Ann. Inst. Statist. Math.
Vol. 42, No. 2, 203–220 (1990)

MEAN CHARACTERISTICS OF GIBBSIAN POINT PROCESSES

SHIGERU MASE

Faculty of Integrated Arts and Sciences, Hiroshima University, Hiroshima 730, Japan

(Received June 20, 1988; revised June 15, 1989)

Abstract. This paper deals with the derivation of an exact expression of mean characteristics of planar global Gibbsian point processes having pair potential functions. The method is analogous to that of the Mayer expansion of grand partition functions, i.e., the reciprocal of the normalizing constant of Gibbsian distribution (well-known in statistical physics). The explicit infinite series expansion of a logarithm of a class of mean quantities with respect to the activity parameter z is derived and the expression of its coefficients is given. The validity of this expansion for a range of z is also shown. Examples of mean characteristics to which this expansion can be applied are given. Finally, a simple numerical example is given in order to show the usage of this expansion as a numerical approximant of mean characteristics.

Key words and phrases: Gibbsian point process, pair potential function, spatial statistics, stochastic geometry, mean characteristics, Mayer expansion, nearest neighbor distribution, spherical contact distribution, numerical approximation.

1. Introduction

Statistical analysis of spatial point patterns has made an impressive advance as summarized in the book of Stoyan *et al.* (1987). This is mainly based on the progress of the theory of spatial point process as an abstract mathematical framework. Through these researches the importance of the Gibbsian point process as a model construction principle has been widely recognized. Besides the fact that it has a long history as a genuine physical model, the attractive feature of the Gibbsian model is that it can offer a variety of complex point patterns starting from simple potential functions which are easily interpretable as attractive and/or repulsive forces acting among points. Formally, a Gibbsian process is defined by giving its (conditional or unconditional) likelihood explicitly. Although this likelihood is apparently very simple, it includes a considerable combinatorial complexity. This complexity is indispensable in a sense, since the likelihood must represent all of the mutual dependencies among points, but presents difficult theoretical as well as numerical problems. As to the partition function Ξ, i.e., the reciprocal of the normalizing constant of the likelihood, physicists have developed a method to derive a series expansion of $\log \Xi$ called the Mayer expansion; i.e., the expansion in powers of the activity (fugacity) parameter z, which is feasible for numerical evaluations in several cases. Ogata and Tanemura (1984) applied this expansion to compute the maximum likelihood estimator of potential functions.

On the other hand, in the second order statistical analysis or the nonparametric analysis of the Gibbsian process, one sometimes needs to compute various mean characteristics of the process or its Palm distribution. Presently almost all these quantities are calculated as corresponding sample quantities of simulated point patterns, since theoretical computations are laborious or, simply no such methods are known. Without doubt, the simulation method is flexible and indispensable in the practical analysis based on Gibbsian models. But, at the same time, we cannot forget that the generation of a Gibbsian point pattern is asymptotic in nature and is a delicate technique. A useful, though possibly not easy, way of assessment is to compare simulated quantities with corresponding theoretical ones.

In this paper we will show that Mayer's method is also applicable to a certain class of mean quantities and will derive expansions of their logarithms. It is important to distinguish between the *fugacity* or *activity expansion* and the *Mayer expansion* of a quantity f. The first means the expansion in powers of z of f itself and the latter is the expansion in powers of z of the logarithm of f. In some cases the Mayer expansion has a simpler structure than the corresponding activity expansion and, hence is more suitable to numerical works. Our result is particularly applicable to mean characteristics such as nearest neighbor distance distribution functions and spherical contact distribution functions, which are basic in stochastic geometry; their theoretical expressions for the Gibbsian process have not been known. Also we will give a complete concise expression of all coefficients in terms of certain cluster integrals associated with block graphs of arguments. Higher order coefficients may be of little practical importance, but still have some theoretical interest.

The idea of the activity expansion (or the virial expansion which is closely related to it) of mean characteristics, especially of correlation functions important in physics, is never new in physics. For example, Ruelle (1969) discusses a special calculus on formal series and derives the activity expansion of correlation functions. However, this is not the Mayer expansion but the expansion of a correlation function itself. Coefficients of this expansion are not explicitly given and seem not to be simplified further. Westcott (1982) applied the underlying idea of the activity expansion to get bounds on the distribution function of the minimum nearest neighbor distance of hard disk systems on the surface of a sphere. The virial expansion of some characteristics such as a radial distribution function can be derived from that of a pressure using a physical relation (see Grandy (1987), Chapter 9). This also leads to an expansion of a characteristic, but not of its logarithm. An exception is the remarkable paper of Minlos and Pogosian (1977). They considered an expansion which is essentially the same as the formula (3.4) below, but without taking into account the conditioning by outer configurations. Also they neither discussed its relation to mean characteristics nor gave its explicit expression.

Finally, it must be stressed that all existing results deal with means in the physical sense. In physics it has been the custom to take first expectations with respect to a Gibbsian process on a bounded region and then take their limits as a region expands (thermo-dynamical limits). This is conceptually very different from ours, that is, expectations being taken with respect to global Gibbsian processes from the first. Although our intuition suggests that two viewpoints yield the same objects under appropriate conditions, this change of viewpoints is never a mathematical pedagogism and requires an additional analysis. Many basic concepts in stochastic geometry can be defined and justified rigorously only when we take account of stationary or motion-invariant, hence necessarily global, processes. An example is the nearest neighbor distribution. It can be defined rigorously only with respect to the Palm distribution of a stationary point process. Furthermore, we will show the validity of expansion with respect to global Gibbs distributions for a certain range of the activity parameter.

2. Hard-core Gibbsian point process

Let $\boldsymbol{B}$ and $\boldsymbol{B}_0$ be Borel sets and bounded Borel sets of $\boldsymbol{R}^2$, respectively. The set of locally finite (i.e., intersections with every sets in $\boldsymbol{B}_0$ are finite) subsets of $\boldsymbol{R}^2$ is denoted by $\mathscr{M}$. $\mathscr{M}$ is endowed with the σ-field $\mathscr{F}$ generated by functions $N_B(\mu) = \#(\mu \cap B)$, $B \in \boldsymbol{B}_0$. Also let $\mathscr{F}_G$, $G \in \boldsymbol{B}$, be the sub-σ-field of $\mathscr{F}$ generated by functions N_B, $B \in \boldsymbol{B}_0$ and $B \subset G$.

