TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 263, Number 1, January 1981

NONSTANDARD ANALYSIS AND LATTICE STATISTICAL MECHANICS: A VARIATIONAL PRINCIPLE

BY

A. E. HURD[1]

Dedicated to the memory of Abraham Robinson

ABSTRACT. Using nonstandard methods we construct a configuration space appropriate for the statistical mechanics of lattice systems with infinitely many particles and infinite volumes. Nonstandard representations of generalized equilibrium measures are obtained, yielding as a consequence a simple proof of the existence of standard equilibrium measures. As another application we establish an extension for generalized equilibrium measures of the basic variational principle of Landord and Ruelle. The same methods are applicable to continuous systems, and will be presented in a later paper.

1. Introduction. The early history of statistical mechanics is characterized by calculations which establish elaborate interrelationships between the fundamental notions but the calculations are mostly formal-convergence questions arising when the number of particles and volume approach infinity are ignored. The step into rigorous statistical mechanics was taken in the works of Van Hove, Yang and Lee, Fisher and Ruelle on the existence of thermodynamic limits [**6**]. Since that time the thrust has been to place not only the thermodynamic quantities, but the whole edifice of statistical mechanics on a rigorous footing. An important step in that development involves defining an appropriate "limiting" configuration space. To quote Minlos [**11**]: "This question, despite its mathematical sophistication, is of great physical interest: essentially it is a matter of being able to describe correctly an 'infinite physical system', that is, a system containing infinitely many particles .... This description ought to be such that all the thermodynamic quantities usually obtained as limits of parameters of the finite ensemble serve as the corresponding parameters of the infinite ensemble." This can be interpreted in the following way. The basic real world of statistical mechanics resides in a family $\mathcal{H} = \langle (X_i, \mathcal{F}_i) \rangle$ $(i \in I)$ of measurable spaces $(X_i, \mathcal{F}_i)$, each of which is finite in an appropriate sense (finite volume, finite number of particles) and possesses considerable attendant structure (measures correlation functions, dynamics, etc.). [Indeed, one could argue that systems containing infinitely many particles do not actually exist in nature, but only families of finite systems of arbitrarily large cardinality.]

Received by the editors May 8, 1979.

AMS (MOS) subject classifications (1970). Primary 02H25, 82A05; Secondary 60B99, 28A20.

Key words and phrases. Lattice statistical mechanics, generalized equilibrium (Gibbs) state, variational principle, limiting configuration space.

[1]Supported in part by NRC grant A-8198.

Then statistical mechanics is the investigation of limiting *measure theoretic* phenomena within $\mathcal{H}$. One wants to define a "limiting" measurable space on which the limiting measure theoretic phenomena can be investigated easily. From this point of view, any such limiting space is an "ideal" structure whose sole role is to make the analysis of limiting phenomena in $\mathcal{H}$ more amenable. The limiting space should satisfy two properties:

(a) There is a class of limiting measure theoretic entities (e.g. measures measurable transformations, etc.) on the limiting space such that each family of corresponding entities on the $(X_i, \mathcal{F}_i)$, which converges in an appropriate sense, corresponds to a well-defined limiting entity in that class.

(b) Any result stated in terms of the limiting entities of (a) should be translatable to a limiting measure theoretic result on the real world of the family $\mathcal{H}$.

At this point in time candidates for the correct limiting space in lattice and continuous have gained common acceptance (see e.g., [7], [11], [16]). But carrying out (a) for these spaces sometimes involves considerable technical difficulties. For example there is the problem of limiting dynamics: in continuous statistical mechanics one wants to define a limiting measurable transformation on the limiting space which is induced by the Newtonian dynamics on the $(X_i, \mathcal{F}_i)$. This problem, essential to nonequilibrium statistical mechanics, is still in an unsatisfactory state (see [7] and the references therein).

Many of these technical problems arise because the commonly accepted limiting spaces, which we denote generically by $(X, \mathcal{F})$, are too small to accommodate (a) easily. We present here the foundations of an approach to rigorous statistical mechanics using nonstandard analysis. The starting point is the introduction of a new limiting space $(\mathbf{X}, \boldsymbol{\mathcal{F}})$ which plays the same role as $(X, \mathcal{F})$ but is large enough so that (a) holds. Even though $(\mathbf{X}, \boldsymbol{\mathcal{F}})$ is very large it also, perhaps surprisingly, satisfies (b). Although it is not essential to do so, we show in addition that each measure on $(X, \mathcal{F})$ induces a measure on $(\mathbf{X}, \boldsymbol{\mathcal{F}})$ in a canonical way. Thus the measure theoretic structure on $(\mathbf{X}, \boldsymbol{\mathcal{F}})$ can in some sense be regarded as an extension of that on $(X, \mathcal{F})$. On the other hand, there are measures on $(\mathbf{X}, \boldsymbol{\mathcal{F}})$ which do not in any way correspond to measures on $(X, \mathcal{F})$, so that the measure theoretic structure of $(\mathbf{X}, \boldsymbol{\mathcal{F}})$ is much richer than that of $(X, \mathcal{F})$. As a consequence it is important to remark that results on $(\mathbf{X}, \boldsymbol{\mathcal{F}})$ do not necessarily yield results on $(X, \mathcal{F})$; however, they always yield limiting statements on the real world of $\mathcal{H}$ by (b). Indeed, it is precisely because $(\mathbf{X}, \boldsymbol{\mathcal{F}})$ is a richer structure that there is a possibility of proving results not envisaged in the usual analysis on $(X, \mathcal{F})$.

Our limiting configuration space $(\mathbf{X}, \boldsymbol{\mathcal{F}})$ is defined using the nonstandard models first introduced by Abraham Robinson in 1961 [14]. Given $\mathcal{H} = \langle (X_A, \mathcal{F}_A) \rangle$ $(A \in I)$, where I is a directed set usually indexing number of particles and volume, then in an appropriately large nonstandard model we consider a nonstandard space $(\mathbf{X}_\Omega, \boldsymbol{\mathcal{F}}_\Omega)$ where Ω is an infinite element of the nonstandard extension $\mathbf{I}$ of I. By using a central result called the Transfer Principle [2] or Leibniz' Principle [17], it is easy to investigate the structure of $(\mathbf{X}_\Omega, \boldsymbol{\mathcal{F}}_\Omega)$. Indeed it follows by transfer that $(\mathbf{X}_\Omega, \boldsymbol{\mathcal{F}}_\Omega)$ possesses nonstandard analogues of all of the structure (dynamics, correlation functions, etc.) which exist on each of the $(X_A, \mathcal{F}_A)$. $\boldsymbol{\mathcal{F}}_\Omega$ is an algebra of subsets

of $\mathbf{X}_\Omega$ and, following ideas of Loeb [10], we let $\mathbf{X} = \mathbf{X}_\Omega$, and $\mathfrak{F}$ be the σ-algebra generated by $\mathfrak{F}_\Omega$. The resulting *standard* measurable space is, in some conventional sense, very large, but still satisfies (a) and (b). The nonstandard entities on $(\mathbf{X}_\Omega, \mathfrak{F}_\Omega)$ can then be used to induce measure theoretic entities on $(\mathbf{X}, \mathfrak{F})$. Note that establishing (b) amounts precisely to showing how to bring results on $(\mathbf{X}, \mathfrak{F})$ back down to results in the real world of $\mathcal{H}$ and makes the nonstandard technique more than just a formal exercise in large "ideal" structures.

After reviewing the standard theory of lattice statistical mechanics in §2, we introduce the basic nonstandard ideas, and, in particular, the notion of generalized equilibrium state, in §3. In §4 we establish properties (a) and (b) for measures on $(\mathbf{X}, \mathfrak{F})$, and connections between measures on the standard $(X, \mathfrak{F})$ and those on $(\mathbf{X}, \mathfrak{F})$. In particular, the notion of generalized equilibrium state is shown to be an extension of the usual notion, and a new proof of the existence of standard equilibrium states results. In §5 we present a typical application of the nonstandard techniques by proving a variational principle for the generalized equilibrium states. The result provides a partial justification for the definition of these states. The nonstandard proof is a transfer of finitistic arguments and thus essentially elementary.

These techniques have many other applications, not only to statistical mechanics, but also to stochastic processes. In a later paper we will present applications to continuous models in statistical mechanics. In particular we will show how the problem of limiting dynamics can be handled in the nonstandard setting. In earlier papers [12] Ostebee, Gambardella and Dresden have applied nonstandard techniques to study the thermodynamic limit. Recently Helms and Loeb [4] have considered the stochastic theory of infinite particle systems from a nonstandard viewpoint. Finally, Anderson [1] and Keisler [6] have developed a nonstandard approach to stochastic integral equations.

The basic ideas in this paper were announced at the Symposium on Abraham Robinson's Theory of Infinitesimals held at the University of Iowa (May 31–June 5, 1977), and at the 1977 summer meeting of the American Mathematical Society in Seattle [5]. I wish to express my indebtedness to Marvin Shinbrot and G. V. Ramanathan for introducing me to statistical mechanics. In addition I want to thank Marvin Shinbrot for his constant encouragement and advice. Finally, many thanks to the referee for suggestions which led to improvements in an earlier version of this paper.

2. Standard lattice statistical mechanics. In this section we review the structure of lattice statistical mechanics. For reference see §5 in [13]. Throughout the paper we use the notation $F(X, \mathfrak{F})$, $M(X, \mathfrak{F})$ and $P(X, \mathfrak{F})$, or simply $F(\mathfrak{F})$, $M(\mathfrak{F})$ and $P(\mathfrak{F})$ when X is understood, to denote the sets of real valued measurable functions, positive measures, and probability measures, respectively, on a measurable space $(X, \mathfrak{F})$.

Let S be a countable set representing the sites. Denote by $\mathcal{C}$ the collection of all finite subsets of S, partially ordered and directed upwards by inclusion $\subseteq$. Let I be a conveniently chosen cofinal subset of $\mathcal{C}$. For example if $S = Z$, the integers, then I could be the set of all subsets of Z of the form $\{x \in Z \mid -n \leqslant x \leqslant n\}$. For

convenience we suppose that the possible states at each $s \in S$ come from a fixed finite set Y (this is the most important case in applications). The standard limiting configuration space is the set $X = Y^S$, the set of all maps $x: S \to Y$, each of which is called a configuration. Similarly, for any subset A of S we let $X_A = Y^A$. There are natural projection maps $p_A: X \to X_A$, and $p_{AB}: X_B \to X_A$, $A \subseteq B$, defined by restriction, e.g. $p_A(x)(s) = x(s)$, $s \in A$. We will write x_A for $p_A(x)$, $x \in X$, or $p_{AB}(x)$, $x \in X_B$, letting the context make clear the distinction.

If $A \in \mathcal{C}$ then X_A is a finite set and we let $\mathcal{F}_A$ be the collection of all (finite) subsets of X_A. There is thus a natural projective family[2] $\langle (X_A, \mathcal{F}_A), p_{AB} \rangle$ $(A, B \in I)$ associated with our lattice model. Let $\mathcal{F}_{(A,B)} = p_{AB}^{-1}(\mathcal{F}_A)$ and $\mathcal{F}_{(A)} = p_A^{-1}(\mathcal{F}_A)$. If $\mathcal{F}$ is the σ-algebra on X generated by the algebra $\mathcal{F}_0 = \bigcup \mathcal{F}_{(A)}$ $(A \in I)$ of cylinder sets, then $\langle (X, \mathcal{F}), p_A \rangle$ $(A \in I)$ is the projective limit of the projective family $\langle (X_A, \mathcal{F}_A), p_{AB} \rangle$ $(A, B \in I)$, and is the standard limiting probability space for lattice models. Also, for $\Lambda \in \mathcal{C}$, $\mathcal{F}^\Lambda$ is the σ-subalgebra of $\mathcal{F}$ generated by $\bigcup \mathcal{F}_{(A-\Lambda)}$ $(A \in I)$, and is denoted by $\mathcal{F}_\Lambda$ in **[13]**. In analogy with $\mathcal{F}^\Lambda$ we define $\mathcal{F}_A^\Lambda = p_{(A-\Lambda)A}^{-1}(\mathcal{F}_{A-\Lambda})$. Note that

$$p_{AB}^{-1}(\mathcal{F}_A^\Lambda) \subset \mathcal{F}_B. \tag{2.1}$$

To define equilibrium (Gibbs) states we need to introduce interaction potentials into the picture. Let $b \in X$ be a reference configuration (whose potential energy will be zero). An *interaction potential* is a family $\langle \phi_\Lambda \rangle$ $(\Lambda \in \mathcal{C})$ of maps ϕ_Λ: $X_\Lambda \to R$ such that

$$\phi_\Lambda(w) = 0 \quad \text{if } w(t) = b(t) \text{ for some } t \in \Lambda \tag{2.2}$$

(with the convention that if $\Lambda = \varnothing$, the empty set, then $\phi_\Lambda = 0$).

