International Statistical Review (1998), **66**, 1, 1–27, Printed in Mexico

Random Sets: Models and Statistics

Dietrich Stoyan

Institut für Stochastik, TU Bergakademie, Freiberg, Germany

Summary

This paper surveys aspects of the theory of random closed sets, focussing on issues of practical and current interest. First, some historical remarks on this part of probability theory are made, where the important role of Georges Matheron is emphasized. Then, fundamental characteristics of the distribution of random closed sets are introduced. The very important Boolean model serves as an example for discussing mathematical and statistical problems. A number of further models is then considered, namely excursion sets of random fields, the system of edges of the Poisson Voronoi tessellation and various random systems of non-overlapping spheres. Finally, some ideas of particle statistics are presented, including some models of random compact sets.

Key words: Aumann mean; Boolean model; Contact distribution function; Covariance; Excursion set; Fibre process; Particle; Point process; Poisson hard core process; Poisson process; Random closed set; Random compact set; Sphere packing; Stienen model; Voronoi tessellation.

1 Introduction

Until the 1970s random sets were only a marginal or exotic part of probability theory. In some papers the terms "random set" or "region ... depending on chance" appeared, see e. g. Kolmogoroff (1933, 1946) and Robbins (1945, 1946). Furthermore, in a number of papers written mainly by non-mathematicians and widely scattered in the literature, particular models of random sets were studied and some formulae were derived that today are considered as fundamental. This was done without a rigorous theoretical base; there was no rigorous definition of random sets.

This situation has changed completely since the publication of the fundamental and seminal book by Matheron (1975). This book has laid the fundamentals of the theory of random closed sets, provided the suitable measure-theoretic machinery and offered the fundamental theorems. Since then random set theory became an inherent part of probability theory. Many mathematicians and statisticians have thus been encouraged to study problems of (and by means of) random sets. Various successful applications have led to systematisation, simplification, commonness and experience.

An early application of random compact sets appeared in the work of Hildenbrand (1974) on large economics. Therein, random sets were used for modelling the demand of sets of consumers. The Aumann mean (see section 9) was used for the characterization of their mean behaviour. Also in the theory of multivalued functions, ideas of the theory of random sets play an important role, see Salinetti & Wets (1986) and Aubin & Frankowska (1990). Such functions are used in the theory of optimal control and optimization. Chinese researchers have written books on set-valued martingales and stochastic processes in general, see Li (1994) and Zhang, Wang & Gao (1996). Random sets appear also in various statistical problems. An example is Gill, van der Laan & Robins (1997) who discuss statistics for grouped, censored and missing data, which all can be considered to be set-valued data. Random sets are also important in the theory of random fuzzy sets and in fuzzy data analysis, see Bandemer & Näther (1992), Goodman & Nguyen (1985), Kruse & Meyer (1987) and Wang

(1985). A further important application field of (still more potential than actual) of random set theory is particle statistics to be discussed in section 9.

Random sets (typically finite) are also a fundamental for general data fusion, see Goodman, Nguyen & Walker (1991) and Mahler (1993, 1994). The problem here is to determine the locations and identities of unknown targets on the basis of sensor data (obtained by radar, sonar, infrared and other sensors). Effects of clutter, jamming, measurement noise, false alarms, missed detections and target manoeuvres make random set methods applicable and useful.

Last but not least, random closed sets are a fundamental tool in solving problems of geometrical statistics and image analysis. This is quite close to the original ideas of Matheron and his colleagues. Today, the theory of random sets is the core of stochastic geometry. In this context, random sets are stochastic models of irregular or random geometrical structures. The perhaps most important field of application of random set methods in this sense is in materials science. In many cases the microstructure of materials can be described by random set models. This is important since engineers know from many experiments that there is a close relationship between the geometry of the microstructure and physical or technical bulk properties, while the geometrical-statistical analyses carried out today by image analysers are much easier to perform than physical experiments.