Let Φ: $[0, +\infty) \to (-\infty, +\infty]$ be a measurable function (a pair potential function) which satisfies the following conditions:

(A1) Hard-core condition: There is a constant r_0 such that $\Phi(r) = +\infty$ for $0 \leq r \leq r_0$.

(A2) The following integral exists:

$$c_0 = \int_0^\infty |e^{-\Phi(r)} - 1| r dr .$$

(A3) There is a non-increasing function $\Phi_0(r)$, $r \geq r_0$, such that $|\Phi(r)| \leq \Phi_0(r)$, $\lim_{r\to\infty} r^2 \Phi_0(r) = 0$, and

$$c_1 = \int_{r_0}^{\infty} r\Phi_0(r)dr < \infty .$$

Using the constant r_0 above, we define the space $\mathscr{M}_0$ of subsets of $\boldsymbol{R}^2$ every two points of which are apart at least r_0. We will need the following estimate in the sequel.

LEMMA 2.1. *If* $\{x_n\} \in \mathscr{M}_0$ *and* $R = \min_n |x_n| \geq r_0$, *then the inequality*

$$\left| \sum_{n\geq 1} \Phi(|x_n|) \right| \leq c_2\Phi_1(R)$$

holds, where $c_2 = \pi/((\pi/3 - \sqrt{3}/4)r_0^2)$, $R = \min_{n\geq 1} |x_n|$ *and*

$$\Phi_1(R) = R^2\Phi_0(R-0) + 2\int_R^{\infty} r\Phi_0(r)dr .$$

PROOF. Consider two disks $D_1 = \{x; |x| < s\}$, $s \geq r_0$ and D_2. The radius of D_2 is r_0 and its center is on the boundary of D_1. The intersection $D_1 \cap D_2$ is divided into two parts D_1' and D_2' by the segments combining two intersection points of boundary circles. Assume the center of D_2 is in D_2'. The area of D_1' has the minimum value $(\pi/3 - \sqrt{3}/4)r_0^2$ when $s = r_0$. Let $T(s)$ be the number of points x_n which are in D_1. From the preceding remark $T(s) \leq c_2 s^2$ for $s \geq r_0$ and $T(R) = 0$.

Let $R' > R$ be a point of continuity of Φ_0. Fix a sequence $R = s_1 < s_2 < \cdots < s_n = R'$, then

$$\begin{aligned}
S' &= \sum_{i=1}^{n-1} \sum\{|\Phi(|x_n|)|; s_i \leq |x_n| < s_{i+1}\} \\
&\leq \sum_{i=1}^{n-1} \Phi_0(s_i)[T(s_{i+1}) - T(s_i)] \\
&= \sum_{i=2}^{n-1} [\Phi_0(s_{i-1}) - \Phi_0(s_i)]T(s_i) + \Phi_0(s_{n-1})T(s_n) \\
&\leq c_2 \sum_{i=2}^{n-1} s_i^2[\Phi_0(s_{i-1}) - \Phi_0(s_i)] + c_2 s_n^2\Phi_0(s_{n-1}) .
\end{aligned}$$

Since this is valid for all $\{s_i\}$, S' is majorated by

$$c_2\left[R'^2\Phi_0(R') - \int_R^{R'} s^2 d\Phi_0(s) \right].$$

Applying the integration in parts formula for Stieltjes integrals (see, e.g., Saks (1937), Chapter 3), and letting $R' \to \infty$, the assertion follows.

The interaction $U(\mu)$ and the mutual interaction $U(\mu; \nu)$, $\mu, \nu \in \mathscr{M}$ are defined by

$$U(\mu) = \frac{1}{2} \Sigma \{\Phi(|x-y|); x, y \in \mu, x \neq y\},$$

$$U(\mu; \nu) = \Sigma \{\Phi(|x-y|); x \in \mu, y \in \nu\}.$$

Empty sums should be taken to be zero. From the preceding lemma the interaction U satisfies the *stability condition*, that is, for every finite $\mu \in \mathscr{M}$,

$$(2.1) \qquad U(\mu) \geq -b\#\mu,$$

where the *stability constant* b can be taken to be $c_2\Phi_1(r_0)/2$, for example. Note that $U(\mu) = \infty$ if $\mu \notin \mathscr{M}_0$. Furthermore, if $\mu \in \mathscr{M}$ is finite and $\nu \in \mathscr{M}_0$, then

$$(2.2) \qquad |U(\mu; \nu)| \leq 2b\#\mu.$$

This estimate (2.2) follows if we apply the estimate of the last proof on each sum $\Sigma\{\Phi(|x-y|);\ y \in \nu\}$, $x \in \nu$.

Note that $U(\mu; \nu) = \infty$ if either $\mu \notin \mathscr{M}_0$ or $R < r_0$. Denote $\{x_1, x_2, \ldots, x_n\}$ by $(x)_n$ and $dx_1 dx_2 \cdots dx_n$ by $d(x)_n$. Also we will use notations $(x)'_n = \{x_2, \ldots, x_n\}$ and $d(x)'_n = dx_2 \cdots dx_n$. Define functions

$$\psi(\mu) = \exp(-U(\mu)),$$

$$\psi(\mu; \nu) = \exp(-U(\mu; \nu)),$$

$$\psi(\mu|\nu) = \psi(\mu)\psi(\mu; \nu).$$

For $\forall$ positive z, $\forall\ G \in \boldsymbol{B}_0$ and $\forall\ \mu \in \mathscr{M}_0$, the grand partition function Ξ is defined by the formula

$$(2.3) \qquad \Xi = \Xi(\Phi, z, G, \mu \cap G^c) = 1 + \sum_{n=1}^{\infty} \frac{z^n}{n!} \int_{G^n} \psi((x)_n | \mu \cap G^c) d(x)_n,$$

which has a finite value because of (2.1) and (2.2). Actually, Ξ is a polynomial in z because of the hard-core condition. Now we can state the definition of a Gibbsian point process.