Later on we will need the following assumptions. Let

$$\|\phi_\Lambda\| = \sup|\phi_\Lambda(y)| \ (y \in X_\Lambda) \quad \text{and} \quad V(\Lambda) = \{C \in \mathcal{C} | C \cap \Lambda \neq \varnothing\}.$$

2.1. Assumption. For each $t \in S$, $\Sigma\|\phi_\Lambda\|$ $(\Lambda \in V(\{t\}))$ is finite.

2.2. Assumption. There is a number K so that $\Sigma\|\phi_\Lambda\|$ $(\Lambda \in V(\{t\})) < K$.

Assumption 2.1 is always in force in the standard development **[13**, Proposition 5.2], and 2.2 is a uniform version of 6.1 which holds when the interaction potential is translation invariant. For $A \in I$, the set

$$\mathcal{C}_A = \{\Lambda \in \mathcal{C} | \Lambda \subseteq A\} \tag{2.3}$$

is an initial segment of $\mathcal{C}$. Using the interaction potentials we define a family $\langle g_A^\Lambda \rangle$ $(\Lambda \in \mathcal{C}_A, A \in I)$ of potential energy functions g_A^Λ: $X_A \to R$ by

$$g_A^\Lambda(x) = \sum \phi_C(x_C) \qquad (C \in V_A(\Lambda)) \tag{2.4}$$

where $V_A(\Lambda) = \{C \in \mathcal{C}_A | C \cap \Lambda \neq \varnothing\}$. In terms of the function g_A^Λ we define the

[2]A projective family $\langle (X_A, \mathcal{F}_A), p_{AB} \rangle$ $(A, B \in I)$ consists of the measurable spaces $(X_A, \mathcal{F}_A)$ $(A \in I)$ and measurable maps $p_{AB}: X_B \to X_A$ satisfying $p_{AB} \circ p_{BC} = p_{AC}(A \subseteq B \subseteq C)$. $\langle (X, \mathcal{F}), p_A \rangle$ $(A \in I)$ is the projective limit of $\langle (X_A, \mathcal{F}_A), p_{AB} \rangle$ $(A, B \in I)$ if the $p_A: X \to X_A$ are measurable and $p_{AB} \circ p_B = p_A$, $A \subseteq B$. If $\mu \in M(X_B, \mathcal{F}_B)$ we define the measure $p_{AB}(\mu) \in M(X_A, \mathcal{F}_A)$ by $p_{AB}(\mu)(F) = \mu(p_{AB}^{-1}(F))$, $F \in \mathcal{F}_A$.

functions $f_A^\Lambda: X_A \to R$ ($\Lambda \in \mathcal{C}_A$, $A \in I$) by

$$f_A^\Lambda(x) = Z_A^\Lambda(x_{A-\Lambda})^{-1} \exp(g_A^\Lambda(x)) \tag{2.5}$$

where

$$Z_A^\Lambda(y) = \sum \exp g_A^\Lambda(y_{A-\Lambda} \times w) \qquad (w \in X_\Lambda), y \in X. \tag{2.6}$$

The Z_A^Λ are called *partition functions*. Finally we define the family of kernels $\langle \pi_A^\Lambda \rangle$ ($\Lambda \in \mathcal{C}_A$, $A \in I$), $\pi_A^\Lambda: X_A \times \mathcal{F}_A \to R$, by

$$\pi_A^\Lambda(x, F) = \sum f_A^\Lambda(x_{A-\Lambda} \times w) \qquad (w \in W_A^\Lambda(x, F)) \tag{2.7}$$

where $W_A^\Lambda(x, F) = \{w \in X_\Lambda | x_{A-\Lambda} \times w \in F\}$. Notice that for given F, $\pi_A^\Lambda(x, F)$ depends only on the restriction of x to $A - \Lambda$ and that

$$\pi_A^\Lambda(y, \{x\}) = \begin{cases} f_A^\Lambda(x) & \text{if } x_{A-\Lambda} = y_{A-\Lambda}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.8}$$

It is easy to check, as in [13], that $\pi_A^\Lambda(x, \cdot)$ is a probability measure on $\mathcal{F}_A$ for each $x \in X_A$ and $\Lambda \in \mathcal{C}_A$.

Any measure $\nu \in P(\mathcal{F}_A^\Lambda)$ is specified by a density $\delta: X_{A-\Lambda} \to R$ where $\delta(x) \geqslant 0$ and $\sum \delta(x)$ ($x \in X_{A-\Lambda}$) $= 1$. Then the family $\langle t_A^\Lambda \rangle$ ($A \in I$, $\Lambda \in \mathcal{C}_A$) of maps t_A^Λ: $P(\mathcal{F}_A^\Lambda) \to P(\mathcal{F}_A)$ given by

$$t_A^\Lambda(\nu)(F) = \sum \delta(y_{A-\Lambda}) \pi_A^\Lambda(y, F) \qquad (y \in X_A) \tag{2.9}$$

satisfies

$$r_A^\Lambda \circ t_A^\Lambda = \text{identity}, \qquad \Lambda \in \mathcal{C}_A, \tag{2.10}$$

and

$$t_A^\Lambda \circ r_A^\Lambda \circ t_A^{\tilde{\Lambda}} = t_A^{\tilde{\Lambda}}, \qquad \Lambda \subseteq \tilde{\Lambda} \text{ in } \mathcal{C}_A, \tag{2.11}$$

where r_A^Λ: $P(\mathcal{F}_A) \to P(\mathcal{F}_A^\Lambda)$ is *restriction*. The map t_A^Λ is the finitistic analogue of the map t_Λ defined in §2 of [13]. The measure $t_A^\Lambda(\nu)$ has a density σ which assigns to each $x \in X_A$ the probability

$$\sigma(x) = t_A^\Lambda(\nu)(\{x\}) = \delta(x_{A-\Lambda}) f_A^\Lambda(x). \tag{2.12}$$

3. Nonstandard lattice models. We assume that the reader is familiar with the foundations of nonstandard analysis as presented in [2], [15] or [17]. Let $\mathfrak{M}$ be a superstructure based on a set which includes all of the standard sets encountered in §2. The nonstandard analysis will be carried out in a κ-saturated nonstandard model $^*\mathfrak{M}$ of $\mathfrak{M}$ [17] where κ is sufficiently large (to be specified shortly), but at least $\aleph_1$ so that $^*\mathfrak{M}$ is denumerably comprehensive.

Star transforms of standard entities τ will be denoted by $^*\tau$, but internal entities will be denoted by bold face symbols. For example, if A is a standard set then its *-transform in $^*\mathfrak{M}$ will be denoted by $^*A$, but if $\sigma = \{A_i | i \in I\}$ is a collection of standard sets A_i indexed by I, then we write $^*\sigma = \boldsymbol{\sigma} = \{\mathbf{A}_i | i \in \mathbf{I}\}$ so that $\mathbf{I} = {}^*I$ and for $i \in I$ (such an i is called *standard*), $\mathbf{A}_i = {}^*A_i$.

As usual, $^\circ r$ or $\mathrm{st}(r)$ will be used to denote the standard part of a finite number $r \in {}^*R$ where R is the set of real numbers and we write $r \simeq s$ if $^\circ(r - s) = 0$. If

$r \in {}^*R$ is positive or negative infinite, we also put ${}^\circ r = \infty$ or ${}^\circ r = -\infty$, respectively.

The projective family $\langle (X_A, \mathcal{F}_A), p_{AB} \rangle$ $(A, B \in I)$ has a *-transform

$$^*\{\langle (X_A, \mathcal{F}_A), p_{AB} \rangle\ (A, B \in I)\} = \langle (\mathbf{X_A}, \boldsymbol{\mathcal{F}}_\mathbf{A}), \mathbf{p_{AB}} \rangle \qquad (\mathbf{A}, \mathbf{B} \in \mathbf{I}).$$

By transfer we see immediately that each $\boldsymbol{\mathcal{F}}_\mathbf{A}$ is an internal algebra of subsets of $\mathbf{X_A}$, but is not necessarily closed under countable unions (since this is an external operation) and so is not necessarily a σ-algebra in the standard sense. We will call $(\mathbf{X_A}, \boldsymbol{\mathcal{F}}_\mathbf{A})$ an *internally measurable space*. Similarly, the internal maps $\mathbf{p_{AB}}$: $\mathbf{X_B} \to \mathbf{X_A}$ $(\mathbf{A}, \mathbf{B} \in \mathbf{I})$ are measurable in the sense that $\mathbf{p_{AB}^{-1}(F)} \in \boldsymbol{\mathcal{F}}_\mathbf{B}$ for all $\mathbf{F} \in \boldsymbol{\mathcal{F}}_\mathbf{A}$ and $\mathbf{p_{AB}} \circ \mathbf{p_{BC}} = \mathbf{p_{AC}}$ for $\mathbf{A} \subseteq \mathbf{B} \subseteq \mathbf{C}$ in $\mathbf{I}$. On each internally measurable space $(\mathbf{X_A}, \boldsymbol{\mathcal{F}}_\mathbf{A})$ there are the sets $\mathbf{F}(\mathbf{X_A}, \boldsymbol{\mathcal{F}}_\mathbf{A})$, $\mathbf{M}(\mathbf{X_A}, \boldsymbol{\mathcal{F}}_\mathbf{A})$, and $\mathbf{P}(\mathbf{X_A}, \boldsymbol{\mathcal{F}}_\mathbf{A})$ of internal $^*R$-valued measurable functions, internal $^*R$-valued measures, and internal $^*R$-valued probability measures, respectively. If $\boldsymbol{\mu}$: $\boldsymbol{\mathcal{F}}_\mathbf{A} \to {}^*R^+$ (the set of positive reals) is in $\mathbf{M}(\mathbf{X_A}, \boldsymbol{\mathcal{F}}_\mathbf{A})$, then $\boldsymbol{\mu}$ is finitely additive in the usual sense, and if $\boldsymbol{\mu}$ is in $\mathbf{P}(\mathbf{X_A}, \boldsymbol{\mathcal{F}}_\mathbf{A})$ then $\boldsymbol{\mu}(\mathbf{X_A}) = 1$. All of these facts are immediate consequences of the transfer principle.

We next show how to find a nonstandard replacement for the standard limiting probability space $(X, \mathcal{F})$ of §2. Now the collection $\mathcal{C}$ of finite subsets of S is partially ordered and directed upwards by $\subseteq$ which is a concurrent binary relation on $\mathcal{C} \times \mathcal{C}$, and we conclude (since $^*\mathfrak{M}$ is an enlargement) that there exist sets $\Omega \in {}^*\mathcal{C}$ so that $A \subseteq \Omega$ for all $A \in \mathcal{C}$; such sets are called *infinite*. For technical reasons, having to do with the limiting significance of measures on $(\mathbf{X}_\Omega, \boldsymbol{\mathcal{F}}_\Omega)$ which will be taken up in the next section, it is necessary to choose Ω to be a sufficiently large infinite set, a question to which we now turn.

Let $\mathrm{cof}(\mathcal{C})$ be the collection of all cofinal subsets $\mathcal{C}' \subseteq \mathcal{C}$. To each $\mathcal{C}'$ we associate a fixed infinite element $\Lambda_{\mathcal{C}'}$ in $^*\mathcal{C}'$, which exists by the argument of the previous paragraph. The set of all such infinite elements is a subset of the infinite elements of $^*\mathcal{C}$ which we denote by $\mathcal{C}_{\lim}$. Note that $\mathrm{card}(\mathrm{cof}(\mathcal{C})) = \mathrm{card}(\mathcal{C}_{\lim})$. Given a $\Lambda \in \mathcal{C}$, a pair $\psi_\Lambda = \{\langle \mu_A^\Lambda \rangle, \langle f_A^\Lambda \rangle\}$ $(A \in I$ and $\Lambda \in \mathcal{C}_A)$, where $\mu_A^\Lambda \in M(X_A, \mathcal{F}_A)$ and $f_A^\Lambda \in F(X_A, \mathcal{F}_A)$ is called *regular* if $\lim \mu_A^\Lambda(f_A^\Lambda)$ $(A \to \infty)$ in I exists.[3] The set of all regular pairs will be denoted by $\mathcal{R}$. We let κ be a cardinal strictly greater than $\mathrm{card}(\mathcal{C} \cup \mathrm{cof}(\mathcal{C}) \cup \mathcal{R})$. Given the sets S and Y, a suitable κ can be determined once and for all. We will suppose that $^*\mathfrak{M}$ is κ-saturated.

If r^Λ is the limit of the regular pair ψ_Λ then given an $\varepsilon > 0$ in R there is an element $A(\varepsilon, \psi_\Lambda, \Lambda) \supseteq \Lambda$ in I so that $|\mu_A^\Lambda(f_A^\Lambda) - r^\Lambda| < \varepsilon$ if $A \supseteq A(\varepsilon, \psi_\Lambda, \Lambda)$. By transfer to $^*\mathfrak{M}$ we see that given $\varepsilon > 0$ in $^*R$, $\Lambda \in {}^*\mathcal{C}$, and $\boldsymbol{\psi}_\Lambda \in {}^*\mathcal{R}$ there is an element $\mathbf{A}(\varepsilon, \boldsymbol{\psi}_\Lambda, \Lambda)$ in $^*I$ so that $|\boldsymbol{\mu}_\mathbf{A}^\Lambda(\mathbf{f}_\mathbf{A}^\Lambda) - \mathbf{r}^\Lambda| < \varepsilon$ whenever $\mathbf{A} \supseteq \mathbf{A}(\varepsilon, \boldsymbol{\psi}_\Lambda, \Lambda)$, where $\mathbf{r}^\Lambda$ is the extension of r^Λ and $\boldsymbol{\psi}_\Lambda = \{\langle \boldsymbol{\mu}_A \rangle, \langle \mathbf{f}_\mathbf{A} \rangle\}$. In particular this is true for each standard $\boldsymbol{\psi}_\Lambda \in {}^*\mathcal{R}$ (i.e. of the form $\boldsymbol{\psi}_\Lambda = {}^*\psi_\Lambda$) and each $\Lambda \in \mathcal{C} \cup \mathcal{C}_{\lim}$. Let $\mathcal{R}'$ denote the standard elements in $^*\mathcal{R}$. Then $\mathrm{card}(\mathcal{R}') = \mathrm{card}(\mathcal{R})$. Now let $\varepsilon > 0$ in

[3]We write $\mu(f) = \int_X f\, d\mu$ for $\mu \in M(X, \mathcal{F})$ and $f \in F(X, \mathcal{F})$ and $\mu(F) = \mu(\chi_F)$ where χ_F is the characteristic function of $F \in \mathcal{F}$. Also $\lim r_A$ $(A \to \infty$ in $\mathcal{C}') = r$ if given $\varepsilon > 0$ in R there is an $A_0 \in \mathcal{C}'$ so that $|r_A - r| < \varepsilon$ if $A \supseteq A_0$, $A \in \mathcal{C}'$, where $\mathcal{C}'$ is cofinal in $\mathcal{C}$.