This paper surveys some of the ideas of modern random set theory. It emphasizes the geometrical aspects, according to this author's working field and experience. The main reference to random sets is still the famous book by Matheron (1975). There are many newer books dealing with special aspects of the theory; in particular the survey-like exposition in Stoyan, Kendall & Mecke (1995) and a book on limit theorems for unions of random sets (Molchanov, 1993) and one on Boolean model statistics (Molchanov, 1997a). However, a rigorous modern book covering the general theory including proofs is still missed.

2 Some Remarks on Historical Aspects

Random sets appear early in probability theory, in the context of geometrical probability. The random points, discs and segments (in Buffon's terms: random needles) are nothing else but random sets. All statisticians are familiar with particular random sets, namely confidence intervals.

Also in the famous book by Kolmogoroff (1933, 1946), the notion of random sets appears. The father of modern probability theory calculates the mean area of a bounded random set by means of the Fubini theorem. He writes (in the English translation of 1946):

"We shall give an application of this theorem to one example in geometric probability. Let G be a measurable region of the plane whose shape depends on chance; in other words, let us assign to every elementary event ξ of a field of probability a definite measurable region G. We shall denote by J the area of the region G, and by $P(x, y)$ the probability that the point (x, y) belongs to the region G. Then

$$\mathrm{E}(J) = \int\int P(x, y)\mathrm{d}x\mathrm{d}y \ . \ "$$

In modern notation, we would perhaps write

$$\mathrm{E}(A(G)) = \int\int \mathbf{1}_{G(\omega)}(x)\mathrm{d}x\mathrm{P}(\mathrm{d}\omega) = \int p(x)\mathrm{d}x \qquad (1)$$

where $p(x) = \mathrm{P}(x \in G), x \in \mathbb{R}^2$. The function $p(x)$ is today sometimes called the *covering function*.

Kolmogoroff defined, as we would expect, a random region as a mapping from a probability space into a suitable phase space. However, he did not introduce one of the necessary measure-theoretic ingredients, the σ-algebra of the state space. He also did not consider the problem of measurability of the function p. In Robbins (1944, 1945) formula (1) is discussed again, and therefore it is sometimes called Robbins' formula. Robbins simply *supposed* that p is measurable. When the state space is the

system of all measurable bounded subsets of the given space (e.g. $\mathbb{R}^2$), a suitable σ-algebra seems to be unknown even today. Modern mathematicians decided to be modest and to consider random *closed* sets. (Sometimes also open sets appear, see e.g. Zähle, 1984.) In the case considered by Kolmogoroff a random *compact* set would probably be sufficient. In this case, a suitable σ-algebra is the Borel σ-algebra corresponding to the Hausdorff metric in the system $\mathbb{K}$ of all compact subsets of $\mathbb{R}^2$. By the way, both concepts, that of Borel σ-algebra and that of Hausdorff metric, were already known in 1933.

Today we know that a convenient σ-algebra for random closed sets is *Matheron's σ-algebra $\mathcal{F}$*. The theory was developed in the 1960s and 1970s, see Matheron (1967, 1975). When I met Georges Matheron on October 9, 1996, in Fontainebleau during a meeting celebrating his work, he made it clear that he had developed the theory of random closed sets completely independently of the work of others. He had had the idea of hit-and-miss events already in the 1960s and he learned later that he had in fact studied capacities in the sense of Choquet (1953, 1954); also Kendall (1974) had similar ideas about the fundamentals of the theory of random sets. Nevertheless, Matheron prefers to call the hit-and-miss topology "Choquet topology"; we should call it *Choquet-Matheron topology*.

The main idea of Matheron's theory is to characterize the distribution of a random closed set not by measures or contents but instead by hit-or-miss events, by the events of intersection with test sets. Today, random measures are viewed as being structures associated to or accompanying random sets. Examples of such measures are the volume measure, V_Ξ, defined by

$$V_\Xi(B) = \nu_d(\Xi \cap B) \qquad \text{for Borel } B\,,$$

and so-called curvature measures, where ν_d denotes the Lebesgue measure on $\mathbb{R}^d$. Two examples of such measures are given in section 5.