DEFINITION 2.1. (Preston (1976) and Nguyen and Zessin (1979)) A point process $\boldsymbol{P} = \boldsymbol{P}_{\Phi, z}$ on $(\mathscr{M}, \mathscr{F})$ is a Gibbsian point process corresponding to the potential Φ and the activity z if the conditional probability $\boldsymbol{P}(A|\mathscr{F}_{G^c})$ is given by

$$(2.4)\qquad \Xi^{-1}\left[1+\sum_{n=1}^{\infty}\frac{z^n}{n!}\int_{G^n}\chi_A((x)_n\cup(\mu\cap G^c))\psi((x)_n|\mu\cap G^c)d(x)_n\right]$$

$\boldsymbol{P}$-a.s. μ for $\forall\, G\in\boldsymbol{B}_0$ and $\forall\, A\in\mathscr{F}$. Equivalently, $\boldsymbol{P}$ is Gibbsian iff, for all measurable, non-negative functions $h(x,\mu)$,

$$(2.5)\qquad \int_{\mathscr{M}}\sum_{x\in\mu}h(x,\mu)\boldsymbol{P}(d\mu)=z\int_{\boldsymbol{R}^2}\int_{\mathscr{M}}h(x,\{x\}\cup\mu)\psi(\{x\}|\mu)dx\boldsymbol{P}(d\mu)\,.$$

A point process $\boldsymbol{P}$ is said to be stationary (resp. rotation-invariant) if $\boldsymbol{P}(A)$, $A\in\mathscr{F}$, is unchanged for every translations (resp. rotations) of A. If $\boldsymbol{P}$ is both stationary and rotation-invariant, it is said to be motion-invariant. The existence of motion-invariant Gibbsian point processes for every $z\geq 0$ and Φ with conditions (A1)–(A3) is proven in Preston (1976). Also it can be shown that these $\boldsymbol{P}$ are supported by $\mathscr{M}_0$ because of the hard-core condition (A1).

Remark 2.1. In literature, the activity is also called a *fugacity* (see Grandy (1987)). Also the constant $-\log z$ is called a *chemical potential* (Grandy (1987)), or a *chemical activity* is Stoyan *et al.* (1987).

3. Mean formula for Gibbsian point process

In the following the symbol μ denotes both an element and a random element in $\mathscr{M}$. Many mean characteristics of a Gibbsian point process are, or contain as parts, expectations of the form

$$(3.1)\qquad \boldsymbol{E}\left\{\prod_{x\in\mu}g(x)\right\}.$$

Our aim is to express this expectation in terms of Φ and z so as to obtain a numerically calculable formula. This will be done by applying Mayer's device to conditional means in place of the unconditional one $\boldsymbol{E}\left\{\prod_{\mu}g(x)\right\}$ itself. The function $g(x)$ should satisfy the condition (A4) below and can be fairly arbitrary. Several important examples will be given in the last section.

Let $\mathscr{G}_n$ and $\mathscr{C}_n$ be the sets of all labeled graphs and all labeled connected graphs on $\{1,2,\ldots,n\}$, $n\geq 1$, respectively. With each $H\in\bigcup_n\mathscr{G}_n$ we associate a real number, a weight, $W(H)$. Let F_n be the sum of $W(H)$

for $\forall\ H \in \mathscr{G}_n$ and f_n be the sum of $W(H)$ for $\forall\ H \in \mathscr{C}_n$. The next lemma, the first Mayer theorem, enables us to express F_n's in terms of simpler f_n's.

LEMMA 3.1. (Uhlenbeck and Ford (1963)) *If a weight W has two properties*:

(1) $W(H_1) = W(H_2)$ *if H_1 and H_2 are equivalent up to labeling and*

(2) $W(H) = \Pi\, W(C)$ *where the product goes over all the connected components C of H, then*

$$1 + \sum_{n=1}^{\infty} F_n \frac{z^n}{n!} = \exp\left(\sum_{n=1}^{\infty} f_n \frac{z^n}{n!} \right).$$

In another word, $F_n = Y_n(f_1, f_2, \dots, f_n)$ where Y_n is the n-th (complete exponential) Bell polynomial. Bell polynomials have the generating function:

$$\exp\left(\sum_{n=1}^{\infty} x_n \frac{t^n}{n!} \right) = 1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} Y_n(x_1, \dots, x_n)\ .$$

Also we will need partial exponential Bell polynomials $B_{n,k}$ which are defined by the generating relation:

$$\exp\left(u \sum_{n=1}^{\infty} x_n \frac{t^n}{n!} \right) = 1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} \left\{ \sum_{k=1}^{n} u^k B_{n,k}(x_1, \dots, x_{n-k+1}) \right\}.$$

As to the definition of Bell polynomials, see, for example, Comtet (1974).

Denote $\exp(-\Phi(|x_i - x_j|)) - 1$ by ϕ_{ij} and let $\phi_H(x)_n$ be the product of ϕ_{ij} for all edges (i, j) of a graph $H \in \mathscr{G}_n$. Define the weight W_μ for each $\mu \in \mathscr{M}_0$ by

$$W_\mu(H) = \int_{G^n} \phi_H(x)_n \psi((x)_n; \mu \cap G^c)\, d(x)_n\ .$$

Also let $\phi(x)_n$ be the sum of ϕ_H over all $H \in \mathscr{C}_n$. It is easy to see that W_μ satisfies conditions of the first Mayer theorem and, hence,

$$\log \Xi = \upsilon \sum_{n=1}^{\infty} r_n z^n\ , \tag{3.2}$$

where υ is the area of G and $r_n = r_n(\Phi, G, \mu \cap G^c)$ is given by

$$\frac{1}{\upsilon n!} \int_{G^n} \phi(x)_n \psi((x)_n; \mu \cap G^c)\, d(x)_n\ . \tag{3.3}$$

Next, let $g(x)$ be an arbitrary function with the condition:

(A4) there is a bounded and non-increasing function $g_0(r)$ such that $|g(x)-1| \leq g_0(|x|)$, $\lim_{r\to\infty} r^2 g_0(r) = 0$ and $\int_0^\infty r g_0(r)dr < \infty$.