$^*\mathcal{R}$ be a *fixed* infinitesimal. For each $\psi_\Lambda \in \mathcal{R}'$ and $\Lambda \in \mathcal{C} \cup \mathcal{C}_{\lim}$ we find an element $\mathbf{A}(\psi_\Lambda, \Lambda) = \mathbf{A}(\varepsilon, \psi_\Lambda, \Lambda)$ in $^*I$ as above and call the set of all such elements $\mathcal{A}$. Then $\operatorname{card}(\mathcal{A}) < \kappa$ and the relation $\subseteq$ is concurrent on $\mathcal{A}$. Since $^*\mathcal{M}$ is κ-saturated we can find an $\Omega \in {}^*I$ so that $\mathbf{A} \subseteq \Omega$ for all $\mathbf{A} \in \mathcal{A}$. Ω will be fixed throughout the paper. From the above discussion we see that

$$\mu_\Omega^\Lambda(\mathbf{f}_\Omega^\Lambda) \approx \mathbf{r}^\Lambda \tag{3.1}$$

for each regular pair $\psi_\Lambda = \{\langle \mu_A^\Lambda \rangle, \langle f_A^\Lambda \rangle\}$ and $\Lambda \in \mathcal{C} \cup \mathcal{C}_{\lim}$. This fact will be used in the next section.

With Ω fixed as indicated we write $\mathbf{X}$ for $\mathbf{X}_\Omega$ and denote $\mathfrak{F}_\Omega$ by $\mathfrak{F}_0$ in analogy with the algebra $\mathfrak{F}_0$ of cylinder sets of the standard limiting space X. Denoting the map $\mathbf{p}_{A\Omega}$: $\mathbf{X} \to \mathbf{X}_A$, $\mathbf{A} \subseteq \mathcal{C}_\Omega$, by $\mathbf{p}_A$, we see that $\mathfrak{F}_0$ consists of all *internal cylinder* sets, i.e., sets of the form $\mathbf{p}_\mathbf{A}^{-1}(\mathbf{G})$, $\mathbf{G} \in \mathfrak{F}_\mathbf{A}$, $\mathbf{A} \subseteq \mathcal{C}_\Omega$. Similarly all of the entities introduced on $(X_A, \mathfrak{F}_A)$ in the last section (e.g. $\mathfrak{F}_A^\Lambda$, g_A^Λ, f_A^Λ, Z_A^Λ, R_A^Λ, etc.) extend to internal entities $\mathfrak{F}_\Omega^\Lambda$, $\mathbf{g}_\Omega^\Lambda$, etc. Since Ω will be fixed in the ensuing discussion, we will suppress the dependence on Ω of transferred quantities, writing $\mathfrak{F}^\Lambda$, $\mathbf{g}^\Lambda$, $\mathbf{f}^\Lambda$, $\mathbf{Z}^\Lambda$, $\boldsymbol{\pi}^\Lambda$, and $\mathbf{t}^\Lambda$ for $\mathfrak{F}_\Omega^\Lambda$, $\mathbf{g}_\Omega^\Lambda$, etc. To avoid a typesetter's nightmare we will use standard face for symbols which should be boldface. Thus in $\mathbf{Z}^\Lambda(y)$ we write y rather than $\mathbf{y}$ for an element in $X_{\Omega-\Lambda}$ (we are already using Λ rather than $\boldsymbol{\Lambda}$ for an element in $\mathcal{C}$). Similarly we will write $\mathbf{Z}^\Lambda(y) = \Sigma \exp \mathbf{g}^\Lambda(y \times \omega)$ but, strictly speaking, Σ, exp and $\times$ should be boldface since they are nonstandard transfers of operations and functions defined in the standard projective family $\langle (X_A, \mathfrak{F}_A), p_{AB} \rangle$ $(A, B \in I)$. The context will settle all possible sources of confusion.

As it stands, the internal space $(\mathbf{X}, \mathfrak{F}_0)$ is not a standard measurable space since $\mathfrak{F}_0$ is only an algebra and not a σ-algebra of subsets of X. Similarly the set $\mathbf{M}(\mathbf{X}, \mathfrak{F}_0)$ consists of $^*R$-valued (not R-valued) finitely additive (not countably additive) "measures". We follow Loeb [10] and associate with $(\mathbf{X}, \mathfrak{F}_0)$ the standard measurable space $(\mathbf{X}, \mathfrak{F}_0)$ by letting $\mathfrak{F}$ be the σ-algebra $\sigma(\mathfrak{F}_0)$ generated by $\mathfrak{F}$.[4]

For each $\boldsymbol{\mu} \in \mathbf{M}(\mathbf{X}, \mathfrak{F})$ we may define the standard part $^\circ\boldsymbol{\mu}$ of $\boldsymbol{\mu}$ on $\mathfrak{F}_0$ by

$$^\circ\boldsymbol{\mu}(\mathbf{F}) = \begin{cases} \operatorname{st} \boldsymbol{\mu}(\mathbf{F}) & \text{if } \boldsymbol{\mu}(\mathbf{F}) \text{ is finite, } \mathbf{F} \in \mathfrak{F}_0, \\ \infty & \text{otherwise.} \end{cases}$$

It is immediate that $^\circ\boldsymbol{\mu}$ is finitely additive on $\mathfrak{F}_0$. Loeb showed that $^\circ\boldsymbol{\mu}$ is, in fact, countably additive if $^*\mathcal{M}$ is $\aleph_1$-saturated (which we have assumed) and so can be extended to $\mathfrak{F}$ by the Carathéodory Extension Theorem and so we have a map

$$L\colon \mathbf{M}(\mathbf{X}, \mathfrak{F}_0) \to M(\mathbf{X}, \mathfrak{F}), \tag{3.2}$$

where $L(\boldsymbol{\mu})$ is the extension of $^\circ\boldsymbol{\mu}$. Loeb also showed that if $\mathbf{f} \in \mathbf{F}(\mathbf{X}, \mathfrak{F}_0)$ and the standard part $^\circ\mathbf{f}$ of $\mathbf{f}$ is defined on $\mathbf{X}$ by

$$^\circ\mathbf{f}(x) = \begin{cases} \operatorname{st} \mathbf{f}(x), & \mathbf{f}(x) \text{ finite,} \\ \infty, & \mathbf{f}(x) \text{ positive infinite,} \\ -\infty, & \mathbf{f}(x) \text{ negative infinite,} \end{cases}$$

then $^\circ\mathbf{f}$ is $\mathfrak{F}$-measurable, i.e. in $F(\mathbf{X}, \mathfrak{F})$.

---

[4]Here we are using boldface to denote an external entity.

We say that $\mathbf{f}$ is *S-integrable with respect to* $\boldsymbol{\mu}$ if $^\circ\mathbf{f}$ is integrable with respect to $L(\boldsymbol{\mu})$ and $^\circ[\boldsymbol{\mu}(\mathbf{f})] = L(\boldsymbol{\mu})(^\circ\mathbf{f})$. Loeb showed that $\mathbf{f}$ is S-integrable with respect to $\boldsymbol{\mu}$ if $\boldsymbol{\mu}(X)$ is finite and $\mathbf{f}$ is bounded by a finite number (and thus in particular for probability measures $\boldsymbol{\mu}$ and $\mathbf{f} = \chi_{\mathbf{F}}$ where $\chi_{\mathbf{F}}$ is the characteristic function of $\mathbf{F} \in \mathfrak{F}_0$). For general conditions on S-integrability see [1]. Thermodynamic quantities usually appear in the form $^\circ[\boldsymbol{\mu}(\mathbf{f})]$ and the question of S-integrability only arises if we want to identify this number as the integral $L(\boldsymbol{\mu})(^\circ\mathbf{f})$; usually this will be unnecessary.

It should be understood that the internal space $(\mathbf{X}, \mathfrak{F}_0)$ and the standard space $(\mathbf{X}, \mathfrak{F})$ work as a pair. The internal structure of $(\mathbf{X}, \mathfrak{F}_0)$ is very rich and can be effectively used in analysis; indeed, every entity which can be defined on $(X_A, \mathfrak{F}_A)$, $A \in I$, has an internal analogue on $(\mathbf{X}, \mathfrak{F}_0)$, but in most cases these internal entities on $(\mathbf{X}, \mathfrak{F}_0)$ cannot be transposed to $(\mathbf{X}, \mathfrak{F})$. However, as we have seen, the measure theory can be transposed. Thus $(\mathbf{X}, \mathfrak{F})$ is the repository of limiting measure theoretic information on the underlying projective family $\langle (X_A, \mathfrak{F}_A), p_{AB} \rangle$ $(A, B \in I)$ and this information is the central concern in statistical mechanics. We take $\langle (\mathbf{X}, \mathfrak{F}), \mathbf{p}_{\mathbf{A}} \rangle$ $(\mathbf{A} \in \mathbf{I})$ as our analogue of the limiting configuration structure $\langle (X, \mathfrak{F}), p_A \rangle$ $(A \in I)$ used in [13].

Now we turn to the definition of equilibrium states (Gibbs states in [13]) associated with a given interaction potential $\langle \phi_A \rangle$ $(A \in I)$.

3.1. Definition. (1) The *nonstandard equilibrium states* for $\langle \phi_A \rangle$ $(A \in I)$ are all finitely additive set functions in $\mathbf{P}(\mathbf{X}, \mathfrak{F}_0)$ of the form $\mathbf{t}^\Lambda(\boldsymbol{\nu})$ with Λ an infinite element of $\mathcal{C}_\Omega$ and $\boldsymbol{\nu} \in \mathbf{P}(\mathfrak{F}^\Lambda)$. We denote the set of all nonstandard equilibrium states by $\mathbf{G}(\phi_A)$.

(2) The *generalized equilibrium states* for $\langle \phi_A \rangle$ $(A \in I)$ are all measures in $P(\mathbf{X}, \mathfrak{F})$ of the form $L(\mathbf{t}^\Lambda(\boldsymbol{\nu}))$ for Λ and $\boldsymbol{\nu}$ as above. The set of all generalized equilibrium states will be denoted by $G(\phi_A)$.

Recall from [13] that a standard equilibrium state μ satisfies $t_\Lambda \circ r_\Lambda(\mu)(F) = \mu(F)$ for all $\Lambda \in \mathcal{C}$. In our situation we have an analogous fact; if $\boldsymbol{\mu} = \mathbf{t}^{\tilde{\Lambda}}(\boldsymbol{\nu})$, $\tilde{\Lambda}$ infinite, is a nonstandard equilibrium state, then from (2.11) we have

$$\mathbf{t}^\Lambda \circ \mathbf{r}^\Lambda(\boldsymbol{\mu}) = \boldsymbol{\mu}, \qquad \Lambda \subseteq \tilde{\Lambda}. \tag{3.3}$$

The set $\mathbf{G}(\phi_A)$ of nonstandard equilibrium states for $\langle \phi_A \rangle$ $(A \in I)$ consists of all finitely additive $^*R$-valued set functions on $(\mathbf{X}, \mathfrak{F}_0)$ of the form $\mathbf{t}^\Lambda(\boldsymbol{\nu})$ where $\Lambda \in \mathcal{C}_\Omega$ is an infinite internal subset of Ω and $\boldsymbol{\nu} \in \mathbf{P}(\mathfrak{F}^\Lambda)$. By transfer of (2.12), each has a density of the form

$$\boldsymbol{\sigma}_{\boldsymbol{\delta}}^\Lambda(x) = \boldsymbol{\delta}(x_{\Omega - \Lambda})\mathbf{f}^\Lambda(x) \tag{3.4}$$

for $x \in \mathbf{X}$, where $\boldsymbol{\delta}$: $\mathbf{X}_{\Omega - \Lambda} \to {}^*R$ is an internal function satisfying $\boldsymbol{\delta} \geqslant 0$ and $\sum \boldsymbol{\delta}(y)$ $(y \in \mathbf{X}_{\Omega - \Lambda}) = 1$. The corresponding nonstandard equilibrium state $\boldsymbol{\nu}_{\boldsymbol{\delta}}^\Lambda$ satisfies

$$\boldsymbol{\nu}_{\boldsymbol{\delta}}^\Lambda(\mathbf{F}) = \sum \boldsymbol{\sigma}_{\boldsymbol{\delta}}^\Lambda(x) \qquad (x \in \mathbf{F}) \tag{3.5}$$

for $\mathbf{F} \in \mathfrak{F}_0$. Finally, the generalized equilibrium states are all of the form $\nu_{\boldsymbol{\delta}}^\Lambda = L(\boldsymbol{\nu}_{\boldsymbol{\delta}}^\Lambda)$ for some $\nu_{\boldsymbol{\delta}}^\Lambda \in \mathbf{G}(\nu_A)$.

A particularly important case is obtained when the function δ is concentrated at

a point $y_0 \in \mathbf{X}_{\Omega-\Lambda}$ as will be seen in the next section. In that case we write $\boldsymbol{\sigma}_y^\Lambda$ and $\boldsymbol{\nu}_y^\Lambda$ for $\boldsymbol{\sigma}_\delta^\Lambda$ and $\boldsymbol{\nu}_\delta^\Lambda$ respectively if $y \in \mathbf{X}$ satisfies $y_{\Omega-A} = y_0$; obviously

$$\boldsymbol{\sigma}_y^\Lambda(x) = \boldsymbol{\pi}^\Lambda(y, \{x\}). \tag{3.6}$$

It should be remarked that the nonstandard partition functions $\mathbf{Z}^\Lambda(y)$ exist only in the nonstandard world of $(\mathbf{X}, \mathfrak{F}_0)$, and there would be no purpose whatever in taking the standard part of these functions since it would usually be infinite. However these nonstandard partition functions can be used in intermediate calculations in the nonstandard world to obtain formulas for thermodynamic quantities which do have standard meaning, a topic to which we now turn.

