A generalization of Poisson convergence to "Gibbs convergence" with applications to statistical mechanics

Helv Phys Acta
Vol. 67 (1994)

0018-0238/94/010030-13$1.50+0.20/0

A Generalization of Poisson Convergence to "Gibbs Convergence" with Applications to Statistical Mechanics

Lincoln Chayes
Department of Mathematics
UCLA
Los Angeles, CA 90024
USA
lchayes@math.ucla.edu

David Klein
Department of Mathematics
California State University
Northridge, CA 91330
USA
dklein@vax.csun.edu

(5. I. 1994)

Abstract. We prove a theorem which generalizes Poisson convergence for sums of independent random variables taking the values 0 and 1 to a type of "Gibbs convergence" for strongly correlated random variables. The theorem is then used to develop a lattice-to-continuum theory for statistical mechanics.

0. Introduction

The Poisson Convergence Theorem (Corollary 1.1 below) has a statistical mechanical interpretation. Let A be the intersection of a fixed rectangle in $\mathbf{R}^d$ with the d-dimensional lattice $n^{-1}\mathbf{Z}^d$ regarded as a subset of $\mathbf{R}^d$. For each site $m \in A$ associate the Bernoulli random variable X_m^n which takes the value 1 if a particle is present at m (with probability p_m^n) and the value 0 otherwise. The distribution of the collection $\{X_m^n\}$ of independent random variables may be thought of as the Gibbs distribution for an ideal gas on the lattice A. If we let n approach infinity, so that the lattice spacing decreases to zero, and if we maintain for each n approximately the same average density of particles in A, then the Poisson Convergence Theorem says that the lattice ideal gas distributions converge weakly to the standard Gibbs distribution for an ideal gas in the continuum.

On physical grounds, one expects that a similar convergence result holds for interacting particles. This would amount to a generalization of the Poisson Convergence Theorem for certain sums of strongly correlated (essentially Gibbs distributed) random variables. The main parameter now becomes a kind of "chemical activity" z^n_m instead of p^n_m. In the case of independent random variables considered in the Poisson Convergence Theorem, $z^n_m = p^n_m(1-p^n_m)^{-1}$ and an analysis based on these quantities produces an estimate on the rate of convergence that is sharper than the standard fare when the sum of the $z^n_m = p^n_m(1-p^n_m)^{-1}$ is less than 2.

Our generalization, in the form of Theorem 1.1, allows us to develop a lattice-to-continuum theory of classical statistical mechanics including some results for the infinite volume case, i.e., a lattice-to-continuum theory for the thermodynamic limit. In particular we find a potentially useful criterion for the existence of a first-order phase transition in hard-core continuum models in terms of related lattice models.

In Section 1, we introduce notation and state and prove our generalization of the Poisson Convergence Theorem. Section 2 is devoted to applications to statistical mechanics.

1. Gibbs Convergence

Let $A \subset R^d$ be a rectangle with volume $|A|$. For each integer n, let $d(n) = \prod_{i=1}^{d} d_i(n)$ where each d_i is an increasing positive integer valued function. Let $|A|z_n = \sum_{m=1}^{d(n)} z^n_m$ where $z^n_m > 0$ for each m and n. Assume that the collection $\{ z^n_m \}$ is chosen so that A may be partitioned into a regular array of d(n) subrectangles $\{ S^n_1, ..., S^n_{d(n)} \}$ with vol. of $(S^n_m) \equiv v(S^n_m) = \frac{z^n_m}{z_n}$. For each m and n, let $q^n_m \in S^n_m$. We will consider a sequence of functions (f_k) satisfying:

Condition 1.1

a. $f_0 \equiv 1$

b. For each $k \geq 1$, $f_k(x_1, x_2, \ldots, x_k)$ is a nonnegative function, Riemann integrable on A^k, satisfying:

  i. f_k is a symmetric function for each k, i.e.,
  $f_k(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(k)}) = f_k(x_1, x_2, \ldots, x_k)$ for any permutation σ.

  ii. There exists a constant C such that $f_k(x_1, x_2, \ldots, x_k) \leq C^k$ for all $k \geq 0$.

  iii. $f_k(x_1, x_2, \ldots, x_k) = 0$ if $x_i = x_j$ for some $i \neq j$.

Remark 1.1 Condition iii above restricts f_k on a set of Lebesgue measure zero and is therefore not necessary in what follows. We include it because it simplifies some of the discussion below and because it is satisfied by our applications of Theorem 1.1.

Theorem 1.1 Let X^n_m, $1 \le m \le d(n)$, be random variables each taking only the values 0 and 1 with density function

$$P\{X^n_1 = a_1, \ldots, X^n_{d(n)} = a_{d(n)}\} \propto f_k(q^n_{i_1}, \ldots, q^n_{i_k}) \prod_{m=1}^{d(n)} (z^n_m)^{a_m} \tag{1.1}$$

where each $a_i = 0$ or 1 and the indicies on the right side are determined by $k = \sum_{m=1}^{d(n)} a_m = \sum_{m=1}^{k} a_{i_m}$. Assume

$$z_n \xrightarrow[n\to\infty]{} z > 0 \quad \text{and} \quad \max_{1 \le m \le d(n)} z^n_m \xrightarrow[n\to\infty]{} 0. \tag{1.2}$$