Of course, any σ-algebra for random closed sets must have as elements sets of closed subsets of the space E considered, which here (and in the following) is taken to be the Euclidean space $\mathbb{R}^d$. Let $\mathbb{F}$ be the system of all closed subsets of E. Particular subsets of $\mathbb{F}$ are those sets $\mathbb{F}_K$ that contain all closed sets intersecting a given compact "test set" K,

$$\mathbb{F}_K = \{F \in \mathbb{F} : F \cap K \neq \emptyset\}\,.$$

The system of all $\mathbb{F}_K$ with $K \in \mathbb{K}$ generates the σ-algebra $\mathcal{F}$. In the case of random closed sets that are almost surely compact, the appropriate σ-algebra $\mathcal{F} \cap \mathbb{K}$ is equivalent to the Borel σ-algebra corresponding to the Hausdorff metric.

Matheron (1975) has shown that the usual set-theoretic operations and real-valued characteristics of sets are measurable with respect to $\mathcal{F}$; in particular, he showed that the covering function p is measurable.

3 Characterizing the Distribution of a Random Closed Set

We all know that the distribution of a (real-valued) random variable X is given by its distribution function,

$$F(x) = \mathrm{P}(X \leq x) \qquad \text{for } x \text{ real}\,.$$

This is related to one of the generators of the Borel σ-algebra on the real axis. In analogy, the distribution of a random closed set Ξ is given by the probabilities

$$T(K) = \mathrm{P}(\Xi \cap K \neq \emptyset) \qquad \text{for } K \in \mathbb{K}\,.$$

Sometimes, T is called the *capacity functional*, because T is a so-called alterating capacity of infinite order in the sense of Choquet (1953, 1954).

It is easy to show that invariance properties of random closed sets can be expressed in terms of

the capacity functional. A random closed set Ξ is called *stationary* if its distribution is translation invariant, i.e. Ξ and $\Xi_h = \{y : y = x + h, x \in \Xi\}$ have the same distribution for all $h \in E$. This is equivalent to

$$T(K) = T(K + h)$$

holding for all $K \in \mathbb{K}$ and all $h \in E$. Isotropy is an analogous property connected with rotations around the origin. Also here, a characterization by means of the capacity functional is possible.

Later work has shown the following interesting facts:

(a) The capacity functional also determines weak convergence of random closed sets, see Norberg (1984), ergodicity and mixing, see Heinrich (1992), and can be used to characterize the comparability of random closed sets (Norberg, 1992).

(b) In particular cases, also other probabilities characterize the distribution of random sets. An example is the case of random convex compact sets. Here it suffices to consider the probabilities $P(\Xi \subset K)$ for K convex and compact. Sometimes this functional is called *belief measure*. (For a finite space E also the *mass-assigned function* $P(\Xi = S)$ is used, which is defined on the system of all finite subsets of E. It clearly determines the distribution of Ξ.)

The system $\mathbb{K}$ is very large so that in non-trivial cases it is hopeless to specify $T(K)$ for all $K \in \mathbb{K}$. A natural idea is to consider only K's belonging to subsets of $\mathbb{K}$. In particular, families of compact sets K which are parametrized (preferably) by real parameters are used. This leads to functions that can be presented graphically. Such families of compact subsets of E are e.g. the systems of

all singletons x for $x \in E$,
all pairs x, y for x and y in E,
all triplets x, y, z for x, y and z in E,
all segments and
all spheres.

The first case corresponds to the covering function defined by

$$p(x) = P(x \in \Xi) .$$

If the random closed set is stationary, then $p(x)$ is a constant, which here is denoted by p. This value can be interpreted as the *volume fraction*, i.e.

$$p = E(v_d(\Xi \cap U))$$

where $U = [0, 1]^d$. For a compact random set, the function $p(x)$ can be considered as a kind of mean or expectation of Ξ.

The function C given by

$$C(x, y) = P(x \in \Xi, y \in \Xi)$$

is called *covariance* (not covariance *function*). In the stationary and isotropic case, the covariance only depends on the distance between the two points x and y.