To begin, we present a nonstandard analogue of the information gain of two measures as defined in §7 of [**13**]. Suppose that $\mu = L(\boldsymbol{\mu})$ and $\nu = L(\boldsymbol{\nu})$ are two measures in $P(\mathbf{X}, \mathfrak{F})$, where $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ are in $\mathbf{P}(\mathbf{X}, \mathfrak{F}_\Omega)$. Then $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ are internal, *-finitely additive, $^*R$-valued (probability) measures on $(\mathbf{X}, \mathfrak{F}_0)$. For $x \in \mathbf{X}$, $\{x\}$ is in $\mathfrak{F}_0$ and so we can define densities $\boldsymbol{\rho}$: $\mathbf{X} \to {}^*R$ and $\boldsymbol{\sigma}$: $\mathbf{X} \to {}^*R$ associated with $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ by putting $\boldsymbol{\rho}(x) = \boldsymbol{\mu}(\{x\})$ and $\boldsymbol{\sigma}(x) = \boldsymbol{\nu}(\{x\})$. For any $\Lambda \in \mathcal{C}_\Omega$ we put $\boldsymbol{\mu}^\Lambda = \mathbf{p}_{\Lambda\Omega}(\boldsymbol{\mu})$, $\boldsymbol{\nu}^\Lambda = \mathbf{p}_{\Lambda\Omega}(\boldsymbol{\nu})$. These measures have densities $\boldsymbol{\rho}^\Lambda$ and $\boldsymbol{\sigma}^\Lambda$ on $\mathbf{X}_\Lambda$. Supposing that $\boldsymbol{\sigma}^\Lambda(w) \neq 0$ for $w \in \mathbf{X}_\Lambda$, we define

$$\begin{aligned} \mathbf{H}^\Lambda(\mu, \nu) &= \sum \boldsymbol{\Phi}(\mathbf{g}_\Lambda)(w_\Lambda)\boldsymbol{\sigma}(w) \qquad (w \in \mathbf{X}) \\ &= \sum \boldsymbol{\Phi}(\mathbf{g}_\Lambda)(w)\boldsymbol{\sigma}^\Lambda(w) \qquad (w \in \mathbf{X}_\Lambda), \end{aligned} \tag{3.7}$$

where

$$\Phi(f)(\cdot) = f(\cdot) \log f(\cdot) \tag{3.8}$$

and

$$\mathbf{g}_\Lambda(w) = \frac{\boldsymbol{\rho}^\Lambda(w)}{\boldsymbol{\sigma}^\Lambda(w)} \qquad (w \in \mathbf{X}_\Lambda) \tag{3.9}$$

and we put $p \log p = 0$ if $p = 0$. Notice that we also have

$$\mathbf{H}^\Lambda(\mu, \nu) = \sum \boldsymbol{\Psi}(\mathbf{g}_\Lambda)(w_\Lambda)\boldsymbol{\sigma}(w) \qquad (w \in \mathbf{X}) \tag{3.10}$$

where

$$\Psi(f) = \Phi(f) - f + 1. \tag{3.11}$$

The function $\boldsymbol{\Psi}$ satisfies $\boldsymbol{\Psi} \geqslant 0$, $\boldsymbol{\Psi}(\mathbf{f}) = 0$ only if $\mathbf{f} = 1$ and is *-strictly convex (all of these by transfer). It follows that

$$\mathbf{H}^\Lambda(\mu, \nu) \geqslant 0 \tag{3.12}$$

and as on p. 116 of [**13**] we have

$$\mathbf{H}^\Lambda(\mu, \nu) \leqslant \mathbf{H}^{\tilde{\Lambda}}(\mu, \nu), \qquad \Lambda \subseteq \tilde{\Lambda}. \tag{3.13}$$

Next we assume that there exists a fixed finite measure m defined on $\mathcal{C}$ (i.e., m: $\mathcal{C} \to R^+$ with $m(\Lambda) < \infty$ for all $\Lambda \in \mathcal{C}$) satisfying $m(\Lambda) \geqslant |\Lambda|$, the cardinality of Λ. The measure m extends to $\mathbf{m}$: $\mathcal{C} \to {}^*R^+$ and we define

$$\mathbf{h}^\Lambda(\mu, \nu) = \frac{1}{\mathbf{m}(\Lambda)}\mathbf{H}^\Lambda(\mu, \nu) \tag{3.14}$$

for any $\Lambda \in \mathcal{C}_\Omega$.

Next we introduce thermodynamic quantities. With $\boldsymbol{\mu}$ as before, we define the *nonstandard entropy* of μ (or $\boldsymbol{\mu}$) on the element $\Lambda \in \mathcal{C}_\Omega$ by

$$\mathbf{s}^\Lambda(\mu) = -\frac{1}{\mathbf{m}(\Lambda)} \sum \boldsymbol{\rho}^\Lambda(w) \log \boldsymbol{\rho}^\Lambda(w) \qquad (w \in \mathbf{X}_\Lambda). \tag{3.15}$$

Similarly we define the *nonstandard specific energy* for μ on Λ with respect to the distribution $y \in \mathbf{X}$ by

$$\mathbf{e}_y^\Lambda(\mu) = -\frac{1}{\mathbf{m}(\Lambda)} \sum \mathbf{g}^\Lambda(y_{\Omega-\Lambda} \times w)\boldsymbol{\rho}^\Lambda(w) \qquad (w \in \mathbf{X}_\Lambda). \tag{3.16}$$

($\mathbf{e}_y^\Lambda$ implicitly depends on the interaction potentials Φ_Λ, but we are not concerned with that dependence here.) The *nonstandard pressure* on Λ with respect to $y \in \mathbf{X}$ is defined by

$$\mathbf{P}_y^\Lambda = \frac{1}{\mathbf{m}(\Lambda)} \log \mathbf{Z}^\Lambda(y). \tag{3.17}$$

Finally we define the *nonstandard specific free energy* of μ on Λ with respect to y by

$$\mathbf{f}_y^\Lambda(\mu) = \mathbf{e}_y^\Lambda(\mu) - \mathbf{s}^\Lambda(\mu). \tag{3.18}$$

There is an important relationship between the information gain and the specific free energy which will be used in §5, namely if $\boldsymbol{\nu} = \boldsymbol{\nu}_y^\Lambda$ then $\boldsymbol{\sigma}^\Lambda(w) = \mathbf{f}(y_{\Omega-\Lambda} \times w_\Lambda)$ and so

$$\mathbf{h}^\Lambda(\mu, \nu) = \mathbf{f}_y^\Lambda(\mu) + \mathbf{P}_y^\Lambda. \tag{3.19}$$

Next we show that under Assumption 2.2, the above quantities are always finite.

3.2. LEMMA. *The quantity* $\mathbf{s}^\Lambda(\mu)$ *is always finite. Under Assumption* 2.2, *the quantities* $\mathbf{p}_y^\Lambda$, $\mathbf{f}_y^\Lambda(\mu)$, *and* $\mathbf{e}_y^\Lambda(\mu)$ *are also finite.*

PROOF. (i) To bound $\mathbf{s}^\Lambda(\mu)$ we proceed as follows. Since $\boldsymbol{\rho}^\Lambda(w) \leqslant 1$ we have $0 \leqslant \mathbf{s}^\Lambda(\mu)$. Now from the convexity of the function $f(x) = x \log x$ we see that $x \log x \geqslant x - 1$ on $0 \leqslant x$ which, setting $x = p/q$, yields

$$p \log p - p \log q \geqslant p - q \tag{3.20}$$

for $p \geqslant 0$ and $q > 0$. Putting $q = |\mathbf{X}_\Lambda|^{-1}$, $p = \boldsymbol{\rho}^\Lambda(w)$ and summing over $w \in \mathbf{X}_\Lambda$ we get

$$\mathbf{s}^\Lambda(\boldsymbol{\mu}) \leqslant \frac{1}{|\Lambda|}\left[-\log |X_\Lambda|^{-1}\right] = \log |Y|$$

since $\log|X_\Lambda| = |\Lambda| \log |Y|$.

(ii) Under Assumption 2.2 we next bound the pressure $\mathbf{P}_y^\Lambda$. As in the proof of Proposition 5.2 in **[13]**, we have

$$|g_A^\Lambda(x)| \leqslant \sum |\Phi_C(x_C)| \ (C \in V_A^\Lambda) \leqslant K|\Lambda| \tag{3.21}$$

and so by transfer $|\mathbf{g}^\Lambda(x)| \leqslant K|\Lambda|$ for $\Lambda \in \mathcal{C}_\Omega$. Thus

$$\mathbf{Z}^\Lambda(y) \leqslant \exp(K|\Lambda|)|\mathbf{X}_\Lambda|$$

and similarly $\mathbf{Z}^\Lambda(y) \geqslant \exp(-K|\Lambda|)|\mathbf{X}_\Lambda|$. Thus

$$-K \leqslant \mathbf{P}_y^\Lambda \leqslant K + \log |Y|.$$

(iii) Under Assumption 2.2,

$$\begin{aligned}|\mathbf{e}_y^\Lambda(\mu)| &\leqslant \frac{1}{|\Lambda|}\sum |\mathbf{g}^\Lambda(y \times w)|\rho^\Lambda(w) \qquad (w \in \mathbf{X}_\Lambda)\\ &\leqslant \frac{1}{|\Lambda|}K|\Lambda| = K.\end{aligned}$$

This completes the proof. □

3.3. Definition. We define the *generalized entropy* $s^\Lambda(\mu)$, *generalized specific energy* $e_y^\Lambda(\mu)$, *generalized pressure* P_y^Λ, and *generalized specific free energy* $f_y^\Lambda(\mu)$ (on Λ with respect to $y \in \mathbf{X}$) of a measure $\mu = L(\boldsymbol{\mu})$ by the equations $s^\Lambda(\mu) = {}^\circ\mathbf{s}^\Lambda(\mu)$, $e_y^\Lambda(\mu) = {}^\circ\mathbf{e}_y(\mu)$, $P_y^\Lambda = {}^\circ\mathbf{P}_y^\Lambda$ and $f_y^\Lambda(\mu) = {}^\circ\mathbf{f}_y^\Lambda(\mu)$ respectively.

The generalized quantities of Definition 3.3 are all ordinary real numbers by Lemma 3.2.

4. Limiting significance of measures on $(\mathbf{X}, \mathfrak{F})$. In the first part of this section we indicate the limiting significance of a certain class $M_s(\mathbf{X}, \mathfrak{F})$ of measures in $L(\mathbf{M}(\mathbf{X}, \mathfrak{F}_0))$ for the underlying standard projective $\langle (X_A, \mathfrak{F}_A), p_{AB}\rangle$ $(A, B \in I)$, by showing that (a) and (b) of the Introduction hold for measures in $M_s(\mathbf{X}, \mathfrak{F})$. Secondly, we investigate connections between the measure theory (and equilibrium states in particular) on $(\mathbf{X}, \mathfrak{F})$ and that on the standard limiting configuration space $(X, \mathfrak{F})$. Thirdly, we show that there is a close connection between the generalized thermodynamic quantities introduced in §3 and the standard ones introduced in **[13]**.

The class $M_s(\mathbf{X}, \mathfrak{F})$ consists of all measures of the form $L(\boldsymbol{\mu}_\Omega^{\Lambda\mathcal{C}})$ where $\langle \mu_B^\Lambda\rangle$ $(B \in I, A \subseteq B$ and $A \in \mathcal{C}')$ and $\mathcal{C}'$ is cofinal in $\mathcal{C}$. Property (a) of the Introduction is established in the following theorem.

4.1. Theorem. *Let* $\{\langle \mu_B^A\rangle, \langle f_B^A\rangle\}$ $(B \in I,\ A \in \mathcal{C}',\ A \subseteq B)$, *with* $\mu_B^A \in M(X_B, \mathfrak{F}_B)$ *and* $f_B^A \in F(X_B, \mathfrak{F}_B)$, *be given, where* $\mathcal{C}'$ *is cofinal in* $\mathcal{C}$. *Suppose that* $\lim \mu_B^A(f_B^A)$ $(B \to \infty$ *in* $I) = r^A$ *exists for each* $A \in \mathcal{C}'$ *and* $\lim r^A$ $(A \to \infty$ *in* $\mathcal{C}')$ $= r$. *Then* $r = \mathrm{st}[\boldsymbol{\mu}_\Omega^{\Lambda\mathcal{C}}(\mathbf{f}_\Omega^{\Lambda\mathcal{C}})]$.

Proof. Let $\mathbf{r}^A$ $(A \in \mathcal{C}')$ be the extension of r^A $(A \in \mathcal{C}')$. By (3.1), $\mathbf{r}^{\Lambda\mathcal{C}} = \boldsymbol{\mu}_\Omega^{\Lambda\mathcal{C}}(\mathbf{f}_\Omega^{\Lambda\mathcal{C}})$ and $\mathbf{r}^{\Lambda\mathcal{C}} \approx r$ since $r = \lim r^A$ $(A \to \infty$ in $\mathcal{C}')$. □

Note that if $\mathbf{f}_\Omega^{\Lambda\mathcal{C}}$ is S-integrable with respect to $\boldsymbol{\mu}_\Omega^{\Lambda\mathcal{C}}$ (and in particular for finite $\boldsymbol{\mu}_\Omega^{\Lambda\mathcal{C}}(\mathbf{X})$ and finitely bounded $f_\Omega^{\Lambda\mathcal{C}}$) then we can replace $\mathrm{st}[\boldsymbol{\mu}_\Omega^{\Lambda\mathcal{C}}(\mathbf{f}_\Omega^{\Lambda\mathcal{C}})]$ by $\mu(f)$ where $\mu = L(\boldsymbol{\mu}_\Omega^{\Lambda\mathcal{C}})$ and $f = {}^\circ\mathbf{f}_\Omega^{\Lambda\mathcal{C}}$. Similar considerations apply throughout this section and will not be explicitly stated.

