating matter in three spatial dimensions confined to a finite container. We have explicitly constructed the exact canonical equilibrium measure on the configurational space of the particles. This has been achieved by first considering systems with smoothed-out gravitational potential and then taking the weak* limit of the exact Newtonian potential. The limiting measure is a linear superposition of single-point Dirac measures, meaning that every member system of the canonical ensemble is in the completely collapsed state, i.e., a single material point. This result holds both for finite systems (Section 2) and in the thermodynamic mean-field limit (Section 3).

We have also briefly discussed the analogous problem in two-dimensional physical space, and we have shown the existence of two well-separated temperature regimes. For very high temperatures the canonical equilibrium measure on the configurational space of the particles is absolutely continuous with respect to Lebesgue measure, and for very low temperatures we obtain again a superposition of single-point Dirac measures, describing the completely collapsed state. It is an open question whether in the intermediate-temperature regime the ensemble consists of completely collapsed systems or not. Here and below, "high" and "low temperature" are always meant with respect to a typical temperature characteristic of the system [see text pertaining to (2.66), and also text above (4.6)].

That shows that, in a sense, it is problematic to state that there cannot exist a thermodynamic equilibrium state for classical self-gravitating matter, as is commonly the case in the literature (cf. the discussion in the introduction). Rather, based on the results obtained in this paper it is proposed here to allow for a broader class of thermodynamic equilibrium states than is discussed usually, and to allow, generally speaking, also probability measures on the phase space that can be obtained as the weak limit of a sequence of usual thermodynamic equilibrium states, provided they behave reasonably in the limit $N \to \infty$ (see the discussion in the introduction). To discriminate between the usual equilibrium states and those which exist only in the weak closure it is proposed further to introduce the more refined distinction between strong and weak thermodynamic equilibrium, depending on whether the corresponding thermodynamic potential pertaining to the measure exists or not, respectively. In this sense, classical self-gravitating matter has a weak thermodynamic equilibrium state but no strong one.

In Section 4 we have investigated the thermodynamic mean-field limit for the related but physically more realistic problem of systems with smoothed-out gravitational interactions. These systems can be considered as a classical approximation to gravitating fermionic systems, in the sense that by cutting out the singularity of the Newtonian interaction potential the stabilizing quantum effects are mimicked. However, the classical "particles" can also be considered as cosmic objects, stars in a galaxy, for instance, the modification of the Newtonian interactions extending then over the size of a star. We inquired into the solution properties of the isothermal mean-field equation (3.40) as derived rigorously in ref. 15. Employing methods developed for nonlinear fixed-point problems in ordered Banach spaces,[30] we presented a constructive technique which allows us to compute solutions of (3.40) in arbitrary simply connected domains Λ. We proved that all the solutions computed by that technique are at least locally stable, meaning at least thermodynamically metastable, and that they belong to the parameter regime of moderately high temperatures up to infinity, connecting differentiably to a uniqueness regime existing[15] for high temperatures. Specializing to spherical systems, we have rigorously shown that a first-order phase transition will occur at temperatures located well inside the existence regime of the well-known isothermal Emden gas spheres, and we have presented strong evidence that the phase transition connects a slightly inhomogeneous gas phase and a phase consisting of a highly condensed core and a dilute atmosphere. The results of Section 4 can be viewed as classical counterparts of the thermodynamic Thomas–Fermi limit for gravitating fermions in the canonical ensemble, where the existence of a quantum mechanical gravitational phase transition was found numerically in ref. 34, an analytical proof of which was given in refs. 35 and 36.