Define

$$S_n = X^n_1 + \cdots + X^n_{d(n)}. \tag{1.3}$$

Define a nonnegative integer valued random variable S by the density function,

$$P(S = k) = \frac{\frac{z^k}{k!}\int_{A^k} f_k(x_1, \ldots, x_k)\, dx_1 \cdots dx_k}{\sum_{n=0}^{\infty} \frac{z^n}{n!}\int_{A^n} f_n(x_1, \ldots, x_n)\, dx_1 \cdots dx_n} \tag{1.4}$$

Then $S_n \Rightarrow S$, i.e., S_n converges weakly to S.

proof.

$$P(S_n = k) = \sum_{a_1 + \cdots + a_{d(n)} = k} P\{X^n_1 = a_1, \ldots, X^n_{d(n)} = a_{d(n)}\} \tag{1.5}$$

$$= \frac{(z_n)^k}{Z(n)} \sum_{\{i_1, \ldots, i_k\} \subset \{1, 2, \ldots, d(n)\}} f_k(q^n_{i_1}, \ldots, q^n_{i_k}) \prod_{m=1}^{k} v(S^n_{i_m}) \tag{1.6}$$

where $Z(n)^{-1}$ is the constant of proportionality in (1.1). By Condition 1.1,

$$P(S_n = k) = \frac{\frac{(z_n)^k}{k!} \sum_{i_1=1}^{d(n)} \cdots \sum_{i_k=1}^{d(n)} f_k(q^n_{i_1}, \ldots, q^n_{i_k}) \prod_{m=1}^{k} v(S^n_{i_m})}{\sum_{k=0}^{d(n)} \frac{(z_n)^k}{k!} \sum_{i_1=1}^{d(n)} \cdots \sum_{i_k=1}^{d(n)} f_k(q^n_{i_1}, \ldots, q^n_{i_k}) \prod_{m=1}^{k} v(S^n_{i_m})} \tag{1.7}$$

The numerator in (1.7) is a Riemann sum converging to the numerator in (1.4). The convergence of the denominator in (1.7) to the denominator in (1.4) follows from Condition 1.1 and the Lebesque Dominated Convergence Theorem. ∎

Corollary 1.1 (Poisson Convergence Theorem) Let $\{d(1), d(2),...\}$ be an increasing sequence of positive integers. For $1 \le m \le d(n)$, let $X^n_m = 1$ with probability p^n_m and $X^n_m = 0$ with probability $1-p^n_m$. Let $\{X^n_1, ..., X^n_{d(n)}\}$ be independent. Assume
$\sum_{m=1}^{d(n)} p^n_m \xrightarrow[n\to\infty]{} z > 0$ and $\max_{1\le m\le d(n)} p^n_m \xrightarrow[n\to\infty]{} 0$. Let $S_n = X^n_1 + \cdots + X^n_{d(n)}$. Then

$$P(S_n = k) \xrightarrow[n\to\infty]{} e^{-z}\frac{z^k}{k!}. \tag{1.8}$$

proof. Let $A = [0, 1] \subset R$. Define $z^n_m = p^n_m(1-p^n_m)^{-1}$. Then (1.2) holds. Also,

$$P\{X^n_1 = a_1, ..., X^n_{d(n)} = a_{d(n)}\} = \prod_{m=1}^{d(n)}[a_m p^n_m + (1-a_m)(1-p^n_m)] = \frac{\prod_{m=1}^{d(n)}(z^n_m)^{a_m}}{\prod_{m=1}^{d(n)}(1+z^n_m)} \propto \prod_{m=1}^{d(n)}(z^n_m)^{a_m}$$

where $a_i = 0$ or 1. Choose the $q^n_m \in S^n_m$ to be distinct, but otherwise arbitrary, and $f_n(x_1, x_2, \ldots, x_n) \equiv 1$ whenever the points $\{x_i\}$ are distinct (and otherwise $f_n(x_1, x_2, \ldots, x_n) = 0$). By Theorem 1.1,

$$P(S_n = k) \xrightarrow[n\to\infty]{} \frac{\frac{z^k}{k!}\int_{[0,1]^k} 1\ dx_1\cdots dx_k}{\sum_{n=0}^{\infty}\frac{z^n}{n!}\int_{[0,1]^n} 1\ dx_1\cdots dx_n} = e^{-z}\frac{z^k}{k!} \tag{1.9}$$

since $f_n(x_1, x_2, \ldots, x_n) \equiv 1$ almost surely. ∎

Remark 1.2 Our method of proof provides an estimate for the rate of convergence for Corollary 1.1. For simplicity and with no loss of generality, let $d(n) = n$. Since $f_n \le 1$,

$$\frac{(z_n)^k}{k!}\sum_{i_1=1}^{n}\cdots\sum_{i_k=1}^{n} f_k(q^n_{i_1},\ldots,q^n_{i_k})\prod_{m=1}^{k} v(S^n_{i_m}) \le \frac{(z_n)^k}{k!} \tag{1.10}$$