A generalisation to three or more points is natural — however, closed-form expressions for higher-order covariances are known only for very special models. Third-order covariances play an important role in the study of physical bulk properties of two-phase random materials, see Jeulin (1998) and Torquato (1991). For some particular Boolean models, formulae for probabilities of the form

$$P(x_1 \in \Xi, x_2 \in \Xi, x_3 \in \Xi)$$

can be given, namely in the planar case for circular, rectangular and Poisson polygonal grains.

In the cases where K is the ball of radius r centred at the origin o, $K = b(o, r)$, or the segment of length r with one endpoint in the origin, $K = s(o, r)$, for a stationary random closed set Ξ with positive volume fraction typically functions such as $P(\Xi \cap b(x, r) \neq \emptyset)$ are not used. Rather, it is more natural to consider the intersections of the ball with Ξ when the centre does not belong to Ξ. The functions

$$H_s(r) = P(\Xi \cap b(o, r) \neq \emptyset \mid o \notin \Xi)$$

$$H_l(r) = P(\Xi \cap s(o, r) \neq \emptyset \mid o \notin \Xi) \qquad \text{for } r \geq 0$$

are called *spherical* and *linear* contact distribution functions. In the anisotropic case H_l depends on the direction of the segment $s(o, r)$. In practically interesting cases, these functions are indeed distribution functions. They have a nice geometrical interpretation: $H_s(r)$ is the distribution of the random distance from o to the nearest point of the random set Ξ, under the condition that the origin o is not in Ξ. Analogously, H_l belongs to the directional distance from o to Ξ, where the direction is given by the segment direction.

4 The Poisson Process

As is typical in applied probability, specific stochastic models play an important role also in the context of random sets. Three such models in the stationary case, for infinitely extended structures, are (i) the Boolean model, (ii) the Stienen model and (iii) the system of edges or faces of the Poisson Voronoi tessellation. All these models are constructed by means of the stationary (or homogeneous) Poisson (point) process. Generalisations to other point processes are possible; such models can still be simulated, but it seems to be hopeless to derive formulae in interesting cases. Sometimes also non-stationary variants of these models are of interest.

The homogeneous Poisson process is a point process with completely randomly scattered points, see Kingman (1993) and Stoyan, Kendall & Mecke (1995). It is a stationary and isotropic point process, with a distribution depending only on one parameter λ. The random number of points in a bounded Borel set A has a Poisson distribution with parameter $\mu = \lambda v_d(A)$. Since the mean of the Poisson distribution coincides with its parameter, the meaning of λ is clear: it is the mean number of points per unit volume. The set of all points of the Poisson process is a stationary and isotropic random closed set.

5 The Boolean Model

5.1 Formulae

The *Boolean model* is the perhaps most famous and most frequently used random set model. It is a mathematically rigorous formulation of the idea of an infinite system of randomly scattered particles. So it is a fundamental model for geometrical probability and stochastic geometry. The Boolean model has a long history; the first papers on the Boolean model appeared in the beginning of the 20th century, see the references in Stoyan, Kendall & Mecke (1995). The name "Boolean model" was coined in Matheron's school in Fontainebleau to discriminate this set-theoretic model from (other) random fields appearing in geostatistical applications.

The Boolean model is constructed from two components: a system of grains and a system of germs. The germs are the points $x_1, x_2, \ldots$ of a Poisson process of intensity λ. The grains form a sequence K_n of i.i.d. random compact sets. Typical examples are spheres, discs, segments and Poisson polyhedra (see Section 9). A further random compact set K_0 having the same distribution as the K_n is sometimes called the 'typical grain'.