The existence of a gravitational-type phase transition in the classical canonical ensemble has been proposed in previous work.[8,14] However, it seems that so far in all the work that addressed the problem to determine explicitly the critical temperature where the phase transition will occur as compared to the existence regime of the Emden isothermal gas spheres (ref. 14 did not address that problem), the critical temperature was identified with $1/\beta^{**}$ (Fig. 1), the analysis being based either on a discussion of the solution curve of Eq. (4.16)[8] and/or on the concept of local stability.[4,9,22] Both these concepts, however, cannot give the final answer to the question of what is the thermodynamic equilibrium state, since (1) Eq. (4.16) does not contain all the relevant solutions for the problem, as shown in Sections 3 and 4, and (2) a local stability analysis alone gives no information on global stability, which is the relevant concept for the statistical mechanics equilibrium state in the first place (e.g., ref. 15). In fact, as follows from the analysis of Section 4, the locally stable isothermal gas spheres, or their equivalents, are generally not the global minimizers of the relevant free-energy functional. The critical β_{tr} at which the phase transition occurs obeys $O(\varepsilon) \leqslant \beta_{\mathrm{tr}} \leqslant O(|\varepsilon \ln \varepsilon|)$, with $\varepsilon \ll 1$ by several orders of magnitude for the nearly-Newtonian gravitational interactions. So

$\beta_{tr} \ll 1$, whereas $\beta^{**} \approx 3$ (e.g., ref. 33). Thus, the actual point of phase transition differs extremely from β^{**}.

In the cited references[4,8,9,22] the identification of the transition temperature with $1/\beta^{**}$ is considered to be exact in the mean-field limit; possible deviations from that result are attributed to correlation effects that come in from finite-N corrections (see especially the corresponding discussions in refs. 4, 8, and 9). The proof that $\beta_{tr} \ll \beta^{**}$ (Section 4) uses, however, the mean-field limit. Thus, what has actually been determined in refs. 4, 8, 9, and 22 is a temperature regime where thermodynamically metastable self-gravitating equilibria exist, together with the critical point $1/\beta^{**}$ beyond which they cease to exist. Those metastable states might also play an important role, of course, reminiscent of a supercooled state of the van der Waals gas.

There are several remaining important questions directly related to the problems treated here which deserve a resolution. Here we comment at least briefly on some of them. A rigorous treatment of these problems has to come from future work.

I. We have investigated only the canonical ensemble. It is an interesting question how the microcanonical measure will look. Concerning the limit $\varepsilon \to 0$, since energy has to be conserved, it seems unlikely that the projected microcanonical measure on the configurational space for an isolated finite ($N < \infty$) system will converge exactly to the same limit (1.1) as does the projected canonical measure. One of the basic ingredients that allows the complete collapse to occur in the canonical ensemble is the possibility to give up an infinite amount of energy to the surrounding world, which is not possible in the microcanonical ensemble. So there must be some doubt about the equivalence of the ensembles, even for $\varepsilon \neq 0$. In the quantum mechanical situation it is known[37] that the various ensembles are not equivalent. Of course, our reasoning cannot replace rigorous considerations.

II. Concerning the systems with smoothed-out interactions, it is of interest to know whether the phase transition in the canonical ensemble bridges a region of negative specific heat in the microcanonical ensemble, which has been found to be the case in the quantum mechanical situation.[34] Clearly, the quantum mechanical results suggest that this will be true also in the classical case. There is some further evidence from exactly soluble approximate problems[14,38,39] which show some of the characteristic features expected for realistic classical self-gravitating matter, especially a gravitational-like phase transition in the canonical ensemble which is associated with a region of negative specific heat in the microcanonical ensemble. So one might indeed expect that feature to occur also for the exact classical problem with regularized Newtonian interactions.

At this point we mention explicitly the model of ref. 39. These authors discussed N point particles moving independently of one another on the surface of a sphere of variable radius. The spherical surface itself is confined between two concentric spherical walls of radii r_0 and $r_e \gg r_0$, and is subject to a central $-r^{-1}$ potential, the coupling strength being proportional to the square of the total mass of the combined system particles + spherical surface. That model is exactly soluble in both microcanonical and canonical ensemble, for both classical and quantum mechanical situations. It is perhaps worth noting that in that model the critical temperature β_T^{-1} for the phase transition in the classical canonical ensemble is given by $\beta_T = O(|(r_0/r_e)\ln(r_0/r_e)|)$ to the leading order in r_0/r_e. That is in remarkable agreement with our estimate for $\beta_{\rm tr}$ for the exact problem of classical particles with regularized interactions.