Replacing f_n by 1 on the left side of (1.10) and subtracting the volume of all subrectangles along any diagonal of $[0, 1]^k$ gives

$$\frac{(z_n)^k}{k!}\sum_{i_1=1}^{n}\cdots\sum_{i_k=1}^{n} f_k(q^n_{i_1},\ldots,q^n_{i_k})\prod_{m=1}^{k} v(S^n_{i_m}) \ge \frac{(z_n)^k}{k!}\left[1-\binom{k}{2}\sum_{i=1}^{n}\left(\frac{z^n_i}{z_n}\right)^2\right] \tag{1.11}$$

for $k \ge 2$. Straightforward manipulations then give,

$$P(S_n = k) \le \frac{e^{-z_n}\dfrac{(z_n)^k}{k!}}{\left[1-\dfrac{(z_n)^{n+1}}{(n+1)!}\right]\left[1-\frac{1}{2}\sum_{i=1}^{n}(z_i^n)^2\right]} \tag{1.12}$$

and

$$P(S_n = k) \ge e^{-z_n}\frac{(z_n)^k}{k!}\left[1-\binom{k}{2}\sum_{i=1}^{n}\left(\frac{z_i^n}{z_n}\right)^2\right] \tag{1.13}$$

for $k \ge 2$. In the case that $z_n = \sum_{m=1}^{n} z_m^n = z$, by (1.13),

$$P(S = k) - P(S_n = k) \le \frac{1}{2}e^{-z}\frac{z^{k-2}}{(k-2)!}\sum_{i=1}^{n}(z_i^n)^2 \tag{1.14}$$

for $k \ge 2$, and otherwise the left side is ≤ 0. If $A = \{k : P(S = k) - P(S_n = k) \ge 0\}$, then

$$\sum_{k\in A}[P(S = k) - P(S_n = k)] \le \sum_{k\in A}\frac{1}{2}e^{-z}\frac{z^{k-2}}{(k-2)!}\sum_{i=1}^{n}(z_i^n)^2 < \frac{1}{2}\sum_{i=1}^{n}(z_i^n)^2. \tag{1.15}$$

Therefore, the total variation norm,

$$\sum_{k=0}^{\infty}|P(S = k) - P(S_n = k)| < \sum_{i=1}^{n}(z_i^n)^2 \tag{1.16}$$

By contrast, using different methods, C. Stein [S], eq. (43) pg. 89 (see also Chen [C], Hodges and LeCam [H-L], and Durret [D]), has shown that if $\sum_{i=1}^{n} p_i^n = z$, then

$$\sum_{k=0}^{\infty}|P(S = k) - P(S_n = k)| \le 2\min(z^{-1}, 1)\sum_{i=1}^{n}(p_i^n)^2 \tag{1.17}$$

Therefore, when $\sum_{m=1}^{n} z_m^n = z$ and $(1-\max_{1\le m\le d(n)} p_m^n)^{-2} < \min(2, 2/z)$, our estimate is sharper than (1.17). We note that $\sum_{i=1}^{n} p_i^n = z$ and $\sum_{m=1}^{n} z_m^n = z$ are mutually exclusive and our estimate (1.16) can also be derived using the methods of [H-L].

2. Lattice to Continuum Statistical Mechanics

We begin with a description of the finite volume continuum theory of classical statistical mechanics.

For a Borel measurable subset $\Lambda \subset \mathbf{R}^d$, let $X(\Lambda)$ denote the set of all locally finite subsets of Λ. $X(\Lambda)$ represents configurations of identical particles in Λ. We let $\emptyset$ denote the empty configuration. Let B_Λ be the σ-field on $X(\Lambda)$ generated by all sets of the form $\{s\in X(\Lambda): |s \cap B| = m\}$, where B runs over all bounded Borel subsets of Λ, m runs over the set of nonnegative integers, and $|\cdot|$ denotes cardinality. We let $(\Omega, S) = (X(\mathbf{R}^d), B_{\mathbf{R}^d})$. For a configuration $x \in \Omega$, let $x_\Lambda = x \cap \Lambda$.

A Hamiltonian H is an S measurable map from the set of finite configurations Ω_F in Ω to $(-\infty, \infty]$ of the form

$$H(x) = \sum_{N=2}^{|x|} \sum_{\substack{y \subset x \\ |y|=N}} \varphi_N(y) \tag{2.1}$$

where the function φ_N on configurations of cardinality N is an N-body potential. The configuration x in (2.1) is coordinatized by $x = \{x_1, x_2, \dots, x_{|x|}\}$. For $x \in X(\Lambda)$, we will sometimes write $H_\Lambda(x)$ instead of H(x). Define the interaction energy between $x \in X(\Lambda)$ and $s \cap \Lambda^c$ by

$$W_\Lambda(x|s) = \sum_{N=2}^{\infty} \sum_{\substack{y \cap x \neq \emptyset \neq y \cap s \\ |y|=N,\, y \subset x \vee s}} \varphi_N(y) \tag{2.2}$$

where we write $x \vee s$ to mean the configuration $x \cup (s \cap \Lambda^c)$. We will sometimes write $W(x \mid s)$ when x and s are located in disjoint regions. Define

$$H_\Lambda(x \mid s) = H_\Lambda(x) + W_\Lambda(x \mid s) \tag{2.3}$$

For a bounded Borel set Λ, let $|\Lambda|$ denote the Lebesgue measure of Λ. The symbol $|\ |$ may therefore represent cardinality or Lebesgue measure, but the meaning will always be clear from the context. For each $i \in \mathbf{Z}^d$, let

$$Q_i = \{r \in \mathbf{R}^d : r^k - 1/2 \leq i^k < r^k + 1/2, k=1,\dots,d\}$$

so that the unit cubes $\{Q_i\}$ partition $\mathbf{R}^d$. Define $|x_i| \equiv |x_{Q_i}| = |x \cap Q_i|$.