Theorem 4.1 will be central to the results in this section. In spite of its simplicity, it makes apparent the considerable technical advantages which result from working on $(\mathbf{X}, \mathfrak{F})$ rather than the standard limiting space $(X, \mathfrak{F})$. To see this we review the question of establishing the analogue of a special case of the theorem on $(X, \mathfrak{F})$. Suppose that $\mu_B^A = \mu_B$ is independent of A and that $\lim \mu_B(f_B^A)$ $(B \to \infty$ in $\mathcal{C})$ exists for all $A \in \mathcal{C}$, $A \subseteq B$ and all f_B^A of the form $f_B^A = \chi p_{AB}^{-1}(F)$ with $F \in \mathfrak{F}_A$. The problem is to find a measure $\mu \in M(X, \mathfrak{F})$ so that $\mu(p_A^{-1}(F)) = \lim \mu_B(f_B^A)$ $(B \to \infty$ in $\mathcal{C})$ for all A and $F \in \mathfrak{F}_A$. A result of this sort is established by the

Kolmogoroff theorem under the *additional* assumptions that the measures $\langle \mu_B \rangle$ $(B \in \mathcal{C})$ are probability measures and form a *consistent* family, i.e., that $\mu_A(F) = \mu_B(p_{AB}^{-1}(F))$ for $B \supseteq A$ and $F \in \mathcal{F}_A$ (in which case $\mu_B(f_B^A)$ is independent of B and $\lim \mu_B(f_B^A)$ $(B \to \infty$ in $\mathcal{C})$ automatically exists). In the more general situation considered by Lenard in [9], in which $\langle (X_A, \mathcal{F}_A), p_{AB} \rangle$ $(A, B \in I)$ is replaced by a more general family $\langle (X_i, \mathcal{F}_i), p_{ij} \rangle$ $(i, j \in I)$ where $(X_i, \mathcal{F}_i)$ $(i \in I)$ is a measurable space and I is partially ordered and directed upwards, the analogue of the Kolmogoroff result does not even hold under the assumption of consistency unless further assumptions of a topological nature are made on the spaces $(X_i, \mathcal{F}_i)$ and the measures, leading to other technical complications. However, our result can be immediately generalized to that situation.

An even more serious difficulty is that there is no corresponding result without the assumption of consistency of $\langle \mu_A \rangle$ $(A \in \mathcal{C})$, and it is precisely this situation that arises in the case of interacting particle systems, and, in particular, in proving the existence of equilibrium states. In the general situation we have considered (in which μ_B^A can depend on A) the problem is usually overcome by technical artifices which eventually reduce the problem to the Kolmogoroff theorem, but this involves technical complications which are evident in the proof of the existence of equilibrium states in [13] (see also [7]). In our development the existence of equilibrium states is immediate, and in addition, we have a simple representation for each equilibrium state.

As an immediate corollary of Theorem 4.1 we have the following connection with measures on $(X, \mathcal{F})$.

4.2. Corollary. *If* $\mu^A \in M(X, \mathcal{F})$ *and* $f^A \in F(X, \mathcal{F})$ $(A \in \mathcal{C}')$ *and* $\lim \mu_B^A(f_B^A)$ $(B \to \infty$ *in* $I) = \mu^A(f^A)$ *for all* $A \in \mathcal{C}'$ *and* $\lim \mu^A(f^A)$ $(A \to \infty$ *in* $\mathcal{C}')$ *exists then*

$$\mathrm{st}\left[\boldsymbol{\mu}_\Omega^{\Lambda_{\mathcal{C}}}(\mathbf{f}_\Omega^{\Lambda_{\mathcal{C}}})\right] = \lim \mu^A(f^A) \qquad (A \to \infty \text{ in } \mathcal{C}').$$

Systematic use of the following result yields property (b) of the Introduction.

4.3. Theorem. *With* μ_B^A *and* f_B^A *as in Theorem* 4.1, *we have* $\mathrm{st}[\boldsymbol{\mu}_\Omega^{\mathcal{C}}(\mathbf{f}_\Omega^{\mathcal{C}})] = \lim \mu_B^A(f_B^A)$, $A \to \infty$, $B \to \infty$, $(A \in \mathcal{C}''$, $B \in I')$ *where* I' *is cofinal in* I *and* $\mathcal{C}''$ *is cofinal in* $\mathcal{C}'$.

Proof. Let $\varepsilon > 0$ in R be given, and let $A_0 \in \mathcal{C}'$ and $B_0 \in I$ be fixed. With $r = \mathrm{st}[\boldsymbol{\mu}_\Omega^{\Lambda_{\mathcal{C}}}(\mathbf{f}_\Omega^{\Lambda_{\mathcal{C}}})]$, the following internal statement is true in ${}^*\mathfrak{M}$: "there exists an $A \supseteq A_0$ in $\mathcal{C}'$ and a $B \supseteq B_0$ in I so that $|r - \mu_B^A(f_B^A)| < \varepsilon$" (take $A = \Lambda_{\mathcal{C}}$ and $B = \Omega$). By transfer down to $\mathfrak{M}$ we see that there exists $A \supseteq A_0$ in $\mathcal{C}'$ and $B \supseteq B_0$ in I so that $|r - \mu_B^A(f_B^A)| < \varepsilon$ which establishes the result by a standard argument. □

Next we show how to transfer measures from $(X, \mathcal{F})$ to $(\mathbf{X}, \boldsymbol{\mathcal{F}})$ and vice versa. The maps

$$\Psi\colon M(X, \mathcal{F}) \to M(\mathbf{X}, \boldsymbol{\mathcal{F}}) \quad \text{and} \quad \Phi\colon M^f(\mathbf{X}, \boldsymbol{\mathcal{F}}) \to M^f(X, \mathcal{F})$$

(where the superscript f denotes the set of finite measures) are defined as follows: Let $\tilde{\mu} \in M(X, \mathcal{F})$; then the *projection* $p_A(\tilde{\mu}) = \mu_A$ defined by $\mu_A(F) = \mu(p_A^{-1}(F))$,

$F \in \mathcal{F}_A$, yields a consistent family $\langle \mu_A \rangle$ $(A \in I)$ and we let $\Psi(\tilde{\mu}) = L(\boldsymbol{\mu}_\Omega)$. Similarly, let $\mu \in M^f(\mathbf{X}, \mathfrak{F})$. The map $\mathbf{p}_A\colon \mathbf{X} \to X_A$, $A \in I$, is measurable, and the probability measures $(1/\mu(\mathbf{X}))\mathbf{p}_A(\mu) = \nu_A$ yields a consistent family. By the Kolmogoroff theorem there is a measure $\tilde{\nu} \in M(X, \mathcal{F})$ so that $\nu_A = \mathbf{p}_A(\tilde{\nu})$ and we let $\Phi(\mu) = \mu(\mathbf{X})\tilde{\nu}$.

The following lemma shows that $\tilde{\mu}$ and $\Psi(\tilde{\mu})$ and μ and $\Phi(\mu)$ agree on "standard" cylinder sets. To make this precise, we define the collection $\mathfrak{F}_s$ of standard cylinder sets in $\mathfrak{F}_0$ by

$$\mathfrak{F}_s = \{\mathbf{F} \in \mathfrak{F}_0 \mid \text{there exists } G \in \mathcal{F}_A, A \in I, \text{ and } \mathbf{F} = \mathbf{p}_A^{-1}(G)\}.$$

The map

$$\Xi\colon \mathcal{F}_0 \to \mathfrak{F}_s$$

defined by $\Xi(F) = \mathbf{p}_A^{-1}(G)$ if $F = p_A^{-1}(G)$ is easily seen to be a Boolean isomorphism.

4.4. Lemma. *The equations $\tilde{\mu}(F) = \Psi(\tilde{\mu})(\Xi(F))$ and $\mu(\Xi(F)) = \Phi(\mu)(F)$ hold for all $F \in \mathcal{F}_0$, $\tilde{\mu} \in M(X, \mathcal{F})$ and $\mu \in M^f(\mathbf{X}, \mathfrak{F})$.*

Proof. Let $F = p_A^{-1}(G)$, $G \in \mathcal{F}_A$, $A \in I$. Since $p_B(\tilde{\mu}) = \mu_B$, $B \in I$ is a consistent family, we have $\mu_B(p_{AB}^{-1}(G)) = \tilde{\mu}(F)$ for all $B \in I$. By transfer, $\boldsymbol{\mu}_\Omega(\mathbf{p}_A^{-1}(G)) = \tilde{\mu}(F)$ and hence $\tilde{\mu}(F) = \Psi(\tilde{\mu})(\Xi(F))$. The proof of the second identity is similar. □

We now use the mappings just introduced to establish a connection between standard and generalized equilibrium states. First, two results, the first of which will be used in §5.

4.5. Lemma. *Let $A \in \mathcal{C}$ be fixed. Under Assumption* 2.1, *given $\varepsilon > 0$ in R there is a $\Delta \in \mathcal{C}$ so that, if $A \subseteq \Delta \subseteq B \subseteq B'$ with $B, B' \in \mathcal{C}$, we have*

$$\left|\pi_{B'}^A\left(x_{B'}, p_{CB'}^{-1}(G)\right) - \pi_B^A\left(x_B, p_{CB}^{-1}(G)\right)\right| < \varepsilon$$

for any $x \in X$ and any $G \in \mathcal{F}_C$ where $C \subseteq \Delta$.

Proof. Find $\eta > 0$ so that $e^{2\eta} - 1 < \varepsilon$. As in the proof of Proposition 5.2 in [**13**] we see from Assumption 2.1 that $\sum \phi_D(x_D)$ $(D \cap A \neq \varnothing)$ converges uniformly and absolutely. Hence there is a $\Delta \in \mathcal{C}$ so that $|g_{B'}^A(x_{B'}) - g_B^A(x_B)| \leqslant \sum|\phi_D(x_D)|$ $(D \cap A \neq \varnothing,\ D \cap (B' - B) \neq \varnothing) < \varepsilon$ for any $x \in X$ and any pair B, B' satisfying $\Delta \subseteq B \subseteq B'$. For a specific $x \in X$, B and B', let $g_{B'}^A(x_{B'}) - g_B^A(x_B) = \psi(x_{B'})$, and put $F = p_{CB}^{-1}(G)$, $F' = p_{CB'}^{-1}(G)$ so that $F' = p_{BB'}^{-1}(F)$. Then

$$\left|\pi_{B'}^A(x_{B'}, F') - \pi_B^A(x_B, F)\right| \leqslant \sum f_{B'}^A(x_{B'-A} \times w) - f_B^A(x_{B-A} \times w) \qquad (w \in W)$$

where $W = W_{B'}^A(x_{B'}, F') = W_B^A(x_B, F)$ since $C \subseteq B \subseteq B'$. Continuing, we see that

the left-hand side is

$$\leqslant \sum \left| \frac{\exp g_{B'}^{A}(x_{B'-A} \times w)}{Z_{B'}^{A}(x_{B'})} - \frac{\exp g_{B}^{A}(x_{B-A} \times w)}{Z_{B}^{A}(x_{B})} \right| \qquad (w \in W)$$

$$\leqslant \sum \left| \frac{e^{\eta} \exp g_{B}^{A}(x_{B-A} \times w)}{e^{-\eta} Z_{B}^{A}(x_{B})} - \frac{\exp g_{B}^{A}(x_{B-A} \times w)}{Z_{B}^{A}(x_{B})} \right| \qquad (w \in W)$$

$$\leqslant e^{2\eta} - 1$$

and we are through. □

4.6. Definition. Let $\tilde{\mu} \in M(X, \mathfrak{F})$. We say that a family $\langle \mu_B \rangle$ $(B \in \mathcal{C})$ of measures $\mu_B \in M(X_B, \mathfrak{F}_B)$ approximates $\tilde{\mu}$, if, for each $C \in \mathcal{C}$, $\mu_B(p_{CB}^{-1}(G))$ converges to $\tilde{\mu}(p_C^{-1}(G))$ as $B \to \infty$ for each $G \in \mathfrak{F}_C$.

4.7. Lemma. *If $\nu \in P(\mathfrak{F}^A)$ and $\nu_B = p_B(\nu) \in P(\mathfrak{F}_B^A)$ for fixed $A \in \mathcal{C}$, then $\langle t_B^A(\nu_B) \rangle$ $(B \in \mathcal{C})$ approximates $t_A(\nu)$ as defined in* [13].