III. Again concerning the systems with smoothed-out interactions, it is of interest to know also the deviation of the finite-N results from the results obtained from the mean-field limit, at least the correction to the mean-field limit result in the leading power of N. To get an idea of this, consider inequality (2.26) for $F(\mu)$ given by (2.30). Instead of taking the box distribution (2.31) at will, we might rather seek the global minimum of (2.30) with respect to ρ. The free-energy functional (2.30), evaluated at its global minimum, is just the mean-field approximation to the exact free energy; equivalently, it is the finite-N pendant to the mean-field limit expression (3.26). The global minimum exists as a consequence of the work presented in ref. 15. As a necessary condition the minimizing ρ has to solve the Euler–Lagrange equation

$$\rho(\mathbf{r}) = \frac{\exp\{-\beta[\int_A \rho(\mathbf{r}')(N-1)\,V(|\mathbf{r}-\mathbf{r}'|,\varepsilon)\,d^3r' + m\phi(\mathbf{r})]\}}{\int_A \exp\{-\beta[\int_A \rho(\mathbf{r}')(N-1)\,V(|\mathbf{r}''-\mathbf{r}'|,\varepsilon)\,d^3r' + m\phi(\mathbf{r}'')]\}\,d^3r''} \tag{5.1}$$

In the case where the external gravitational potential vanishes, comparing (5.1) with (3.27) and noting the mean-field scaling (3.15a) reveals that the best approximation of $F^{\rm conf}$ from above by $\mathscr{F}(\rho)$ is given for ρ a measure obtained from the mean-field limit, for a limit temperature that deviates by terms of order N^{-1} from the temperature of the finite system. Precisely, $\beta(1-N^{-1}) = \beta_\infty$, where β_∞ denotes the inverse temperature that has to be chosen for the mean-field limit. So far we do not know how good this "best approximation of $F^{\rm conf}$ from above" actually is. According to ref. 4, the leading corrections are in fact $\sim N^{-1}$, that result being obtained from a formal series expansion around the mean-field equilibrium state, which was evaluated up to second order.

In refs. 4 and 22 an equivalent version of (5.1) has been derived by means of a different method based on an application of Jensen's inequality. Because of the similarity of (5.1), in the spherical case, with the Emden gas sphere equation these authors conjectured that an Emden isothermal gas sphere will give the best mean-field approximation to the exact equilibria, which is generally not the case, as shown in Sections 3 and 4 of the present paper. The formal series expansion method mentioned above is not based, however, on the prescription of any special solution of (5.1).

It should be noted, however, that even if the leading corrections to the mean-field results are $\sim N^{-1}$, the actual deviations of the finite-N results need not be small in the neighborhood of the phase transition. It should be noted further that the fine details of the collapsed phase must not be taken too seriously, even if the deviations of the finite systems from the mean-field results are of $O(N^{-1})$ and also small. The reason is simply that our mean-field approach (recall that we took a limit where two-particle interactions are small) is not an adequate description for realistic systems if the typical interparticle distances are of the order of the particle sizes itself, which is the case in the collapsed phase. Short-range repulsive forces [$\gamma \neq 0$ in (2.1)] and correlations will then be important. However, provided that (1) the typical interparticle distances are much larger than the particles sizes itself if the N particles are distributed statistically homogeneously over the domain Λ, and (2) the range r_0 of the repulsive forces is of the size of the particles, i.e., of $O(\varepsilon)$, then the infimum of (2.30) behaves qualitatively exactly as that of (3.26). Note that then the high-temperature regime is described adequately by our mean-field approach, and further that (2.30) obeys the inequality (2.26). Our finding of the existence of a gravitational phase transition in the mean-field limit then implies that the more realistic exact finite-N systems, i.e., where the correlations and repulsive forces are not neglected, also show a kind of smooth phase change. This means that the state of the system changes drastically but differentiably (since N is finite) from a self-gravitating gas phase to a collapsed phase. Drastically means within a very narrow region of β values. (From the above it should be clear that our results might be applicable to dilute self-gravitating systems in space and not to collision-dominated classical gases with strong repulsive short-range forces and weak gravitational interactions.)