We assume that H satisfies the following:

Condition 2.1

a) H is translation invariant

b) H is stable, i.e., $H(x) \geq -K|x|$ for some $K \geq 0$ and all $x \in \Omega_F$

c) H(x) is lower regular. For any Λ_1 and Λ_2 which are each finite unions of unit cubes with $x \subset \Lambda_1$ and $s \subset \Lambda_2$,

$$W(x \mid s) \geq - \sum_{i \in \Lambda_1} \sum_{j \in \Lambda_2} \|i-j\|^{-\lambda} |x_i| |s_j|$$

where $K > 0$, $\lambda > d$ are fixed.

d) H(x) is tempered. There exists $R_o > 0$ such that with the same notation as in part c, assuming Λ_1 and Λ_2 are separated by a distance R_o or more,

$$W(x \mid s) \leq K \sum_{i \in \Lambda_1} \sum_{j \in \Lambda_2} \|i-j\|^{-\lambda} |x_i| |s_j|$$

e) $\exp\{-\varphi_n(x_1, \ldots, x_n)\}$ is Riemann integrable in any closed rectangle of $\mathbf{R}^{dn}$ for all $n \geq 2$.

Temperedness and lower regularity allow $W(x|s)$ to be defined when s is an infinite configuration of particles. We assume in this section that the configuration s is chosen so that $W(x \mid s)$ is finite.

Let $X_N(\Lambda)$ be the set of configurations of cardinality N in Λ and let $T: \Lambda^N \to X_N(\Lambda)$ be the map which takes the ordered N-tuple $(x_1, \ldots, x_N)$ to the (unordered) set $\{x_1, \ldots, x_N\}$. In a natural way T defines an equivalence relation on Λ^N and $X_N(\Lambda)$ may be regarded as the set of equivalence classes induced by T. For n = 1, 2, 3, ..., let $d^n x$ be the projection of nd-dimensional Lebesgue measure onto $X_N(\Lambda)$ under the projection $T: \Lambda^N \to X_N(\Lambda)$. The measure $d^0 x$ assigns mass 1 to $X_0(\Lambda) = \{\emptyset\}$. The unnormalized Poisson measure on $(X(\Lambda), B_\Lambda)$ with parameter z, interpreted here as fugacity, is given by

$$\nu_\Lambda(dx) = \sum_{n=0}^{\infty} \frac{z^n}{n!} d^n x \tag{2.4}$$

If $\Lambda \cap A = \emptyset$ where Λ and A are Borel sets, then $(X(\Lambda), B_\Lambda, \nu_\Lambda) \times (X(A), B_A, \nu_A)$ may be identified with $(X(\Lambda \cup A), B_{\Lambda\cup A}, \nu_{\Lambda\cup A})$ via $x_\Lambda \times x_A = x_\Lambda \cup x_A$.

The grandcanonical partition function in Λ with boundary configuration s is defined by

$$Z_\Lambda(s) = \int_{X(\Lambda)} \exp\{-\beta H(x|s)\} \nu_\Lambda(dx) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \int_{\Lambda^n} \exp\{-\beta H(x_1, \ldots, x_n|s)\}\ dx_1 \cdots dx_n \tag{2.5}$$

where β in inverse temperature. The pressure $p(\beta, z, \Lambda)$ for the Hamiltonian H in Λ is given by

$$\beta p(\beta, z, \Lambda) = \frac{\ln Z(\emptyset)}{|\Lambda|} \tag{2.6}$$

For a a bounded Borel set Λ in $\mathbf{R}^d$ and a configuration s in Λ^c, the finite volume Gibbs state with boundary configuration s for H, $\beta > 0$, and z is

$$\sigma_\Lambda(dx|s) = \frac{\exp\{-\beta H(x|s)\}}{Z_\Lambda(s)} \nu_\Lambda(dx) \tag{2.7}$$

The probability that there are k particles in a Borel subset Γ of Λ may be determined by integrating the characteristic function for the set $\{ x \subset \Lambda: |x \cap \Gamma| = k\}$ with respect to $\sigma_\Lambda(dx \mid s)$.

We now describe lattice theories of statistical mechanics in finite volume in a form suitable for Theorem 2.1 below.

Let $\{S^n_m\}$ be a partition of $\mathbf{R}^d$ by translates of $\prod_{i=1}^{d}(0,\frac{1}{n}]$ by linear combinations of the standard basis vectors with coefficients of the form $\frac{m}{n}$ where $m \in \mathbf{Z}$. For each m and n choose a point $q^n_m \in S^n_m$ and define $Q(n) \equiv \{q^n_m\}$. For example, $Q(n) = \frac{1}{n}\mathbf{Z}^d \subset \mathbf{R}^d$. Let Λ be a Jordan-measurable set in $\mathbf{R}^d$ (i.e. Λ is bounded and the boundary of Λ has Lebesgue measure zero).