Proof. Let $C \in \mathcal{C}$, $G \in \mathfrak{F}_C$, and $\varepsilon > 0$ in R be given. Using Lemma 4.5 we find a Δ so that $C \subseteq \Delta$ and for all B and B' satisfying $\Delta \subseteq B \subseteq B'$,

$$\left| \pi_{B'}^{A}(x_{B'}, F') - \pi_{B}^{A}(x_B, F) \right| < \varepsilon$$

where $F = p_{CB}^{-1}(G)$, $F' = p_{CB'}^{-1}(G)$, and we see immediately that $|t_{B'}^{A}(\nu_{B'})(F') - t_B^A(\nu_B)(F)| < \varepsilon$. Thus $t_B^A(\nu_B)(p_{CB}^{-1}(G))$ converges as $B \to \infty$. Now by Proposition 5.3 in [13], $g_B^A(x_B)$ converges to g^A (uniformly in x), and by an argument similar to that used in Lemma 4.5 we see that $\pi_B^A(x_B, F)$ converges to $\pi_A(x, \hat{F})$ (uniformly in x) where $\hat{F} = p_C^{-1}(G)$, and hence $t_B^A(\nu_B)(p_{CB}^{-1}(G))$ converges to $t_A(\nu)(p_C^{-1}(G))$. □

4.8. Theorem. *Suppose that Assumption* 2.1 *is in force.*

(a) *If $\tilde{\mu}$ is a standard equilibrium state then there is a generalized equilibrium state μ so that $\Psi(\tilde{\mu})$ and μ agree on $\mathfrak{F}_s$.*

(b) *If $\mu = L(\mathbf{t}^{\Lambda}(\boldsymbol{\nu}^{\Lambda}))$, where $\Lambda \in \mathcal{C}_{\Omega}$ and $\Lambda \neq \Omega$ and $\boldsymbol{\nu}^{\Lambda} \in \mathbf{P}(\mathfrak{F}^{\Lambda})$, is a generalized equilibrium state then $\Phi(\mu)$ is a standard equilibrium state.*

Proof. (a) By Proposition 2.2 in [13], $\tilde{\mu}$ is characterized by the fact that $\tilde{\mu} = t_A(\nu^A)$ for each $A \in \mathcal{C}$ where ν^A is the restriction of $\tilde{\mu}$ to $\mathfrak{F}^A$. Now $\nu_B^A = p_B(\nu^A) \in P(X_B, \mathfrak{F}_B^A)$ and we define $\mu = L(\mathbf{t}^{\Lambda}(\boldsymbol{\nu}))$ where $\boldsymbol{\nu} = \boldsymbol{\nu}_{\Omega}^{\Lambda}$ and $\Lambda \in \mathcal{C}_{\lim}$. By Lemma 4.7 and Theorem 4.1 we see that

$$\Psi(\tilde{\mu})(\Xi(F)) = \tilde{\mu}(F) = \lim t_A(\nu^A)(F)\,(A \to \infty \text{ in } \mathcal{C}) = \mu(\Xi(F))$$

for all $F = p_A^{-1}(G)$, $G \in \mathfrak{F}_A$.

(b) Let $\boldsymbol{\mu} = \mathbf{t}^{\Lambda}(\nu)$, $\Lambda \in \mathcal{C}_{\Omega}$ and infinite, $\boldsymbol{\nu} \in \mathbf{P}(\mathfrak{F}_{\Omega}^{\Lambda})$, and put $\mu = L(\boldsymbol{\mu})$ and $\Phi(\mu) = \tilde{\mu}$. We need to show that $\tilde{\mu}(F) = t_A(\nu^A)(F)$ for each $A \in \mathcal{C}$, $F \in p_A^{-1}(G)$, $G \in \mathfrak{F}_A$, where ν^A is the restriction to $\mathfrak{F}^A$ of $\tilde{\mu}$. By (3.3), $\boldsymbol{\mu}(\Xi(F)) = \mathbf{t}^A(\boldsymbol{\nu}^A)(\Xi(F))$ where $\boldsymbol{\nu}^A$ is the restriction of $\boldsymbol{\mu}$ to $\mathfrak{F}^A$. Thus if $B \supseteq A$,

$$\begin{aligned} \tilde{\mu}(F) &= \Phi\big[\mathbf{t}^A(\boldsymbol{\nu}^A)\big](F) = \mathbf{p}_B\big[\mathbf{t}^A(\boldsymbol{\nu}^A)\big]\big(p_{AB}^{-1}(G)\big) \\ &= \lim t_B^A\big(\nu_B^A\big)\big(p_{AB}^{-1}(G)\big)\,(B \to \infty \text{ in } \mathcal{C}) = t_A(\nu^A)\big(p_A^{-1}(G)\big), \end{aligned}$$

the last two equalities by Lemma 4.7. □

This result shows the close connection between equilibrium states in the two models. It also yields a simple proof of the existence of standard equilibrium states (compare [**13**, §3]).

Finally we establish the connection between the nonstandard information gain of §3 and the standard information gain $h(\mu, \nu)$ as defined in §7 of [**13**]. Similar results can be established for the other thermodynamic quantities and will not be explicitly considered here.

It is relatively easy, using the results of §4, to establish a connection between $\mathbf{h}^{\Lambda}(\mu, \nu)$ for infinite Λ and the standard *information gain* $h(\mu, \nu)$ as defined in §7 of [**13**]. For $\tilde{\mu}, \tilde{\nu} \in P(X, \mathfrak{F})$, the function $h(\tilde{\mu}, \tilde{\nu})$ is the limit as $\Lambda \to \infty$ (if it exists) of

$$h_{\Lambda}(\tilde{\mu}, \tilde{\nu}) = \frac{1}{m(\Lambda)} H_{\Lambda}(\mu, \nu), \qquad \Lambda \in \mathcal{C}, \tag{4.1}$$

where

$$H_{\Lambda}(\tilde{\mu}, \tilde{\nu}) = \int \Phi(g_{\Lambda})\, d\nu, \tag{4.2}$$

$g_{\Lambda} = d\tilde{\mu}_{\Lambda}/d\tilde{\nu}_{\Lambda}$ (the Radon-Nikodým derivative) and $\tilde{\mu}_{\Lambda}$ and $\tilde{\nu}_{\Lambda}$ are the restrictions of $\tilde{\mu}$ and $\tilde{\nu}$ to $\mathfrak{F}_{(\Lambda)}$. Preston remarks that it is too much to expect the limit to exist if Λ ranges over all elements in $\mathcal{C}$, and restricts the limit to be taken over all cubes in the case $S = Z^d$. Thus we are really concerned with the limit along certain cofinal subsets $\mathcal{C}' \subseteq \mathcal{C}$.

4.9. THEOREM. *Let $\mathcal{C}'$ be a cofinal subset of $\mathcal{C}$. Suppose we are given collections $\langle \mu_B^A \rangle$ $(B \in I, A \in \mathcal{C}_A)$ and $\langle \nu_B^A \rangle$ $(B \in I, A \in \mathcal{C}_A)$ of measures in $P(X_B, \mathfrak{F}_B)$ such that for fixed A, $\langle \mu_B^A \rangle$ $(B \in I)$ and $\langle \nu_B^A \rangle$ $(B \in I)$ approximate the restrictions $\tilde{\mu}_A$ and $\tilde{\nu}_A$ of two measures $\tilde{\mu}$ and $\tilde{\nu}$ in $P(X, \mathfrak{F})$ to $\mathfrak{F}_{(A)}$. Then if $h(\tilde{\mu}, \tilde{\nu}) = \lim h_A(\tilde{\mu}, \tilde{\nu})$ $(A \to \infty$ in $\mathcal{C}')$, we have*

$$h(\tilde{\mu}, \tilde{\nu}) = {}^{\circ}\mathbf{h}^{\Lambda_{\mathcal{C}}}(\mu, \nu) \tag{4.3}$$

where $\mu = L(\boldsymbol{\mu})$, $\nu = L(\boldsymbol{\nu})$ and $\boldsymbol{\mu} = \boldsymbol{\mu}_{\Omega}^{\Lambda_{\mathcal{C}}}$, $\boldsymbol{\nu} = \boldsymbol{\nu}_{\Omega}^{\Lambda_{\mathcal{C}}}$. Conversely if $\mathbf{h}^{\Lambda_{\mathcal{C}}}(\mu, \nu)$ is finite then

$${}^{\circ}\mathbf{h}^{\Lambda_{\mathcal{C}}}(\mu, \nu) = \lim h_A(\tilde{\mu}, \tilde{\nu}) \qquad (A \to \infty \text{ in } \mathcal{C}'')$$

for some cofinal $\mathcal{C}'' \subseteq \mathcal{C}'$.

PROOF. For fixed $A \in \mathcal{C}'$, a version of the Radon-Nikodým derivative g_A is

$$g_A(x) = \tilde{\mu}_A(p_A^{-1}(\{x_A\}))/\tilde{\nu}_A(p_A^{-1}(\{x_A\})).$$

Thus

$$h_A(\tilde{\mu}, \tilde{\nu}) = \frac{1}{m(A)} \sum \Phi\left[\, \tilde{\mu}_A(p_A^{-1}(\{w\}))/\tilde{\nu}_A(p_A^{-1}(\{w\}))\right](w)\tilde{\nu}_A(p^{-1}(\{w\})) \qquad (w \in X_A).$$

But also, if

$$h_B^A(\mu_B^A, \nu_B^A) = \frac{1}{m(A)} \sum \Phi(g_B^A)(w)\sigma_B^A(w) \qquad (w \in X_A)$$

where $g_B^A(w) = \rho_B^A(w_A)/\sigma_B^A(w_A)$ and ρ_B^A and σ_B^A are the densities of μ_B^A and ν_B^A on X_A, then

$$h_B^A(\mu_B^A, \nu_B^A) = \frac{1}{m(A)} \sum \Phi\left[\mu_B^A(p_B^{-1}(\{w\}))/\nu_B^A(p_{AB}^{-1}(\{w\}))\right]\nu_B^A(p_{AB}^{-1}(\{w\})) \qquad (w \in X_A).$$

Since from Definition 4.6, $\mu_B^A(p_{AB}^{-1}(\{w\}))$ and $\nu_B^A(p_{AB}^{-1}(\{w\}))$ approach $\tilde{\mu}_A(p_A^{-1}(\{w\}))$ and $\tilde{\nu}_A(p_A^{-1}(\{w\}))$ for each $w \in X_A$, we see that $h_B^A(\mu_B^A, \nu_B^A)$ approaches $h_A(\mu, \nu)$ for fixed $A \in \mathcal{C}'$. The result now follows from Corollary 4.2 and Theorem 4.3 □

Note that the assumptions of Theorem 4.9 are satisfied if $\mu_B^A = p_A(\tilde{\mu})$ and $\nu_B^A = p_A(\tilde{\nu})$ (by projectivitity of the family $\langle \mu_A \rangle$ and $\langle \nu_A \rangle$), or if $\tilde{\nu} = t^A(\zeta)$, $\zeta \in P(X, \mathcal{F}^A)$, is a standard equilibrium measure and $\nu_B^A = t_B^A(\zeta_B)$, $\zeta_B = p(\zeta_B^A)$ (by Lemma 4.7).

Notice also that the finiteness of $\underset{\sim}{h}^{\Lambda_{\mathcal{C}}}(\mu, \nu)$ guarantees the existence of $\lim h_\Lambda(\tilde{\mu}, \tilde{\nu})$ along a certain cofinal $\mathcal{C}'' \subset \tilde{\mathcal{C}}'$ but does not establish the uniqueness of the limit. In §8 of [**13**] this question is addressed for $S = Z^d$ and any equilibrium measure ν. A nonstandard approach to this topic could be developed (for related work in continuous statistical mechanics see [**12**]) but we will not pursue it further.

5. A variational principle. One of the more important results in the standard theory of equilibrium states is the variational principle for lattice statistical mechanics first presented by Lanford and Ruelle [**8**], and subsequently extended by other authors (we refer to §§7 and 8 of [**13**] for the standard development). It provides a partial physical justification for the abstract definition of equilibrium states. In this section we will establish a variational principle of the same sort for our generalized equilibrium states, thus showing that they are also intimately related to the physics of lattice statistical mechanics. The discussion will also provide a good example of the use of the nonstandard ideas presented so far.

To begin we need a basic inequality for the information gain of §3. Let Λ, $\Delta \in \mathcal{C}_\Omega$ with $\Lambda \cap \Delta = \varnothing$ and $\Lambda \cup \Delta = D$. Then we have

$$\begin{aligned} \mathbf{H}^{\Lambda\cup\Delta}(\mu, \nu) - \mathbf{H}^{\Lambda}(\mu, \nu) &= \sum \mathbf{\Phi}(\mathbf{q}^{\Lambda,\Delta})(w)\mathbf{g}^{\Delta}(w_\Delta)\boldsymbol{\sigma}(w) \qquad (w \in \mathbf{X}) \\ &= \sum \mathbf{\Psi}(\mathbf{q}^{\Lambda,\Delta})(w)\mathbf{g}^{\Delta}(w_\Delta)\boldsymbol{\sigma}(w) \qquad (w \in \mathbf{X}) \end{aligned} \tag{5.1}$$

where

$$\mathbf{q}^{\Lambda,\Delta}(w) = \begin{cases} \mathbf{g}^{D}(w_D)/\mathbf{g}^{\Delta}(w_\Delta) & \text{if } \mathbf{g}^{\Delta}(w_\Delta) \neq 0 \ (w \in \mathbf{X}), \\ 0 & \text{otherwise.} \end{cases} \tag{5.2}$$

If $\Lambda_1, \ldots, \Lambda_n$ is a *-finite collection of disjoint sets in $\mathcal{C}_\Omega$ with $\Gamma_k = \bigcup_{j=1}^{k} \Lambda_j$, and $\Gamma_n \subset \Lambda'$, we see from (3.13) and (5.1) that

$$\mathbf{H}^{\Lambda'}(\mu, \nu) \geqslant \mathbf{H}^{\Lambda_1}(\mu, \nu) + \sum_{k=2}^{n} \left[\sum \mathbf{\Psi}(\mathbf{q}^{\Lambda_k,\Gamma_{k-1}})(w)\mathbf{g}^{\Gamma_{k-1}}(w_{\Gamma_{k-1}})\boldsymbol{\sigma}(w) \ (w \in \mathbf{X}) \right]. \tag{5.3}$$

These results are simple combinatorial computations, and are the analogues of Lemmas 7.4 and 7.5 in [**13**].