IV. The results derived in the present paper are based on the prescription of the equilibrium measure (2.9). Although (2.9) is the usual canonical equilibrium measure, the crucial question is whether that measure actually describes the time-asymptotic state of the dynamical evolution, in the usual average sense, of classical self-gravitating matter with regularized interactions when that matter is in a finite box and subject to a thermal contact with a heat bath. This presents the problem of the dynamical accessibility of the derived equilibrium structures. Analytical studies of the dynamical evolution of self-gravitating matter are extremely difficult. For the systems with smoothed-out interactions it should be possible, however, to test at least via particle simulation studies whether the systems tend to evolve into core–atmosphere type structures or not, when the temperature is still high enough to allow also the existence of shallow equilibria but is not too high, of course. Note that in particle simulation studies the particles are necessarily smeared out, which is equivalent to having smoothed-out interactions. So far, particle simulation studies (see, for instance, the discussion in ref. 3) do indeed often show a nucleation of the system.

A possible verfication or falsification of the accessibility of the infimum state of the free-energy functional (3.26) has necessarily to take into account the confinement of the system to a finite container. In the first place the container is a mathematical tool that allows the calculation of equilibrium structures. The following shows that a container need not be a completely unrealistic assumption. A large, extended, self-gravitating system in space will surely show both a tendency of (the inner) parts of the system to shrink and a tendency of (the outer) parts to "evaporate." If the time scales for both processes are well separated, then a system in a container can be interpreted as the idealization where the ratio of the evaporation time scale to the collapse time scale becomes infinite. The mere existence of planets, stars, and galactic nuclei might be viewed at as a hint that there is something realistic to that concept.

If we consider self-gravitating matter in all $\mathbb{R}^3$, which is a more realistic model for classical astrophysical applications, considerable doubt has been expressed concerning the applicability of calculating equilibria by means of concepts based on extremizing entropy-like functionals[(7)] (see also the discussion in ref. 40). By studying the influence of weak gravitational collisions[(41,42)] it has also been questioned whether a self-gravitating system should show the tendency to thermalize at all. Considering the dynamical counterpart of the equilibrium mean-field limit, i.e., the collisionless Vlasov dynamics (see ref. 43 for a derivation of the Vlasov dynamics in the case of classical bounded interactions, and ref. 44 for the quantum mechanical pendant) with exact Newtonian interactions, a rigorous proof[(45)] has been given that the gravitational energy of an isolated self-gravitating system in $\mathbb{R}^3$ is bounded away from minus infinity for all times provided (1) the total energy is negative, and (2) the initial distribution function is bounded and has compact support on the single-particle phase space. At least for a collisionless isolated system in all $\mathbb{R}^3$ this is a proof that purely self-gravitating matter with smooth initial data will generally neither develop a point singularity nor thermalize. From these considerations it is to be expected that the concept of a thermal contact on one hand and the inclusion of correlations into the dynamics on the other will play decisive roles in future investigations of whether the collapse of self-gravitating matter is dynamically accessible or not.

ACKNOWLEDGMENT

I thank K. Schindler for a critical reading of the manuscript and for valuable discussion. I have benefitted also from useful comments by D. Dürr and S. Albeverio. Many interesting discussions with H. Wiechen and H. J. Ziegler are gratefully acknowledged. The work of ref. 4 was pointed out to me by T. Neukirch. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 162).