Remark 2.1 The Hamiltonian H(x) restricted to $Q(n) \cap \Lambda$ can be rewritten in a form more commonly associated with lattice models. Let the integer n be fixed. For each lattice site $q^n_m \in Q(n)$ associate the occupation variable (or "spin" variable) s_m which takes the value 1 if a particle is present at q^n_m and takes the value 0 otherwise. Let s denote the configuration $(s_1, \dots, s_{d(n)})$ in the rectangle $Q(n) \cap \Lambda$ such that $s_j = 1$ if and only if $j \in \{i_1, \dots, i_k\}$. Then we may identify

$$H(s) = H(q^n_{i_1},\dots,q^n_{i_k}) = \sum_{m=2}^{k} \sum_{j_1<j_2<\cdots<j_m} J_{j_1 j_2\cdots j_m} s_{j_1} s_{j_2} \cdots s_{j_m}$$

where $J_{j_1 j_2 \cdots j_m} = \varphi_m(q^n_{j_1},\dots,q^n_{j_m})$.

The grandcanonical partition function $Z_\Lambda(n, s)$ for the lattice gas on $Q(n) \cap \Lambda$ with the Hamiltonian H given by (2.1) restricted to $Q(n) \cap \Lambda$, inverse temperature β, and fugacity z is given by

$$Z_\Lambda(n, s) = \sum_{k=0}^{|Q(n)\cap\Lambda|} \left(\frac{z}{n^d}\right)^k \sum_{\{q^n_{i_1},\dots,q^n_{i_k}\}\subset Q(n)\cap\Lambda} \exp\{-\beta H(q^n_{i_1},\dots,q^n_{i_k} | s)\} \tag{2.8}$$

The grandcanonical pressure is then,

$$\beta p_n(\beta, z, \Lambda) = \frac{\ln Z(n,\varnothing)}{|Q(n)\cap\Lambda| n^{-d}} \tag{2.9}$$

The finite volume Gibbs state is defined on the measurable space $(\{0, 1\}^{Q(n)\cap\Lambda}, B_\Lambda(n))$ where $B_\Lambda(n)$ is the σ-field consisting of all subsets of $\{0, 1\}^{Q(n)\cap\Lambda}$. Elements of $\{0, 1\}^{Q(n)\cap\Lambda}$ may be identified in an obvious way with subsets of $Q(n) \cap \Lambda$. The finite volume Gibbs state $\sigma_\Lambda(\ |\ s)_n$ with boundary configuration s is given by

$$\sigma_\Lambda(B|s)_n = \sum_{q\in B} \frac{\exp\{-\beta H(q|s)\}}{Z_\Lambda(n,s)} \left(\frac{z}{n^d}\right)^{|q|} \tag{2.10}$$

where $B \in B_\Lambda(n)$. The probability that there are k particles in a subset Γ of $Q(n) \cap \Lambda$ may be determined by integrating the characteristic function for the set $\{x \subset Q(n) \cap \Lambda : |x \cap \Gamma| = k\}$ with respect to $\sigma_\Lambda(dx\ |\ s)_n$.

Definition 2.1 Let $\Lambda \subset \mathbf{R}^d$ be Jordan-measurable. The sequence of lattice Gibbs states $\{\sigma_\Lambda(\ \mid s)_n\}$ converges weakly to the continuum Gibbs state $\sigma_\Lambda(\ \mid s)$ if for any Jordan-measurable set Γ contained in Λ, the probability according to $\sigma_\Lambda(\ \mid s)_n$ that there are exactly k particles in $Q(n) \cap \Gamma$ converges to the probability according to $\sigma_\Lambda(\ \mid s)$ that there are exactly k particles in Γ, as n approaches infinity.

Theorem 2.1 For a fixed Jordan-measurable $\Lambda \subset \mathbf{R}^d$, as $n \to \infty$,

a) the lattice partition function $Z_\Lambda(n, s)$ converges to the continuum partition function $Z_\Lambda(s)$.

b) the lattice pressures $p_n(\beta, z, \Lambda)$ converge to the continuum pressure $p(\beta, z, \Lambda)$

c) for any s, the lattice Gibbs states $\sigma_\Lambda(\ \mid s)_n$ converge weakly to the continuum Gibbs states $\sigma_\Lambda(\ \mid s)$.

proof. Let $A = \prod_{i=1}^{d}[a_i, b_i]$ be a closed rectangle with integer vertices a_i and b_i containing Λ. For convenience relabel $\{q_m^n\}$ so that $Q(n) \cap A = \{q_1^n, \dots, q_{d(n)}^n\}$.

a) Define random variables $\{X_m^n\}$ associated with the lattice sites $\{q_m^n\}$ taking the values 0 and 1 with distribution $P\{X_1^n = a_1, \dots, X_{d(n)}^n = a_n\} \propto f_k(q_{i_1}^n, \dots, q_{i_k}^n) \prod_{m=1}^{d(n)} (z_m^n)^{a_m}$ where $z_m^n = zn^{-d}$ and $f_k(q_{i_1}^n, \dots, q_{i_k}^n) = \chi_{\Lambda^k}(q_{i_1}^n, \dots, q_{i_k}^n)\exp\{-\beta\, H((q_{i_1}^n, \dots, q_{i_k}^n) \mid s)\}$ and where χ_{Λ^k} is the characteristic function for Λ^k and k is determined as in (1.1). Then f_k satisfies Condition 1.1. If we define as in Theorem 1.1 $S_n = X_1^n + \cdots + X_{d(n)}^n$, then by Theorem 1.1,

$$Z_\Lambda(n, s) \equiv P(S_n = 0)^{-1} \to P(S = 0)^{-1} \equiv Z_\Lambda(s)\,. \tag{2.11}$$

b) This follows immediately from part a and the continuity of the logarithm.