We next suppose that there is a family $H = \{\xi\}$ of bijections ξ: $\mathcal{C} \to \mathcal{C}$ satisfying the conditions

$$A \cap B = \varnothing \text{ implies } \xi(A) \cap \xi(B) = \varnothing \tag{5.4}$$

$$A \subseteq B \text{ implies } \xi(A) \subseteq \xi(B) \quad \text{and} \quad \xi(B - A) = \xi(B) - \xi(A). \tag{5.5}$$

For example, H could be induced from a group of bijections from S to S as on p. 133 of [**13**]. We make the following assumption.

5.1. Assumption. The function m: $\mathcal{C} \to R$ of §3 is invariant under H, i.e. $m(\xi(\Lambda)) = m(\Lambda)$ for all $\xi \in H$, $\Lambda \in \mathcal{C}$.

Associated with H we define the *packing function* n_H: $\mathcal{C} \times \mathcal{C} \to N$ (the natural numbers) by

$$n_H(\Lambda, \tilde{\Lambda}) = \max\{n\text{: there exist } \xi_1, \ldots, \xi_n \in H \text{ so that } \xi_1(\Lambda), \ldots, \xi_n(\Lambda) \text{ are disjoint subsets of } \tilde{\Lambda}\}. \tag{5.6}$$

The set H extends in ${}^*\mathfrak{M}$ to $\mathbf{H}$. Similarly, the function n_H extends to $\mathbf{n_H}$: $\mathcal{C} \times \mathcal{C} \to \mathbf{N}$. We make the further

5.2. Definition. Let $\mathcal{C}'$ be a cofinal subset of $\mathcal{C}$. We say that $\mathcal{C}'$ is *adapted* to H if there exists a strictly positive function α: $\mathcal{C} \to R^+$ so that for each $\Lambda \in \mathcal{C}$ and infinite $\tilde{\Lambda} \in \mathcal{C}'$ we have

$$\mathbf{m}(\Lambda)\mathbf{n_H}(\Lambda, \tilde{\Lambda}) \geqslant \alpha(\Lambda)\mathbf{m}(\tilde{\Lambda}). \tag{5.7}$$

The assumption of adaptedness is slightly weaker than the assumption on p. 134 of [**13**], which in this context immediately yields it by transfer.

A standard ingredient in obtaining a variational principle is the assumption of some sort of translational invariance of the measures involved. It turns out that something slightly weaker will suffice (this is also clear in the standard development).

5.3. Definition. We say that the measures μ and $\nu \in P(\mathbf{X}, \mathfrak{F})$ are mutually near H-invariant if for any $\xi \in \mathbf{H}$ and Λ, $\Delta \in \mathcal{C}_\Omega$ with $\Lambda \cap \Delta = \varnothing$ for which $\xi(\Lambda)$ and $\xi(\Delta) \in \mathcal{C}_\Omega$,

$$\sum \mathbf{\Psi}(\mathbf{q}^{\xi(\Lambda),\xi(\Delta)})(w)\big(\mathbf{g}^{\xi(\Delta)}(w_{\xi(\Delta)})\big)\boldsymbol{\sigma}(w)\ (w \in \mathbf{X}) \simeq 0$$

whenever

$$\sum \mathbf{\Psi}(\mathbf{q}^{\Lambda,\Delta})(w)\mathbf{g}^{\Delta}(w_\Delta)\boldsymbol{\sigma}(w)\ (w \in \mathbf{X}) \simeq 0.$$

In particular, if $\mu = L(\boldsymbol{\mu}_\Omega)$ and $\nu = L(\boldsymbol{\nu}_\Omega)$ for families of measures $\langle \mu_A \rangle$ $(A \in I)$ and $\langle \nu_A \rangle$ $(A \in I)$ on $(X_A, \mathfrak{F}_A)$, then this assumption is satisfied if each μ_A and ν_A is H-invariant.

For the proofs we need the following lemmas. The first is the finite analogue of Lemma 7.3 in [**13**] and is proved by elementary calculus. The third is the analogue of Lemma 7.9 in [**13**].

5.4. Lemma. *Given $\varepsilon > 0$ there is a $\delta > 0$ so that if $\lambda(w)$ $(w \in Q)$ is a discrete probability measure on the finite set Q, and f: $Q \to R^+$ is such that $\sum \Psi(f)(w)\lambda(w)$ $(w \in Q) < \delta$ then $\sum |f(w) - 1|\lambda(w)$ $(w \in Q) < \varepsilon$.*

5.5. LEMMA. *Let* $\mu = \mathbf{t}_\Omega^{\Lambda'}(\nu')$ *be a nonstandard equilibrium state with density* ρ. *If* $\Lambda \subseteq \Lambda'$ *is in* $\mathcal{C}_\Omega$ *and* $\mathbf{F} \in \mathfrak{F}_{(\Lambda)}$, $G \in \mathfrak{F}_{(\Omega-\Lambda)}$ *where* $\mathfrak{F}_{(A)} = \mathbf{p}_{A\Omega}^{-1}(\mathfrak{F}_A)$ *then*

$$\mu(\mathbf{F} \cap \mathbf{G}) = \sum \pi_\Omega(w, \mathbf{F})\rho(w) \qquad (w \in \mathbf{G}).$$

PROOF. Using (2.11) we see that $\mu = \mathbf{t}_\Omega^\Lambda \circ \mathbf{r}_\Omega^\Lambda \circ \mu$ so that

$$\rho(x) = \sum \pi_\Omega^\Lambda(w, \{x\})\rho^{\Omega-\Lambda}(w_{\Omega-\Lambda}) \qquad (w \in \mathbf{X}).$$

Then

$$\begin{aligned}
\mu(\mathbf{F} \cap \mathbf{G}) &= \sum \rho(x) \qquad (x \in \mathbf{F} \cap \mathbf{G}) \\
&= \sum \Big[\sum \pi_\Omega^\Lambda(w, \{x\})\rho^{\Omega-\Lambda}(w_{\Omega-\Lambda})\ (x \in \mathbf{F} \cap \mathbf{G}) \Big] \qquad (w \in \mathbf{X}) \\
&= \sum \pi_\Omega^\Lambda(w, \mathbf{F} \cap \mathbf{G})\rho^{\Omega-\Lambda}(w_{\Omega-\Lambda}) \qquad (w \in \mathbf{X}) \\
&= \sum \pi_\Omega^\Lambda(w, \mathbf{F})\rho^{\Omega-\Lambda}(w_{\Omega-\Lambda}) \qquad (w \in \mathbf{G})
\end{aligned}$$

since $\pi_\Omega^\Lambda(w, \mathbf{F} \cap \mathbf{G}) = 0$ if $w \notin \mathbf{G}$ and $\pi_\Omega^\Lambda(w, \mathbf{F} \cap \mathbf{G}) = \pi_\Omega^\Lambda(w, \mathbf{F})$ if $w \in \mathbf{G}$ since $\mathbf{F} \in \mathfrak{F}_{(\Lambda)}$ and $\mathbf{G} \in \mathfrak{F}_{(\Omega-\Lambda)}$. □

5.6. LEMMA. *Let* $\Lambda, \Delta \in \mathcal{C}_\Omega$ *with* $\Lambda \cap \Delta = \varnothing$. *Suppose that there exists a number* $\alpha_{\Lambda,\Delta} \geqslant 1$ *in* $^*R$ *such that for each* $\mathbf{F} \in \mathfrak{F}_{(\Lambda)}$ *we can find a function* $\mathbf{h}\colon \mathbf{X}_\Delta \to {}^*R$ *with*

$$(\alpha_{\Lambda,\Delta})^{-1}\mathbf{h}(w_\Delta) \leqslant \pi_\Omega^\Lambda(w, \mathfrak{F}) \leqslant \alpha_{\Lambda,\Delta}\mathbf{h}(w_\Delta). \tag{5.8}$$

Then for any generalized equilibrium measures, $\mu = L(\mathbf{t}_\Omega^{\Lambda'}(\nu'))$ $\nu = L(\mathbf{t}_\Omega^{\Lambda''}(\nu''))$ *with* $\Lambda \subseteq \Lambda'$ *and* $\Lambda \subseteq \Lambda''$, *we have*

$$\mathbf{H}^{\Lambda\cup\Delta}(\mu, \nu) - \mathbf{H}^\Delta(\mu, \nu) \leqslant 2 \log \alpha_{\Lambda,\Delta}.$$

PROOF. Let $\mu = \mathbf{t}_\Omega^{\Lambda'}(\nu')$, $\nu = \mathbf{t}_\Omega^{\prime\prime}(\nu'')$ with densities ρ and σ and put $\Lambda \cup \Delta = D$. Then

$$\mathbf{g}^D(w_D) = \rho^D(w_D)/\sigma^D(w_D) = \mu(\mathbf{F} \cap \mathbf{G})/\nu(\mathbf{F} \cap \mathbf{G})$$

where

$$\mathbf{F} = \mathbf{p}_{\Lambda\Omega}^{-1}(\{w_\Lambda\}) \in \mathfrak{F}_{(\Lambda)}, \qquad \mathbf{G} = \mathbf{p}_{\Delta\Omega}^{-1}(\{w_\Delta\}) \in \mathfrak{F}_{(\Omega-\Lambda)}.$$

From Lemma 5.5,

$$\begin{aligned}
\mu(\mathbf{F} \cap \mathbf{G}) &= \sum \pi_\Omega^\Lambda(x, \mathbf{F})\rho(x) \qquad (x \in \mathbf{G}) \\
&\leqslant \alpha_{\Lambda,\Delta} \sum \mathbf{h}(x_\Delta)\rho(x) \qquad (x \in \mathbf{G}) \\
&= \alpha_{\Lambda,\Delta}\mathbf{h}(w_\Delta) \sum \rho(x) \qquad (x \in \mathbf{G}) \\
&= \alpha_{\Lambda,\Delta}h(w_\Delta)\mu(G)
\end{aligned}$$

where the third equality follows from $\mathbf{G} = \mathbf{p}_{\Delta\Omega}^{-1}(\{w_\Delta\})$. Similarly, $\nu(\mathbf{F} \cap \mathbf{G}) \geqslant (\alpha_{\Lambda,\Delta})^{-1}\mathbf{h}(w_\Delta)\nu(G)$ and so $\mathbf{g}^D(w_D) \leqslant (\alpha_{\Lambda,\Delta})^2\mathbf{g}^\Delta(w_\Delta)$ from which $\mathbf{q}^{\Lambda,\Delta}(w) \leqslant (\alpha_{\Lambda,\Delta})^2$. Then

$$\begin{aligned}
\mathbf{H}^D(\mu, \nu) - \mathbf{H}^\Delta(\mu, \nu) &= \sum \Phi(\mathbf{q}^{\Lambda,\Delta})(w)\mathbf{g}^\Delta(w_\Delta)\sigma(w) \qquad (w \in \mathbf{X}) \\
&\leqslant 2 \log \alpha_{\Lambda,\Delta}
\end{aligned}$$

and we are through. □

We now come to the main results of this section.

5.7. THEOREM. *Suppose that $\mathcal{C}'$ is adapted to H, and that $\mu \in P(\mathbf{X}, \mathfrak{F})$ and $\nu = L(\nu_y^{\Lambda_{\mathcal{C}}})$ are mutually near H-invariant. If $f^{\Lambda_{\mathcal{C}}}(\mu) + P_y^{\Lambda_{\mathcal{C}}} = 0$ then there is a generalized equilibrium state ζ so that $\mu(\mathbf{F}) = \zeta(\mathbf{F})$ for all $\mathbf{F} \in \mathfrak{F}_s$.*

PROOF. Let $\Lambda \in \mathcal{C}$ be given. From (3.19) we see that if $\Lambda' = \Lambda_{\mathcal{C}'}$ then $\mathbf{H}^{\Lambda'}(\mu, \nu)/\mathbf{m}(\Lambda') = \varepsilon$ is infinitesimal. Since $\mathcal{C}'$ is adapted to H, there exist a *-finite number $n = \mathbf{n}_{\mathbf{H}}(\Lambda, \Lambda')$ of elements $\xi_1, \ldots, \xi_n$ in $\mathbf{H}$ so that $\xi_i(\Lambda) = \Lambda_i$ $(i = 1, \ldots, n)$ are disjoint, $\Gamma_n \subset \Lambda'$ where $\Gamma_k = \bigcup_{j=1}^{k} \Lambda_j$ and $\mathbf{m}(\Lambda)n \geqslant \alpha(\Lambda)\mathbf{m}(\Lambda')$. Also n is infinite since Λ is finite and Λ' is infinite. Then $\mathbf{H}^{\Lambda'}(\mu, \nu) < \eta n$ where $\eta = \varepsilon\, \mathbf{m}(\Lambda)/\alpha(\Lambda)$ is again infinitesimal. From (5.3) we have

$$\mathbf{H}^{\Lambda_1}(\mu, \nu) + \sum_{k=2}^{n} \Big[\sum \Psi(\mathbf{q}^{\Lambda_k, \Gamma_{k-1}})(w)\mathbf{g}^{\Gamma_{k-1}}(w_{\Gamma_{k-1}})\sigma(w)\ (w \in \mathbf{X}) < \eta n \Big].$$

Thus there exists a k satisfying $n/2 \leqslant k \leqslant n$ (and hence infinite) for which

$$\sum \Psi(\mathbf{q}^{\Lambda_k, \Gamma_{k-1}})(w)\mathbf{g}^{\Gamma_{k-1}}(w_{\Gamma_{k-1}})\sigma(w)\ (w \in \mathbf{X}) < 2\eta$$

and hence the quantity on the left is infinitesimal. For this k we can find a $\xi \in \mathbf{H}$ so that $\xi(\Lambda_k) = \Lambda$ and put $\xi(\Gamma_k) = D$ and $\Delta = D - \Lambda$. Notice that $\Delta = \xi(\Gamma_{k-1})$ by (5.5), and D is infinite since k is infinite. Since μ and ν are mutually near H-invariant,

$$\sum \Psi(\mathbf{q}^{\Lambda, \Delta})(w)\mathbf{g}^{\Delta}(w_\Delta)\sigma(w)\ (w \in \mathbf{X}) \simeq 0$$

and so using the transfer of Lemma 5.4 one can show that

$$\sum |\mathbf{q}^{\Lambda, \Delta}(w) - 1|\mathbf{g}^{\Delta}(w_\Delta)\sigma(w)\ (w \in \mathbf{X}) \simeq 0$$

and so

$$\sum |\rho(w) - \mathbf{g}^{\Delta}(w_\Delta)\sigma(w)|\ (w \in \mathbf{F}) \simeq 0$$

for any $\mathbf{F} \in \mathfrak{F}_\Omega$. Now

$$\sum \mathbf{g}^{\Delta}(w_\Delta)\sigma(w)\ (w \in \mathbf{F}) = \sum \mathbf{E}_\nu(\chi_F|\mathfrak{F}_{(\Delta)})(w)\mathbf{g}^{\Delta}(w_\Delta)\sigma(w) \qquad (w \in \mathbf{X})$$

where $\mathbf{E}_\nu(\chi_F|\mathfrak{F}_{(\Delta)})$ is the transfer of the discrete version of the conditional expectation. But $\mathbf{E}_\nu(\chi_F|\mathfrak{F}_{(\Delta)})(w) = \nu(\mathbf{F} \cap \mathbf{G})/\sigma^{\Delta}(w_\Delta)$ where $G = \mathbf{p}_{\Delta\Omega}^{-1}(\{w_\Delta\})$.