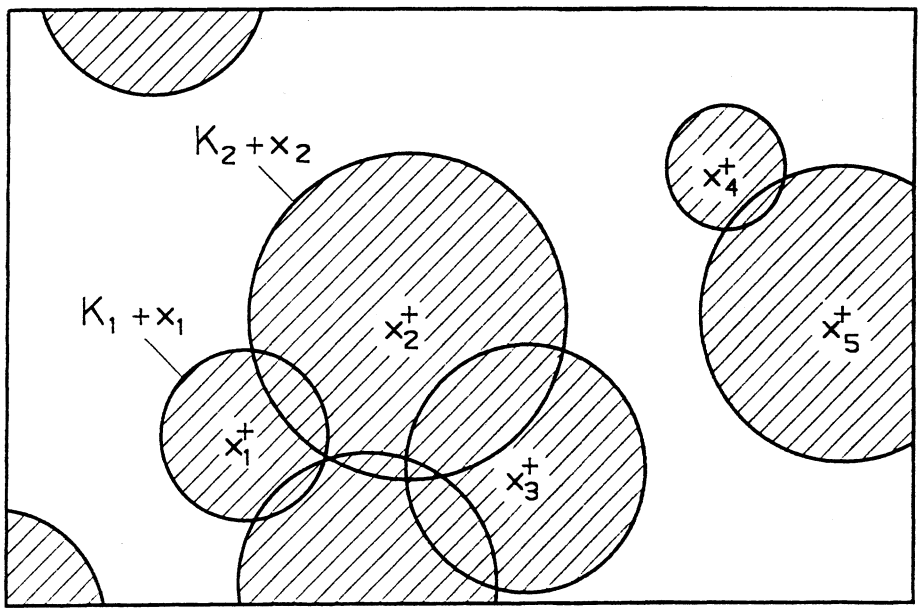

Figure 1. *A part of a Boolean model with germs and grains.*

The Boolean model Ξ is the union of all grains shifted to the germs,

$$\Xi = \bigcup_{n=1}^{\infty} (K_n + x_n) .$$

Figure 1 shows schematically a part of a Boolean model with some germs and translated grains, while Figure 2 gives a typical simulated sample for the case of circular grains.

In the following it is always assumed that the typical grain K_0 is convex; only Boolean models with convex grains are discussed henceforth. But this does not mean that non-convex grains are unimportant. For example, the case where K_0 is a finite point set corresponds to Poisson cluster processes.

The parameters of a Boolean model are the intensity λ and parameters characterizing the typical grain K_0. While for simulations the complete distribution of K_0 is necessary, for a statistical description it often suffices to know that the basic assumption of a Boolean model is acceptable and to have some parameters such as mean area $\overline{A}$, mean perimeter $\overline{U}$ or, if a set-theoretic characterization is needed, the Aumann mean $\mathrm{E}(K_0)$ of K_0; see section 9 for an explanation.

The capacity functional of the Boolean model Ξ is given by the simple formula

$$\mathrm{P}(\Xi \cap K \neq \emptyset) = 1 - \exp(-\lambda \mathrm{E}(\nu_d(K_0 \oplus \check{K}))) \qquad \text{for } K \in \mathbb{K} .$$

Here $\oplus$ denotes Minkowski addition, $A \oplus B = \{a + b : a \in A, b \in B\}$, and $\check{K}$ is the set $\{-k : k \in K\}$. The derivation of this formula is given in Matheron (1975) and Stoyan, Kendall & Mecke (1995). Its structure is quite similar to the emptiness probability of the Poisson process or to the probability that a Poisson random variable does not vanish. It can perhaps be partially explained when applied to the particular case $K_0 = \{o\}$. Then, the Boolean model is nothing else but the