Under this condition, $c_3 = \int_0^\infty |g(x)-1|dx < \infty$. Also, we can show by the same argument as in the proof of Lemma 2.1 that $\sum_{x\in\mu} |g(x)-1|$ is absolutely convergent and bounded for $\forall \mu \in \mathscr{M}_0$. Hence $\prod_\mu g(x)$ is a.s. well-defined and bounded. Corresponding to the previously defined Ξ, we consider $\Xi_g = \Xi_g(\Phi, z, G, \mu \cap G^c)$ given by

$$1 + \sum_{n=1}^{\infty} \frac{z^n}{n!} \int_{G^n} \prod_i g(x_i) \psi((x)_n | \mu \cap G^c) d(x)_n ,$$

which has a finite value and is a polynomial in z. The appropriate weight associated with Ξ_g is

$$W_\mu(H) = \int_{G^n} \prod_i g(x_i) \phi_H(x)_n \psi((x)_n; \mu \cap G^c) d(x)_n$$

and the first Mayer theorem gives the expansion

$$\log \Xi_g = v \sum_{n=1}^{\infty} r_n' z^n , \tag{3.4}$$

where $r_n' = r_n'(g, \Phi, G, \mu \cap G^c)$ is given by

$$\frac{1}{vn!} \int_{G^n} \prod_i g(x_i) \phi(x)_n \psi((x)_n; \mu \cap G^c) d(x)_n . \tag{3.5}$$

THEOREM 3.1. *The conditional expectation of* $\prod_{\mu\cap G} g(x)$ *given* $\mathscr{F}_{G^c}$ *for every* $G \in \boldsymbol{B}_0$ *has the Mayer expansion*

$$\boldsymbol{E}\left\{ \prod_{\mu\cap G} g(x) | \mathscr{F}_{G^c} \right\} = \exp\left(\sum_{n=1}^{\infty} \delta_n^0 \frac{z^n}{n!} \right) \tag{3.6}$$

$\boldsymbol{P}$*-a.s.* μ, *where* $\delta_n^0 = \delta_n^0(g, \Phi, G, \mu \cap G^c)$ *is given by*

$$\delta_n^0 = \int_{G^n} \left[\prod_i g(x_i) - 1 \right] \phi(x)_n \psi((x)_n; \mu \cap G^c) d(x)_n . \tag{3.7}$$

PROOF. Note that, from (2.4), this conditional expectation is the ratio Ξ_g/Ξ. Hence (3.6) follows from (3.2) and (3.3). Also the relation (3.7) follows from (3.3) and (3.5).

Let $\mathscr{B}_n$ and $\mathscr{B}_{0n}$ be the sets of block graphs on $\{1,2,\dots,n\}$ and $\{0,1,2,\dots,n\}$, respectively. Recall that a graph H is a block iff H as well as all the subgraphs with $n-1$ vertices are connected. The set of tree graphs on $\{1,2,\dots,n\}$ is denoted by $\mathscr{T}_n$. Also, we will need the set $\mathscr{D}_{0n}$ consisting of connected graphs on $\{0,1,2,\dots,n\}$ whose restrictions to $\{1,2,\dots,n\}$ are still connected. Denote $g(x_i)-1$ by g_i. It is convenient to denote g_i by ϕ_{0i} and to extend the definition of $\phi_H(x)_n$ for graphs H on $\{0,1,\dots,n\}$. Define three quantities $\varepsilon_n^0=\varepsilon_n^0(g,\Phi,G,\mu\cap G^c)$, $\gamma_n^0=\gamma_n^0(\Phi,G,\mu\cap G^c)$ and $\beta_n^0=\beta_n^0(\Phi,G,\mu\cap G^c)$ by

$$\int_{G^n}\Big[\sum_{H\in\mathscr{B}_{0n}}\phi_H(x)_n\Big]\psi((x)_n;\mu\cap G^c)d(x)_n\,,$$

$$\frac{1}{v}\int_{G^n}\phi(x)_n\psi((x)_n;\mu\cap G^c)d(x)_n\,,$$

$$\frac{1}{v}\int_{G^n}\Big[\sum_{H\in\mathscr{B}_n}\phi_H(x)_n\Big]\psi((x)_n;\mu\cap G^c)d(x)_n\,,$$

respectively. We should consider γ_1^0 to be 1.

THEOREM 3.2. *The limits*

$$\delta_n=\lim\delta_n^0=\int_{\boldsymbol{R}^{2n}}\Big[\sum_{H\in\mathscr{D}_{0n}}\phi_H(x)_n\Big]d(x)_n\,,$$

$$\varepsilon_n=\lim\varepsilon_n^0=\int_{\boldsymbol{R}^{2n}}\Big[\sum_{H\in\mathscr{B}_{0n}}\phi_H(x)_n\Big]d(x)_n\,,$$

$$\gamma_n=\lim\gamma_n^0=\int_{\boldsymbol{R}^{2n-2}}\Big[\sum_{H\in\mathscr{C}_n}\phi_H(x)_n\Big]d(x)_n'\,,$$

$$\beta_n=\lim\beta_n^0=\int_{\boldsymbol{R}^{2n-2}}\Big[\sum_{H\in\mathscr{B}_n}\phi_H(x)_n\Big]d(x)_n'$$

as $G\uparrow\boldsymbol{R}^2$ exist and are independent of μ for $\forall\,\mu\in\mathscr{M}_0$, hence, for $\boldsymbol{P}$-a.s. μ.

PROOF. We will give the proof of the case of δ_n since others can be similarly proven. Let $\|g\|=\sup_x|g(x)|$. Then $\|g\|\geq 1$. The inequality

$$\Big|\prod_i g(x_i)-1\Big|\leq\sum_{i=1}^n\|g\|^{n-i}|g_i|\leq\|g\|^{n-1}\sum_{i=1}^n|g_i| \tag{3.8}$$

is valid. The following estimate is known (see Duneau *et al.* (1975)),

$$|\phi(x)_n| \le e^{2(n-2)b} \sum_{H \in \mathscr{T}_n} \prod_{(i,j) \in H} |\phi_{ij}| \,. \tag{3.9}$$

From (2.2) we can show that $|\psi((x)_n; \mu \cap G^c)|$ is bounded by e^{2bn} and is convergent to 1 as $G \uparrow \boldsymbol{R}^2$ for each $(x)_n$. On the other hand, since $\#\mathscr{T}_n = n^{n-2}$, we have

$$\int_{\boldsymbol{R}^{2n}} |g_k| \sum_{H \in \mathscr{T}_n} \prod_{(i,j) \in H} |\phi_{ij}| d(x)_n \le n^{n-2} e^{2(n-2)b} c_0^{n-1} c_3 \,. \tag{3.10}$$

Hence $\delta_n = \lim \delta_n^0$ exists from the dominated convergence theorem. It is obvious that δ_n is independent of $\mu \in \mathscr{M}_0$. Hence the assertion has been proven.