c) Let Γ be a Jordan measurable subset of Λ and let m be a nonnegative integer. With the notation of Theorem 1.1, let $z_i^n = zn^{-d}$ and

$$f_k(q_{i_1}^n, \dots, q_{i_k}^n) = \chi_{\Lambda^m \vee (\Lambda\setminus\Gamma)^{k-m}}(\{q_{i_1}^n, \dots, q_{i_k}^n\})\exp\{-\beta\, H(q_{i_1}^n, \dots, q_{i_k}^n \mid s)\},$$

where $\chi_{\Lambda^m \vee (\Lambda\setminus\Gamma)^{k-m}}(\{q_{i_1}^n, \dots, q_{i_k}^n\}) = 1$ provided $|\{q_{i_1}^n, \dots, q_{i_k}^n\} \cap \Lambda| = m$ and $|\{q_{i_1}^n, \dots, q_{i_k}^n\} \cap \Lambda\setminus\Gamma| = k - m$; otherwise $\chi_{\Lambda^m \vee (\Lambda\setminus\Gamma)^{k-m}}(\{q_{i_1}^n, \dots, q_{i_k}^n\}) = 0$.

The collection $\{f_k\}$ satisfies Condition 1.1. By Theorem 1.1, $P(S_n = 0)^{-1} \to P(S = 0)^{-1}$. Combining this with part a shows that,

$$Z_\Lambda(n, s)^{-1} \sum_{k=0}^{|Q(n)\cap A|} \left(\frac{z}{n^d}\right)^k \sum_{\{q_{i_1}^n, \dots, q_{i_k}^n\} \subset Q(n)\cap A} f_k(q_{i_1}^n, \dots, q_{i_k}^n)$$

converges as $n \to \infty$ to

$$Z_\Lambda(s)^{-1}\sum_{k=0}^{\infty}\frac{z^k}{k!}\int_{\Lambda^k} f_k(x_1,\ldots,x_k)\,dx_1\cdots dx_k$$

In other words, the probability according to $\sigma_\Lambda(\ \mid s)_n$ that there are exactly m particles in $Q(n)\cap\Gamma$ converges to the probability according to $\sigma_\Lambda(\ \mid s)$ that there are exactly m particles in Γ, as n approaches infinity. ∎

We now show how the finite volume lattice approximation of the continuum pressure may be extended to the infinite volume case, i.e., after taking the thermodynamic limit. For simplicity we assume that our Hamiltonian H is given by a pair potential with a hard-core of radius R. This has the effect of limiting the number of particles which can accumulate in any unit cube Q_i in $\mathbf{R}^d$. The lattice and continuum infinite volume pressures are given respectively by

$$p_n(\beta, z) = \beta^{-1} \lim_{|\Lambda|\uparrow\infty}\frac{\ln Z_\Lambda(n,\varnothing)}{|Q(n)\cap\Lambda| n^{-d}} \tag{2.12}$$

and

$$p(\beta, z) = \beta^{-1} \lim_{|\Lambda|\uparrow\infty}\frac{\ln Z_\Lambda(\varnothing)}{|\Lambda|} \tag{2.13}$$

where limit may be taken via an increasing sequence of cubes centered at the origin.

Theorem 2.2 If H is determined by a pair potential with hard-core radius R satisfying Condition 2.1, then $\lim_{n\to\infty} p_n(\beta, z) = p(\beta, z)$ for each β, $z > 0$.

proof. Let Λ be a cube centered at the origin containing an integer number of unit cubes of the form Q_i. Let Λ be partitioned into cubes $\{\Lambda_k\}$ of equal size and integer dimensions. The size and number of these cubes will be determined below. Assume that Λ is large enough to make the partitions we describe below possible. With this notation we may write,

$$Z_\Lambda(\varnothing) = \int\prod_k \exp\{-\beta H(x\cap\Lambda_k)\}\prod_k \exp\{-\tfrac{1}{2}\beta W(x\cap\Lambda_k | x\cap\Lambda_k^c)\}\nu_\Lambda(dx) \tag{2.14}$$

From the hard-core assumption and Condition 2.1c,

$$W(x\cap\Lambda_k | x\cap\Lambda_k^c) \geq -N^2 K \sum_{i:Q_i\subset\Lambda_k}\ \sum_{j:Q_j\subset\Lambda_k^c} \|i-j\|^{-\lambda} \tag{2.15}$$

for some integer N, and the bound on the right side is the same for all k. Therefore,

$$Z_A(\emptyset) \le \prod_k \exp\left\{\frac{\beta}{2} N^2 K \sum_{i:Q_i \subset A_k} \sum_{j:Q_j \subset A_k^c} \|i-j\|^{-\lambda}\right\} \int \prod_k \exp\{-\beta H(x \cap A_k)\} \nu_A(dx)$$

$$= \prod_k \exp\left\{\frac{\beta}{2} N^2 K \sum_{i:Q_i \subset A_k} \sum_{j:Q_j \subset A_k^c} \|i-j\|^{-\lambda}\right\} \prod_k Z_{A_k}(\emptyset) \tag{2.16}$$

because of the product structure of ν_A. An inequality analogous to (2.16) holds with $Z_A(\emptyset)$ and $Z_{A_k}(\emptyset)$ replaced by $Z_A(n, \emptyset)$ and $Z_{A_k}(n,\emptyset)$ respectively.