Suppose now that $\mathbf{F} \in \mathfrak{F}_{(\Delta)}$ so that $\mathbf{F} = \mathbf{p}_{\Lambda\Omega}^{-1}(\mathbf{F}_0)$ for $\mathbf{F}_0 \in \mathfrak{F}_\Lambda$, and put $\mathbf{F}' = \mathbf{p}_{\Lambda D}^{-1}(\mathbf{F}_0)$. By Lemma 5.5, $\nu(\mathbf{F} \cap \mathbf{G}) = \sum \pi_\Omega^\Lambda(x, \mathbf{F})\sigma(x)$ $(x \in \mathbf{G})$. But by transfer from Lemma 4.5,

$$\pi_D^\Lambda(x_D, \mathbf{F}') \simeq \pi_\Omega^\Lambda(x, \mathbf{F}) \tag{5.9}$$

since D is infinite, and hence

$$\begin{aligned} \nu(\mathbf{F} \cap \mathbf{G}) &\simeq \sum \pi_D^\Lambda(x_D, \mathbf{F}')\sigma(x) \qquad (x \in \mathbf{G}) \\ &= \pi_D^\Lambda(w_D, \mathbf{F}')\sigma^{\Delta}(w_\Delta), \end{aligned}$$

so that $\mathbf{E}_\nu(\chi_F|\mathfrak{F}_{(\Delta)})(w) \simeq \pi_D^\Lambda(\omega_D, \mathbf{F}')$ and we conclude, using (5.9) again, that

$$\sum \rho(w)\ (w \in \mathbf{F}) \simeq \sum \pi_\Omega^\Lambda(w, \mathbf{F})\rho(w)\ (w \in \mathbf{X}).$$

This is true for any $\mathbf{F} \in \mathfrak{F}_{(\Lambda)}$ and hence

$$\sup_{\mathbf{F} \in \mathfrak{F}_{(\Lambda)}} \left| \boldsymbol{\mu}(\mathbf{F}) - \sum \boldsymbol{\pi}_{\Omega}^{\Lambda}(w, \mathbf{F})\boldsymbol{\rho}(w) \ (w \in \mathbf{X}) \right| \simeq 0. \tag{5.10}$$

Now (5.10) is true for any finite Λ and hence by Robinson's Sequential Lemma ([**17**, 6.4.1]), it is true for some infinite $\tilde{\Lambda}$. Then $\boldsymbol{\mu}(\mathbf{F}) = \zeta(\mathbf{F})$ for all $\mathbf{F} \in \mathfrak{F}_{(\tilde{\Lambda})}$ and hence for all $\mathbf{F} \in \mathfrak{F}_s$, where $\zeta = L(\mathbf{t}_{\Omega}^{\tilde{\Lambda}}(\boldsymbol{\mu}_{\Omega - \tilde{\Lambda}}))$ is a generalized equilibrium state. □

The converse of Theorem 5.7 would say that

$$f_y^{\Lambda_{\mathcal{C}}}(\mu) + P_y^{\Lambda_{\mathcal{C}}} = 0 \tag{5.11}$$

for all generalized equilibrium states μ. That this will not be true in general is well known. A recent paper by Föllmer and Snell [**3**] has established a version of the converse by altering the definition of f. We note in passing that at least (5.11) holds if $\mu = L(\boldsymbol{\nu}_y^{\Lambda_{\mathcal{C}}})$. A more general converse can be established under assumptions similar to (7.25) and (7.26) in [**13**].

5.8. Assumption. (i) There exist functions γ_1, γ_2: $\mathcal{C}_{\Omega} \to^* R$ with $\gamma_i(\Lambda)/\mathbf{m}(\Lambda) \simeq 0$ $(i = 1, 2)$ for infinite Λ, such that if $\Lambda \in \mathcal{C}_{\Omega}$ we can find $\tilde{\Lambda} \in \mathcal{C}_{\Omega}$ with $\mathbf{m}(\tilde{\Lambda} - \Lambda) \leqslant \gamma_1(\Lambda)$ so that if $\mathbf{F} \in \mathfrak{F}_{(\Lambda)}$ then there is a function $\mathbf{h}$: $\mathbf{X}_{\tilde{\Lambda} - \Lambda} \to^* R$ with

$$\{\exp(-\gamma_2(\Lambda_{\mathcal{C}}))\}\mathbf{h}(w_{\tilde{\Lambda} - \Lambda}) \leqslant \boldsymbol{\pi}_{\Omega}^{\Lambda_{\mathcal{C}}}(w, \mathbf{F}) \leqslant \exp \gamma_2(\Lambda_{\mathcal{C}})\mathbf{h}(w_{\tilde{\Lambda} - \Lambda}).$$

(ii) With μ a generalized equilibrium state and $\nu = \nu_y^{\Lambda_{\mathcal{C}}}$, the quantity $\mathbf{h}^{\tilde{\Lambda} - \Lambda_{\mathcal{C}}}(\mu, \nu)$ is finite for any $\tilde{\Lambda} \supseteq \Lambda_{\mathcal{C}'}$.

5.9. Theorem. *Under Assumption* 5.8, $f_y^{\Lambda_{\mathcal{C}}}(\mu) + P_y^{\Lambda_{\mathcal{C}}} = 0$ *for all generalized equilibrium states* μ.

Proof. Let $\mu = L(\boldsymbol{\mu})$, $\boldsymbol{\mu} = \mathbf{t}_{\Omega}^{\Lambda'}(\boldsymbol{\nu}')$, and $\nu = L(\boldsymbol{\nu}_y^{\Lambda_{\mathcal{C}}})$. From 5.8(i) find a $\tilde{\Lambda}$ associated with $\Lambda' \cup \Lambda_{\mathcal{C}'}$. By Lemma 5.6,

$$\begin{aligned} \mathbf{f}_y^{\Lambda_{\mathcal{C}}}(\boldsymbol{\mu}) + \mathbf{P}_y^{\Lambda_{\mathcal{C}}} = \mathbf{h}^{\Lambda_{\mathcal{C}}}(\mu, \nu) &\leqslant \frac{1}{\mathbf{m}(\Lambda_{\mathcal{C}'})}\mathbf{H}^{\tilde{\Lambda}}(\mu, \nu) \\ &\leqslant \frac{1}{\mathbf{m}(\Lambda_{\mathcal{C}'})}\left[\mathbf{m}(\tilde{\Lambda} - \Lambda_{\mathcal{C}'})\mathbf{h}^{\tilde{\Lambda} - \Lambda_{\mathcal{C}'}}(\mu, \nu) + 2\gamma_2(\Lambda_{\mathcal{C}'})\right] \\ &\simeq 0 \end{aligned}$$

and we are done. □

The reader is invited to find the local conditions on $(X_A, \mathfrak{F}_A)$ which assure that 7.8 holds; in particular 7.8 will hold under conditions which guarantee (7.25) and (7.26) in [**13**].

6. Concluding remarks. In this paper the ease of analysis on $(\mathbf{X}, \mathfrak{F})$ and the richness of that structure have not been fully exploited since we have mainly been interested in laying the foundations of the theory, and in establishing results corresponding to those in [**13**]. A more detailed comparison of the nonstandard technique with that of [**13**] is instructive, and we conclude with a few remarks in that direction.

First note that the standard existence problem for equilibrium states [**13**, §3] does not occur in our development. Indeed, generalized equilibrium states exist with no assumptions at all on the interaction potential. The crucial assumptions are used in two places: Firstly, to show that each standard equilibrium state corresponds to a generalized one we need Assumption 2.1 (note that Assumption 2.1 is used in [**13**] to show that equilibrium states exist in lattice models). Even without Assumption 2.1 our generalized equilibrium measures would have limiting significance via the results of §4. Assumption 2.2 enters more crucially in both the standard and nonstandard development in establishing the existence (finiteness) of the thermodynamic quantities of §3 and hence the connection with the real world.

Another significant difference between the approach of [**13**] and ours is that in our development each generalized equilibrium measure μ has an *explicit* representation of the form $\mu = L(\mathbf{t}_\Omega^\Lambda(\boldsymbol{\nu}))$ for some infinite Λ and $\boldsymbol{\nu} \in \mathbf{P}(\mathbf{X}_\Omega, \mathfrak{F}_{(\Omega-\Lambda)})$, and so different generalized equilibrium measures (for different Λ and $\boldsymbol{\nu}$) can be effectively compared. In [**13**] on the other hand, an abstract characterization is given for equilibrium measures, and then existence is demonstrated by showing that $\lim \pi_{\Lambda_n}(x, F)$ $(\Lambda_n \to \infty)$ exists for all F in a countable subset $\mathfrak{D}$ of $\mathfrak{F}$ for *some* subsequence $\{\Lambda_n\}$ (see p. 36, [**13**]). From our viewpoint this existence proof singles out a measure of the form $\mu = L(\boldsymbol{\pi}_\Omega^\Lambda(y, \cdot))$ for some infinite Λ. Since the technical complications in [**13**] are considerable, the interrelationships between measures obtained for different subsequences $\{\Lambda_n\}$ would be difficult to investigate.

Consider lastly our proof of the variational principle of §5. Not surprisingly the proof is in some aspects similar to the proof of the corresponding result in [**13**]. But the proofs of all of the necessary lemmas are transfers of finitistic arguments. We are able to use systematically the fact that nonstandard measures on $(\mathbf{X}_\Omega, \mathfrak{F}_\Omega)$ have densities. Thus the nonstandard methods have some attraction, at least from a pedagogical standpoint. In addition, they are likely to be even more effective in continuous statistical mechanics.

References

1. R. M. Anderson, *A non-standard representation for Brownian Motion and Itô integration*, Israel J. Math. **25** (1976), 15–46.

2. M. Davis, *Applied nonstandard analysis*, John Wiley, New York, 1977.

3. H. Föllmer and J. L. Snell, *An "inner" variational principle for Markov fields*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **39** (1977), 187–196.

4. L. L. Helms and P. A. Loeb, *Applications of nonstandard analysis to spin models*, J. Math. Anal. Appl. **69** (1979), 341–352.

5. A. E. Hurd, *Nonstandard models in nonequilibrium statistical mechanics*, Notices Amer. Math. Soc. **24** (1977), A-506.

6. H. Jerome Keisler, *An infinitesimal approach to stochastic analysis* (preprint).

7. O. E. Lanford III, *Time evolution of large classical systems*, Dynamical Systems, Theory and Applications, Lecture Notes in Physics, vol. 38, Springer-Verlag, Berlin and New York, 1974.

8. O. E. Lanford III and D. Ruelle, *Observables at infinity and states with short range correlations in statistical mechanics*, Comm. Math. Phys. **13** (1969), 194–215.

9. A. Lenard, *States of classical mechanical systems of infinitely many particles*. I, Arch. Rational Mech. Anal. **59** (1975), 219–239.

10. P. Loeb, *Conversion from non-standard to standard measure spaces and applications in probability theory*, Trans. Amer. Math. Soc. **211** (1975), 113–122.

11. R. A. Minlos, *Lectures on statistical physics*, Russian Math. Surveys **13** (1968), 137–196.

12. A. Ostebee, P. Gambardella and M. Dresden, *Nonstandard approach to the thermodynamic limit*. I, II, Phys. Rev. A **13** (1976), 878–881; J. Mathematical Phys. **17** (1976), 1570–1578.

13. C. Preston, *Random fields*, Lecture Notes in Mathematics, No. 534, Springer-Verlag, Berlin and New York, 1976.

14. A. Robinson, *Non-standard analysis*, Nederl. Akad. Wetensch. Proc. Ser. A **64** (1961), 432–440.

15. _____, *Non-standard analysis*, North-Holland, Amsterdam, 1970.

16. D. Ruelle, *Statistical mechanics*, W. A. Benjamin, New York, 1969.

17. K. D. Stroyan and W. A. J. Luxemburg, *Introduction to the theory of infinitesimals*, Academic Press, New York, 1976.

Department of Mathematics, University of Victoria, Victoria, British Columbia, Canada