Although integrals which define δ_n's are numerically calculable in principle, their integrands involves too many summands as soon as n increases slightly. Hence it is desirable to express them by relatively simpler quantities such as ε_n's and β_n's. This will be done by employing the idea of the second Mayer theorem. Define auxiliary functions:

$$C(x) = \sum_{n=1}^{\infty} \gamma_n x^n / n!, \qquad B(x) = \sum_{n=2}^{\infty} \beta_n x^n / n! \,,$$

$$R(x) = xC'(x), \qquad S(x) = B'(R(x)) \,.$$

THEOREM 3.3. (Second Mayer theorem) *The following relations hold*:

$$R(x) = x \exp(S(x)) \,, \tag{3.11}$$

$$\sum_{n=1}^{\infty} \delta_n \frac{x^n}{n!} = \sum_{n=1}^{\infty} \varepsilon_n \frac{x^n}{n!} \exp(nS(x)) = \sum_{n=1}^{\infty} \frac{\varepsilon_n}{n!} (R(x))^n \,. \tag{3.12}$$

PROOF. The first equality is the so-called second Mayer theorem and enables us to express γ_n's in terms of β_n's. Its proof is given in Uhlenbeck and Ford ((1963), Chapter 2). Also a simpler proof is given in Harary and Palmer ((1973), Chapter 3 and Theorem 3.1). Although the last proof is given for the particular weight $W \equiv 1$, i.e., the labeled counting, it can be easily generalized to weights satisfying conditions of Lemma 3.1. We shall prove (3.12) along the style of Harary and Palmer's proof. The key is the concept of rooted graphs and the technique of rooted graph enumeration.

Consider the weight

$$W(H)=\begin{cases}\int_{R^{2n}}\phi_H(x)_n d(x)_n & \text{if} \quad H\in \bigcup_{n\geq 1}\mathscr{D}_{0n}\,,\\ \int_{R^{2n}}\phi_H(x)_n d(x)'_n & \text{if} \quad H\in \bigcup_{n\geq 2}\mathscr{C}_n\,.\end{cases}$$

Note that the last integral is independent of x_1. Fix an $H\in\mathscr{D}_{0n}$. If two subgraphs H_1 and H_2 of H are blocks and have two vertices in common, then their union is also a block. So, there is the maximal block subgraph of H containing (i,j) for each edge $(i,j)\in H$. Hence H can be decomposed into block subgraphs, every two of which have at most one vertex in common, and the vertex 0 belongs to only one of them. From this remark, we can show that H can be decomposed uniquely into $H_0, H_1,\dots, H_k$, $0\leq k$, such that

(1) $H_0\in\mathscr{B}_{0m}$ for $1\leq \exists m\leq n$,

(2) each H_i, $i\geq 1$, is connected and has just one vertex $\neq 0$ in common with H_0 and

(3) every two of H_i, $i\geq 1$, have at most one vertex in common which is also a vertex of H_0.

Moreover, it can be shown that $W(H)$ is equal to the product of $W(H_0),\dots, W(H_k)$.

The function $S(x)$ can be interpreted as the exponential generating function of sums of weights for rooted, connected and labeled graphs where the root is unlabeled and just one block subgraph contains the root (see the proof of Theorem 3.1 of Harary and Palmer (1973)). Therefore the exponential generating function of δ_n's is equal to

$$\sum_{n=1}^{\infty}\varepsilon_n\frac{x^n}{n!}\left[\sum_{m=0}^{n}\binom{n}{m}\left\{\sum_{j_1=1}^{\infty}\frac{S(x)^{j_1}}{j_1!}\right\}\cdots\left\{\sum_{j_m=1}^{\infty}\frac{S(x)^{j_m}}{j_m!}\right\}\right],$$

where

(1) the second sum chooses m vertices $v_1,\dots,v_m$ out of non-zero vertices of H_0 from which rooted connected graphs $H_1,\dots,H_k$ hang,

(2) the sum with respect to j_i selects the number of rooted connected graphs which hang from the vertex v_i,

(3) $S(x)^{j_i}$ stands for the labeled enumeration of j_i rooted connected graphs while their roots are unlabeled (because they are identified with v_i) and

(4) the division by $j_i!$ means that j_i graphs are unordered.

The last relation is equivalent to (3.12) and the proof is completed.

COROLLARY 3.1. *Functional relations* (3.11) *and* (3.12) *are solved explicitly as: for* $n\geq 1$

$$(3.13)\qquad \gamma_n = \frac{1}{n} Y_{n-1}(n\beta_2, n\beta_3, \dots, n\beta_n)\,,$$

$$(3.14)\qquad \delta_n = \sum_{i=1}^{n} \varepsilon_i B_{n,i}(1, 2\gamma_2, 3\gamma_3, \dots, (n-i+1)\gamma_{n-i+1})\,.$$

PROOF. Let

$$f(y) = \exp\left(\sum_{k=1}^{\infty} \beta_{k+1} \frac{y^k}{k!}\right),$$

then $f(0) = 1$ and

$$f(y)^n = 1 + \sum_{i=1}^{\infty} \frac{y^i}{i!} Y_i(n\beta_2, n\beta_3, \dots, n\beta_i)\,.$$

The equation (3.11) can be rewritten as $y = xf(y)$ and this functional equation can be solved uniquely by Lagrange's inversion formula as

$$y = p_1 x + p_2 \frac{x^2}{2!} + \cdots, \qquad p_n = \left(\frac{d}{dy}\right)^{n-1} f(y)^n \bigg|_{y=0}$$

(see Harary and Palmer (1973), Chapter 1). Now y must be equal to $R(x)$ and hence $p_n = n\gamma_n$. On the other hand, $p_n = Y_{n-1}(n\beta_2, n\beta_3, \dots, n\beta_n)$ and the first assertion follows.

Next let $f_m(x) = x^m$. From the differential formula of composite functions

$$\left(\frac{d}{dx}\right)^n R(x)^m \bigg|_{x=0} = \sum_{k=1}^{n} B_{n,k}(p_1, \dots, p_{n-k+1}) f_m^{(k)}(0)$$

$$= \begin{cases} m!\,B_{n,m}(p_1, \dots, p_{n-m+1}) & \text{if} \quad m \le n\,, \\ 0 & \text{if} \quad m > n\,. \end{cases}$$

Hence, substituting power series expansions of $R(x)^m$ into the right-hand side of (3.12), we can show (3.14).