To obtain lower bounds for $Z_A(\emptyset)$ and $Z_A(n, \emptyset)$, we partition A differently. Necessarily, the hard-core diameter 2R of the Hamiltonian H is less than R_0, where R_0 is the constant appearing in Condition 2.1d. Define $x \in A$ to be a "corridor point" if x lies within a distance $\frac{1}{2}R_0$ of some A_k not containing x or within a distance $\frac{1}{2}R_0$ of A^c and let C be the collection of all corridor points. Then A is partitioned by the sets C and a collection $\{B_k\}$ of disjoint cubes of equal size with each B_k a proper subset of some A_k. Any two cubes B_i and B_j are separated by a distance of at least R_0. Let $\emptyset_C$ denote the event $\{x \subset A: x \cap C = \emptyset\}$, i.e., there are no particles in C. Then

$$Z_A(\emptyset) \ge \int_{\emptyset_C} \exp\{-\beta H(x)\} \nu_A(dx)$$

$$= \int_{\emptyset_C} \prod_k \exp\{-\beta H(x \cap B_k)\} \prod_k \exp\{-\tfrac{1}{2}\beta W(x \cap B_k | x \cap B_k^c)\} \nu_A(dx) \tag{2.17}$$

By Condition 2.1d and the fact that $B_k \subset A_k$,

$$W(x \cap B_k | x \cap B_k^c) \le N^2 K \sum_{i:Q_i \subset A_k} \sum_{j:Q_j \subset A_k^c} \|i-j\|^{-\lambda} \tag{2.18}$$

It follows that

$$Z_A(\emptyset) \ge \prod_k \exp\left\{-\frac{\beta}{2} N^2 K \sum_{i:Q_i \subset A_k} \sum_{j:Q_j \subset A_k^c} \|i-j\|^{-\lambda}\right\} \int_{\emptyset_C} \prod_k \exp\{-\beta H(x \cap B_k)\} \nu_A(dx)$$

$$= \prod_k \exp\left\{-\frac{\beta}{2} N^2 K \sum_{i:Q_i \subset A_k} \sum_{j:Q_j \subset A_k^c} \|i-j\|^{-\lambda}\right\} \prod_k Z_{B_k}(\emptyset) \tag{2.19}$$

An inequality analogous to (2.19) holds with $Z_A(\emptyset)$ and $Z_{A_k}(\emptyset)$ replaced by $Z_A(n, \emptyset)$ and $Z_{A_k}(n,\emptyset)$ respectively.

Combining (2.16) and (2.19) gives for any k,

$$\frac{1}{|A|} \ln Z_A(\emptyset) - \frac{1}{|A|} \ln Z_A(n, \emptyset) \le$$

$$\frac{1}{|A_k|} \ln Z_{A_k}(\emptyset) - \frac{1}{|A_k|} \ln Z_{B_k}(n,\emptyset) + \frac{1}{|A_k|} \beta N^2 K \sum_{i:Q_i \subset A_k} \sum_{j:Q_j \subset A_k^c} \|i-j\|^{-\lambda} \tag{2.20}$$

Taking the limit as the cube A increases to $\mathbf{R}^d$ gives

$$\beta p(\beta, z) - \beta p_n(\beta, z) \le$$

$$\frac{1}{|A_k|}\ln Z_{A_k}(\emptyset) - \frac{|B_k|}{|A_k|}\frac{1}{|B_k|}\ln Z_{B_k}(n,\emptyset) + \frac{1}{|A_k|}\beta N^2 K \sum_{i:Q_i\subset A_k}\sum_{j:Q_j\subset A_k^c}\|i-j\|^{-\lambda} \quad (2.21)$$

Now given $\varepsilon > 0$, choose fixed A_k and B_k large enough so that :

$$\frac{1}{|A_k|}\beta N^2 K \sum_{i:Q_i\subset A_k}\sum_{j:Q_j\subset A_k^c}\|i-j\|^{-\lambda} < \varepsilon \quad (2.22)$$

$$\frac{|B_k|}{|A_k|} > 1 - \varepsilon \quad (2.23)$$

$$\left|\frac{1}{|A_k|}\ln Z_{A_k}(\emptyset) - \frac{1}{|B_k|}\ln Z_{B_k}(\emptyset)\right| < \varepsilon \quad (2.24)$$

For such fixed A_k and B_k choose n sufficiently large (using Theorem 2.1b) so that

$$\left|\frac{1}{|B_k|}\ln Z_{B_k}(\emptyset) - \frac{1}{|B_k|}\ln Z_{B_k}(n,\emptyset)\right| < \varepsilon \quad (2.25)$$