REFERENCES

1. A. M. Salzberg, *J. Math. Phys.* **6**:158 (1965).
2. G. B. Rybicki, *Astrophys. Space Sci.* **14**:56 (1971).
3. W. C. Saslaw, *Gravitational Physics of Stellar and Galactic Systems* (Cambridge University Press, Cambridge, 1985).
4. P. S. Cally, *Aust. J. Phys.* **34**:267 (1981).
5. J. S. Nicolis, *Dynamics of Hierarchical Systems* (Springer-Verlag, New York, 1986).
6. W. C. Saslaw, *Mon. Not. R. Astron. Soc.* **141**:1 (1968).
7. J. R. Ipser, *Astrophys. J.* **193**:463 (1974).
8. D. Lynden-Bell and R. Wood, *Mon. Not. R. Astron. Soc.* **138**:495 (1968).
9. G. Horwitz and J. Katz, *Astrophys. J.* **222**:941 (1978).
10. V. A. Antonov, *Vest. Leningr. Gas. Univ.* **7**:135 (1962).
11. R. Emden, *Gaskugeln* (Teubner Verlag, Leipzig, 1907).
12. S. Chandrasekhar, *An Introduction to the Study of Stellar Structure* (University of Chicago Press, Chicago, 1939).
13. W. M. Suen and K. Young, *Phys. Rev.* **35**:406 (1987).
14. W. Thirring, *Z. Phys.* **235**:339 (1970).
15. J. Messer and H. Spohn, *J. Stat. Phys.* **29**:561 (1982).
16. D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969).
17. W. Thirring, *Lehrbuch der Mathematischen Physik 4* (Springer-Verlag, Vienna, 1980).
18. E. Hewitt and L. J. Savage, *Trans. Am. Math. Soc.* **80**:470 (1955).
19. L. D. Landau and E. M. Lifshitz, *Course in Theoretical Physics*, Vol. 5 (Pergamon Press, Oxford, 1969).
20. J. Messer, *Temperature Dependent Thomas–Fermi Theory* (Springer, Berlin, 1981).
21. E. H. Lieb and H. T. Yau, *Commun. Math. Phys.* **112**:147 (1987).
22. P. S. Cally and J. J. Monaghan, *J. Math. Phys.* **22**:348 (1981).
23. M. Reed and B. Simon, *Functional Analysis*, 2nd ed. (Academic Press, New York, 1980).
24. Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick, *Analysis, Manifolds, and Physics*, 2nd ed. (North-Holland, New York, 1982).
25. P. C. Hemmer and J. L. Lebowitz, in *Phase Transitions and Critical Phenomena*, Vol. 5b, C. Domb and M. S. Green, eds. (Academic Press, London, 1976).
26. L. van Hemmen, *Fortschr. Phys.* **26**:397 (1978).
27. M. G. Krein and M. A. Rutman, *Am. Math. Soc. Transl.* **10**:199 (1962).

28. J. Glimm and A. Jaffe, *Quantum Physics* (Springer, New York, 1981).
29. H. Fujita, *Bull. Am. Math. Soc.* **75**:132 (1969).
30. H. Amann, *SIAM Rev.* **18**:620 (1976).
31. H. Amann, *J. Diff. Eq.* **21**:363 (1976).
32. I. M. Gel'fand, *Am. Math. Soc. Transl.* **29**:295 (1963).
33. D. D. Joseph and T. S. Lundgren, *Arch. Rat. Mech. Anal.* **49**:241 (1973).
34. P. Hertel and W. Thirring, in *Quanten und Felder*, H. P. Dürr, ed. (Vieweg, Brunswick, 1971).
35. J. Messer, *J. Math. Phys.* **22**:2910 (1981).
36. J. Messer, *Phys. Lett. A* **83**:304 (1981).
37. J. Messer, *Z. Phys.* **33**:313 (1979).
38. P. Hertel and W. Thirring, *Ann. Phys.* (NY) **63**:520 (1971).
39. D. Lynden-Bell and R. M. Lynden-Bell, *Mon. Not. R. Astron. Soc.* **181**:405 (1977).
40. P. T. Landsberg, *J. Stat. Phys.* **35**:159 (1984).
41. I. Prigogine and G. Severne, *Physica* **32**:1376 (1966).
42. H. E. Kandrup, *Astrophys. Space Sci.* **112**:215 (1985).
43. W. Braun and K. Hepp, *Commun. Math. Phys.* **56**:101 (1977).
44. H. Narnhofer and G. L. Sewell, *Commun. Math. Phys.* **79**:9 (1981).
45. H. Wiechen, H. J. Ziegler, and K. Schindler, *Mon. Not. R. Astron. Soc.* **232**:623 (1988).