For example, the first several γ_n's and δ_n's are:

$$\gamma_2 = \beta_2, \qquad \gamma_3 = \beta_3 + 3\beta_2^2, \qquad \gamma_4 = \beta_4 + 12\beta_2\beta_3 + 16\beta_2^3\,,$$

$$\gamma_5 = \beta_5 + 20\beta_2\beta_4 + 15(\beta_3 + 10\beta_2^2)\beta_3 + 125\beta_2^4\,,$$

$$\gamma_6 = \beta_6 + 30\beta_2\beta_5 + 60(\beta_3 + 6\beta_2^2)\beta_4 + 540\beta_2(\beta_3 + 4\beta_2^2)\beta_3 + 1296\beta_2^5\,,$$

$$\delta_1 = \varepsilon_1, \quad \delta_2 = \varepsilon_2 + 2\beta_2\varepsilon_1, \quad \delta_3 = \varepsilon_3 + 6\beta_2\varepsilon_2 + 3(\beta_3 + 3\beta_2^2)\varepsilon_1 ,$$

$$\delta_4 = \varepsilon_4 + 12\beta_2\varepsilon_3 + 12(\beta_3 + 4\beta_2^2)\varepsilon_2 + 4(\beta_4 + 12\beta_2\beta_3 + 16\beta_2^3)\varepsilon_1 ,$$

$$\delta_5 = \varepsilon_5 + 20\beta_2\varepsilon_4 + 30(\beta_3 + 5\beta_2^2)\varepsilon_3 + 20(\beta_4 + 15\beta_2\beta_3 + 25\beta_2^3)\varepsilon_2 + 5(\beta_5 + 20\beta_2\beta_4 + 15\beta_3(\beta_3 + 10\beta_2^2) + 125\beta_2^4)\varepsilon_1 .$$

If we let $g(x) = \exp(-\Phi(|x|))$, then $\varepsilon_i = \beta_{i+1}$. Therefore

$$\delta_1 = \beta_2, \quad \delta_2 = \beta_3 + 2\beta_2^2, \quad \delta_3 = \beta_4 + 9\beta_2\beta_3 + 9\beta_2^3 ,$$

$$\delta_4 = \beta_5 + 16\beta_2\beta_4 + 12(\beta_3 + 8\beta_2^2)\beta_3 + 64\beta_2^4 ,$$

$$\delta_5 = \beta_6 + 25\beta_2\beta_5 + 50(\beta_3 + 5\beta_2^2)\beta_4 + 125\beta_2(3\beta_3 + 10\beta_2^2)\beta_3 + 625\beta_2^5 .$$

Finally note that, if we introduce the scale parameter σ and replace $g(x)$ and $\Phi(r)$ by $g(\sigma^{-1}x)$ and $\Phi(r/\sigma)$, respectively, then δ_n and ε_n will be multiplied by σ^{2n}, and β_n and γ_n will be multiplied by $\sigma^{2(n-1)}$.

4. Mayer expansion of global mean

The family of σ-fields $\{\mathscr{F}_{G^c}\}$ is decreasing to the tail σ-field $\mathscr{F}_\infty$ as $G \uparrow \boldsymbol{R}^2$. Therefore, from the reverse martingale convergence theorem,

$$\lim_{G \uparrow \boldsymbol{R}^2} \boldsymbol{E}\left\{ \prod_{\mu \cap G} g(x) | \mathscr{F}_{G^c} \right\} = \lim_{G \uparrow \boldsymbol{R}^2} \left[\boldsymbol{E}\left\{ \prod_{\mu} g(x) | \mathscr{F}_{G^c} \right\} \Big/ \prod_{\mu \cap G^c} g(x) \right] = \boldsymbol{E}\left\{ \prod_{\mu} g(x) | \mathscr{F}_\infty \right\}$$

for $\boldsymbol{P}$-a.s. μ. Hence we expect the formula of the Mayer expansion of the global mean, that is, the unconditional mean, of $\prod_\mu g(x)$ to be:

$$\boldsymbol{E}\left\{ \prod_{\mu} g(x) \right\} = \exp\left(\sum_{n=1}^{\infty} \delta_n \frac{z^n}{n!} \right). \tag{4.1}$$

But, so far the derivation of the expansion (3.6) is only formal. In this section we shall show that (4.1) is valid at least in some region of z.

THEOREM 4.1. *The Mayer expansion* (4.1) *is valid at least in the region* $z < 1/(c_0 e^{4b+1}\|g\|)$. *If the potential* Φ *is non-negative, the expansion is valid in the region* $z < 1/(c_0 e\|g\|)$.

Let G_n be the disk $\{x;\ |x| \leq n\}$. Fix two numbers z_0, q with

$$0 < z_0 < q < 1/(c_0 e^{4b+1} \|g\|) .$$

Consider the sequence of functions

$$f_n(z) = \Xi_g(\Phi, z, G_n, \mu \cap G_n^c) / \Xi(\Phi, z, G_n, \mu \cap G_n^c) .$$

From the reverse martingale convergence theorem (see Doob (1953))

$$\lim_{n \to \infty} f_n(z_0) = \boldsymbol{E}_{\Phi, z_0} \left\{ \prod_{\mu} g(x) | \mathscr{F}_\infty \right\} \tag{4.2}$$

for $\boldsymbol{P}_{\Phi, z_0}$-a.s. μ. Fix a $\mu \in \mathscr{M}_0$ so that (4.2) holds. From the expansion (3.6) and the estimates of δ_n^0 given in the proof of Theorem 3.2, we can see that $f_n(z)$, considered as functions of complex variable z, are analytic and uniformly bounded for $|z| \leq q$. Therefore, $\{f_n\}$ is normal, that is, compact with respect to uniform convergence topology in $|z| \leq q$ (see, e.g., Hille (1977), Chapter 15) and there is a subsequence $\{f_{n'}\}$ which is uniformly convergent in $|z| \leq q$. The limit function $f(z)$ is also analytic in $|z| \leq q$ and, since the uniform convergence induces convergence of corresponding Taylor coefficients, it must be equal to

$$\exp\left(\sum_{n=1}^{\infty} \delta_n \frac{z^n}{n!} \right)$$

by Theorems 3.1 and 3.2. Finally, from (4.2), $f(z_0) = \boldsymbol{E}_{\Phi, z_0} \left\{ \prod_{\mu} g(x) | \mathscr{F}_\infty \right\}$ and the first assertion is completed. If the potential function Φ is non-negative, we can let the stability constant $b = 0$ in the proof of Theorem 3.2. So the second assertion follows from the first immediately.

Remark 4.1. We can show from the proof of Theorem 3.2 that the series $\sum \delta_n z^n / n!$ is convergent for $|z| < 1/(c_0 e^{2b+1} \|g\|)$.