Combining (2.21) through (2.25) gives,

$$\beta p(\beta, z) - \beta p_n(\beta, z) \le \frac{1}{|A_k|}\ln Z_{A_k}(\emptyset) - (1-\varepsilon)\left[\frac{1}{|A_k|}\ln Z_{A_k}(\emptyset) - 2\varepsilon\right] + \varepsilon \quad (2.26)$$

Hence,

$$\beta p(\beta, z) - \beta p_n(\beta, z) \le \varepsilon\frac{1}{|A_k|}\ln Z_{A_k}(\emptyset) + 3\varepsilon \quad (2.27)$$

Since $\{\frac{1}{|A_k|}\ln Z_{A_k}(\emptyset)\}$ is a bounded sequence for all A_k with integer vertices, it follows that

$$\liminf_{n\to\infty} p_n(\beta, z) \ge p(\beta, z) \quad (2.28)$$

An analogous argument shows that $\limsup_{n\to\infty} p_n(\beta, z) \le p(\beta, z)$ which establishes the theorem. ∎

A probability measure σ on (Ω, S) is a continuum Gibbs state (or infinite volume continuum Gibbs state) for H, β, and z if it satisfies the DLR equations, i.e., if

$$\sigma\left(\int f(x\vee s)\,\sigma_\Lambda(dx \mid s)\right) = \sigma(f) \quad (2.29)$$

for every bounded S-measurable function f from Ω to $\mathbf{R}$ and every bounded Borel set Λ. The definition of infinite volume Gibbs states for the lattice models we consider is completely analogous.

For any of the grandcanonical lattice or continuum models we consider, a first order phase transition is said to occur if the (infinite volume) pressure fails to be differentiable as a function of the chemical potential $\mu \equiv \beta^{-1}\log z$ at some point (β_0, μ_0). In case the pressure is not differentiable with respect to μ at the point (β_0, μ_0), there exist two translation-invariant Gibbs states whose expectations of the number of particles in a unit cube are equal respectively to $D^-p(\beta_0, \mu_0)$ and $D^+p(\beta_0, \mu_0)$, where D^+ (resp. D^-) denotes the right-hand derivative (resp. the left-hand derivative) with respect to μ, and where p may be either the lattice or continuum pressure, which we now regard as a function of β, μ instead of β, z (see, e.g., [K-Y] and the references contained therein). The quantity $D^+p(\beta_0, \mu_0) - D^-p(\beta_0, \mu_0)$ is therefore the gap between the high density and low density states of matter which can co-exist in equilibrium at the values β_0, μ_0 of the inverse temperature and chemical potential.

The following theorem may be useful in proving the existence of first-order phase transitions for continuum models of statistical mechanics.

Theorem 2.3 Assume that the Hamiltonian H is determined by a pair potential with a hard-core and that it satisfies Condition 2.1. If each element in a subsequence of lattice pressures (on lattices of the form Q(n)) exhibits a first order phase transition at β_n, μ_n with the gap between the high density and low density states bounded below by a positive number, and $(\beta_n, \mu_n) \to (\beta, \mu)$, then the continuum pressure exhibits a first order phase transition at β, μ with the same lower bound on the gap between high density and low density states.

proof. It is routine to check that if $(P_k(t))$ is a sequence of convex functions defined on a open interval I of the real line and $(P_k(t))$ converges pointwise to a convex function P(t) and $t_k \to t_0 \in I$, then $P_k(t_k) \to P(t_0)$ and

$$D^-P(t_0) \leq \liminf_{k\to\infty} D^-P_k(t_k) \leq \limsup_{k\to\infty} D^+P_k(t_k) \leq D^+P(t_0). \tag{2.30}$$

(Here $D^+ P(t_0)$ and $D^-P(t_0)$ denote respectively the right and left hand derivatives of the fuction P at t_0.) It follows that $\lim_{n\to\infty} p_n(\beta_n, \mu) = p(\beta, \mu)$ when $\beta_n \to \beta$, and that

$$D^-p(\beta, \mu) \leq \liminf_{n\to\infty} D^-p_n(\beta_n, \mu_n) \leq \limsup_{n\to\infty} D^+p_n(\beta_n, \mu_n) \leq D^+p(\beta, \mu) \tag{2.31}$$

when $\mu_n \to \mu$, where the one-sided derivatives are again taken with respect to μ.∎

REFERENCES

[C] Chen, L.H.Y., "On Convergence of Poisson binomial to Poisson distributions" *Ann. Probab.* 2, 178-180 (1974) and "Poisson approximation for dependent trials" *Ann. Probab.* 3, 534-545 (1975).

[D] Durret, R., *Probability: Theory and Examples* Wadsworth & Brooks, Pacific Grove, California (1991).

[H-L] Hodges, J. L. Jr., LeCam, L., "The Poisson approximation to the binomial distribution" *Ann. Math. Statist.* 31, 737-740 (1960).

[K-Y] Klein, D., Yang, W.S. "A characterization of first-order phase transitions for superstable interactions in classical statistical mechanics" *J. Stat. Phys.* 71 1043-1062 (1993)

[S] Stein, C., "Approximate computation of expectations." *Institute of Mathematical Statistics. Lecture Notes-Monograph Series,* Vol. 7 (1986).