5. Examples

Let us list several examples of mean quantities which can be dealt with using our method. First, let $g(x)$ be of the form

$$g(x) = \psi(\{x\}; (x)_m) = \exp\left[- \sum_{i=1}^{m} \Phi(|x - x_i|) \right]$$

and $g^{(m)}(x_1, \ldots, x_m)$ be corresponding means $\boldsymbol{E}\left\{ \prod_{\mu} g(x) \right\}$. In statistical physics, functions $\rho^{(m)} = z^m \psi((x)_m) g^{(m)}$, z being the activity, are known as m-point correlation functions. If $\boldsymbol{P}$ is motion-invariant, $\rho^{(1)}$ is a constant λ. Also, $\rho^{(2)}(x_1, x_2)$ depends only on $|x_1 - x_2|$ and can be written as $\rho^{(2)}(|x_1 - x_2|)$. The constant λ is the intensity of $\boldsymbol{P}$, that is, the mean number of points per unit area. If C_1 and C_2 are two disks with infinitesimal areas dV_1 and dV_2 and if the distance between centers is t, then $\rho^{(2)}(t) dV_1 dV_2$ is interpreted as the probability $\boldsymbol{P}\{\mu \cap C_1 \neq \emptyset, \mu \cap C_2 \neq \emptyset\}$. The correlation function $\rho^{(2)}$ plays a central role in the second-order analysis of spatial point process. For example, $\rho^{(2)}(t) = \lambda^2 K'(t)/2\pi t$, where K is the second reduced moment function, i.e., Ripley's K-function (for details, see Stoyan *et al.* (1987)).

Next fix a bounded Borel set A and let $g(x)$ be the indicator function of A^c. Then the assumption (A4) is trivially satisfied and we have $\boldsymbol{E}\left\{ \prod_{\mu} g(x) \right\} = \boldsymbol{P}\{\mu \cap A = \emptyset\}$. In particular, if A is $b(0, t)$, the closed disk with center 0 and radius t, then $F(t) = 1 - \boldsymbol{E}\left\{ \prod_{\mu} g(x) \right\} = \boldsymbol{P}\{\mu \cap A \neq \emptyset\}$ is the distribution function of the nearest distance to points of μ from an arbitrary settled point. Stoyan *et al.* (1987) coined the good name *spherical contact distribution function* for this distribution function.

Also, let $g^*(x) = g(x) e^{-\Phi(|x|)}$. Since the reduced Palm distribution $\boldsymbol{P}_0(d\mu)$ is equal to $z/\lambda \cdot \psi(\{0\}|\mu) \boldsymbol{P}(d\mu)$, we have the relation

$$\boldsymbol{E}\left\{ \prod_{\mu} g^*(x) \right\} = \lambda/z \cdot \boldsymbol{E}_0 \left\{ \prod_{\mu} g(x) \right\},$$

that is, we can also deal with mean quantities with respect to the Palm distribution. If $g(x)$ is the indicator function of $b(0, t)^c$ as above, then $D(t) = 1 - z/\lambda \cdot \boldsymbol{E}\left\{ \prod_{\mu} g^*(x) \right\} = \boldsymbol{P}_0\{\mu \cap b(0, t) \neq \emptyset\}$ is called the *nearest neighbor distance distribution function*. This is the distribution function of the nearest distance to points of μ from a typical point of μ. As to the practical role of these distribution functions in stochastic geometry, see, for example, Diggle (1983).

It may be instructive to give a numerical example. Since the purpose is to explain how to compute using the expansion (4.1), no perfection or completeness is intended. As can be seen, the computation is never easy and needs heavy computer work. Let us consider the spherical contact distribution function $F(t)$ of the hard disk system. The function $g(x)$ used is of the form

$$g(x) = g_t(x) = \begin{cases} 1 & \text{if} \quad |x| > t\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Our choice of the potential function is the simplest one, that is,

$$\Phi(t) = \begin{cases} +\infty & \text{if} \quad t \leq 1\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Hence the corresponding point patterns are considered as centers of mutually disjoint unit disks uniformly dispersed on the plane. The point patterns are approximated by the equilibrium state of lattice-valued birth and death Markov chains as explained in Preston (1976) (see also Stoyan *et al.* (1987), Chapter 5). The region $[0,100] \times [0,100]$ is discretized into the square lattice $L = \{(i/100, j/100);\ 0 \leq i, j \leq 10000\}$. The initial point patterns X_0 on L are given by sequential random packing so that the packing density is about the given activity z. With the time span $\Delta t = 10^{-5}$, L-valued Markov chains X_n are generated successively for sufficiently many times. The transition from X_n to X_{n+1} is determined as follows:

(1) $\boldsymbol{P}\{X_{n+1} = X_n \backslash \{x\}\} = \Delta t$ for each $x \in X_n$ (i.e., the death rate $\equiv 1$),

(2) $\boldsymbol{P}\{X_{n+1} = X_n \cup \{x\}\} = z\Delta t$ for each x not within distance 1 from every point of X_n, (i.e., the birth rate $= z\psi(x; X_n)$) and

(3) otherwise $X_{n+1} = X_n$.

Finally, the sample spherical contact distribution function $\hat{F}(t)$ is calculated for various contact distances t. In order to avoid the edge effect, only points in $[20,80] \times [20,80]$ are used.

Theoretical approximations are computed based on the formula (4.1). Well-known numerical results for virial coefficients v_n, which are equal to $-(n-1)\beta_n/n!$, for hard disk Gibbsian process, give values $\beta_2, \ldots, \beta_7$ (see Ree and Hoover (1967)). We compute ε_n by multidimensional numerical integration for each t. Of course, $\varepsilon_1 = -\pi t^2$. The integral defining ε_n in Theorem 3.2 can be simplified further. Let us say that two graphs in $\mathscr{B}_{0n}$ are equivalent if they are the same up to a permutation of vertices $\{1, 2, \ldots, n\}$. Let $\mathscr{B}_{0n}^*$ be a set of representatives of this equivalent relation and $b(H)$ be the power of the equivalent class of H. Then

$$\varepsilon_n = \int_{\boldsymbol{R}^{2n}} \left[\sum_{H \in \mathscr{B}_{0n}^*} b(H)(-1)^{N(H)} \right] d(x)_n\,,$$

where $N(H)$ is the number of those edges $(0, j) \in H$ with $|x_j| < t$ and edges (i, j) with $|x_i - x_j| \leq 1$. Since, for example, $\#\mathscr{B}_{06} = 1{,}014{,}888$ and $\#\mathscr{B}_{06}^* =$ 2,278 the integrand is considerably simplified. Further, a careful choice of representatives makes the numerical evaluation more efficient. Finally, a Monte Carlo evaluation of this integral are done. Usually a Monte Carlo quadrature uses pseudo-random numbers, but we used quasi-random numbers. A sequence $\boldsymbol{x}_n = (x_n^{(1)}, \ldots, x_n^{(d)})$ of d-dimensional quasi-random numbers are generated by the formula $x_n^{(i)} = \{$the fractional part of $n x_0^{(i)}\}$

where $\boldsymbol{x}_0 = (x_0^{(1)}, \ldots, x_0^{(d)})$ is a set of irrational numbers. It is known that with appropriately chosen $\boldsymbol{x}_0$ the efficiency of the quadrature based on quasi-random numbers (the Haselgrove method) is far superior to that based on pseudo-random numbers, at least for sufficiently smooth integrands (see, e.g., Niederreiter (1978)). We took $x_0^{(i)} = 2^{i/(d+1)}$. Though our integrand is not smooth (a linear combination of indicator functions), a preliminary experiment affirmed this superiority. As a result, we got values ε_2, ε_3 and ε_4 numerically for $t = 0.1(0.1)5.0$ and computed values

$$\tilde{F}(t) = 1 - \exp\left(\sum_{k=1}^{4} \varepsilon_k \frac{z^k}{k!} \right).$$

The results are displayed in Fig. 1. The pair of curves $(\tilde{F}(t), \hat{F}(t))$ are shown for $z = 0.1$ (A, B), 0.2 (C, D) and 0.3 (E, F). Though the use of coefficients only up to fourth degree seems to rough, agreements are fairly good for $z = 0.05(0.05)0.25$ (in fact excellent for $z = 0.20$ and 0.25) but not for $z = 0.3$ and over. A common defect of the Mayer series (also of the activity expansion) is that it is convergent only for relatively small values of z. This is because mean values of Gibbsian processes usually have singularities as a function of z, which indicates the phenomenon called the *phase transition*, well-known to physicists and a characteristic property of Gibbsian processes. In fact, it is proven that the hard disk system has a phase transition somewhere between $1/\pi e$ and $1/\pi$ (probably nearer to $1/\pi = 0.318\ldots$), where the system makes the transition from the gas phase to the liquid phase (see Ruelle (1969), Chapter 4). In some cases, a Padé approximation, i.e., approximation by rational functions in z, to a Mayer series show a better agreement for higher values of z.

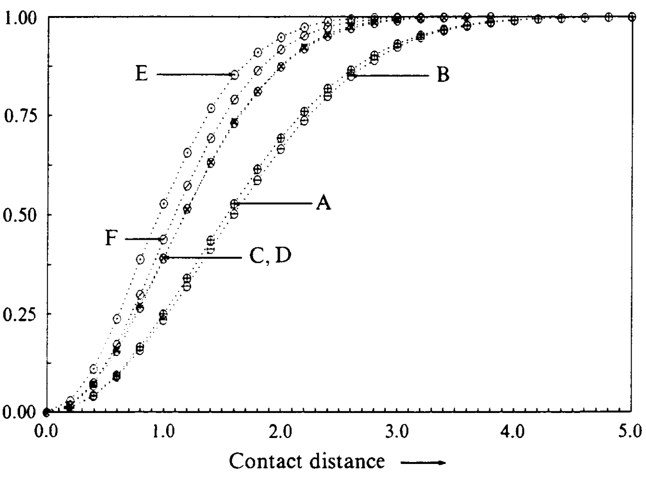

Fig. 1. Spherical contact distributions $(\tilde{F}(t), \hat{F}(t))$ for $z = 0.1$ (A, B), 0.2 (C, D) and 0.3 (E, F).

Acknowledgements

The author is very grateful to Dr. D. Stoyan and the referees for useful comments.

REFERENCES

Comtet, L. (1974). *Advanced Combinatorics*, Reidel, Dordrecht.

Diggle, P. J. (1983). *Statistical Analysis of Spatial Point Patterns*, Academic Press, New York.

Doob, J. L. (1953). *Stochastic Processes*, Wiley, New York.

Duneau, M., Soulliard, B. and Logolnitzer, D. (1975). Decay of correlations for infinite-range interactions, *J. Math. Phys.*, **16**, 1662–1666.

Grandy, W. T., Jr. (1987). *Foundations of Statistical Mechanics*, Vol. 1, *Equilibrium Theory*, Reidel, Dordrecht.

Harary, F. and Palmer, E. M. (1973). *Graphical Enumeration*, Academic Press, New York.

Hille, E. (1977). *Analytic Function Theory*, Vol. 1, 2, Chelsea, New York.

Minlos, R. A. and Pogosian, S. K. (1977). Estimates of ursell functions, group functions, and their derivatives, *Teoret. Mat. Fiz.*, **31**, 199–213.

Nguyen, X. X. and Zessin, H. (1979). Integral and differential characterizations of the Gibbs process, *Math. Nachr.*, **88**, 105–115.

Niederreiter, H. (1978). Quasi-Monte Carlo methods and pseudo-random numbers, *Bull. Amer. Math. Soc.*, **84**, 957–1041.

Ogata, Y. and Tanemura, M. (1984). Approximation of likelihood functions in estimating interaction potentials from spatial point patterns, *J. Roy. Statist. Soc. Ser. B*, **46**, 496–518.

Preston, C. (1976). *Random Fields*, Springer, New York.

Ree, F. H. and Hoover, W. G. (1967). Seventh virial coefficients for hard spheres and disks, *J. Chem. Phys.*, **46**, 4181–4197.

Ruelle, D. (1969). *Statistical Mechanics Rigorous Results*, Benjamin Inc., New York.

Saks, S. (1937). *Theory of the Integral*, Hafner Publishing Co., New York.

Stoyan, D., Kendall, W. S. and Mecke, J. (1987). *Stochastic Geometry and Its Applications*, Wiley, New York.

Uhlenbeck, G. E. and Ford, G. W. (1963). *Lectures in Statistical Mechanics*, American Mathematical Society, Providence, Rhode Island.

Westcott, M. (1982). Approximations to hard-core models and their application to statistical analysis, *J. Appl. Probab.*, **19A**, 281–292.