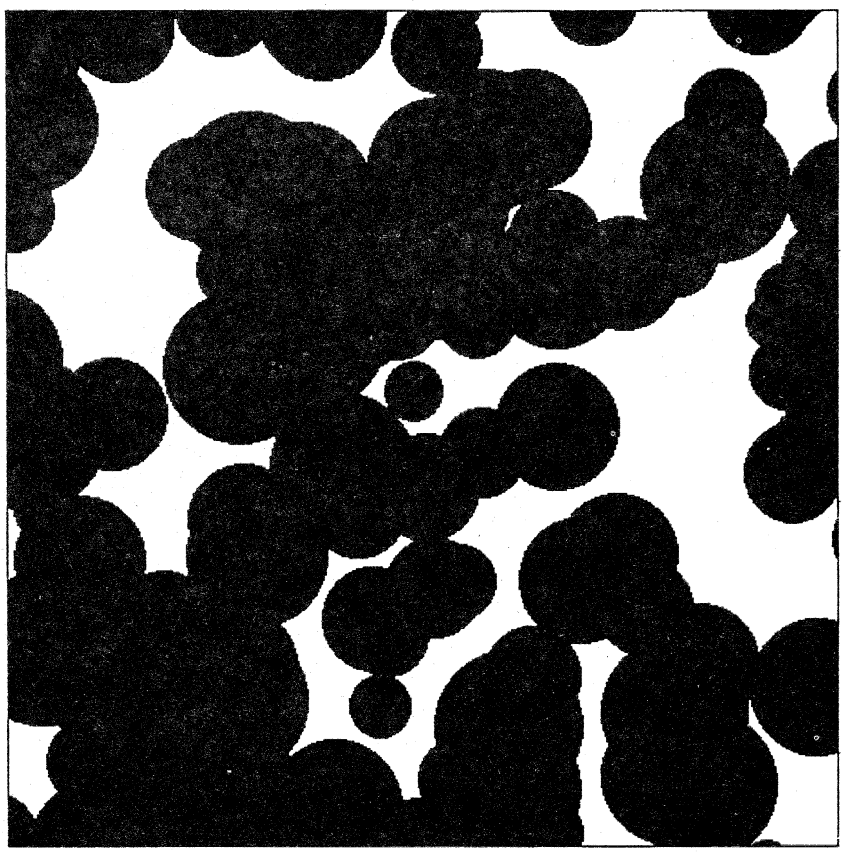

Figure 2. *A simulated sample of a Boolean model with circular grains.*

random set consisting of all points of the Poisson process of germs. Consequently,

$$P(\Xi \cap K \neq \emptyset) = 1 - \exp(-\lambda \nu_d(K)) ,$$

which coincides with the general formula for $K_0 = \{o\}$.

The calculation of the capacity functional of a Boolean model poses a non-trivial geometrical problem, viz. the determination of the mean

$$E(\nu_d(K_0 \oplus \check{K})) .$$

Here, integral geometry helps if K is convex. If K_0 is a ball, then the classical Steiner formula gives the result. If K_0 is not spherical but isotropic (with distribution invariant with respect to rotations around the origin o), a generalisation of this formula by Matheron leads to a formula in which the so-called Minkowski functionals appear. For example, in the three-dimensional case it is

$$E(\nu_3(K_0 \oplus \check{K})) = \overline{V} + 2\overline{b}(K)\overline{S} + 2\overline{\overline{b}}S(K) + V(K)$$

where $\overline{b}(K)$, $S(K)$ and $V(K)$ are average breadth, surface area and volume of K, and $\overline{\overline{b}}$, $\overline{S}$ and $\overline{V}$ are the corresponding means of K_0.

For non-convex K, the numerical determination of $T(K)$ is rather difficult. Already the numerical determination of the covariance $C(r)$, which belongs to the case where the set K is a pair of points of distance r, is a difficult problem unless K_0 is a ball or Poisson polyhedron. (The Poisson polyhedron

is the typical cell of the stationary Poisson hyperplane tessellation of the space, see Stoyan, Kendall & Mecke, 1995.)

By means of the formula for the capacity functional and the generalised Steiner formula it is easy to give formulae for the spherical and linear contact distributions. The spherical contact distribution function satisfies

$$H_s(r) = 1 - \exp(-\lambda \sum_{k=1}^{d} \binom{d}{k} \overline{W}_k r^k) \qquad \text{for } r \geq 0$$

where the $\overline{W}_k$ are the expectations of the Minkowski functionals of K_0. The linear contact distribution function H_l is an exponential distribution function with parameter $\lambda \frac{b_{d-1}}{b_d} \overline{W}_1$, where b_k is the volume of the unit sphere of $\mathbb{R}^k$.

5.2 *Statistics*

In this section, spatial statistical analysis for a stationary random set is discussed by considering the particular example of the planar Boolean model. Three methods are considered: the minimum contrast method, the method of intensities and the method of accompanying processes, restricting attention to the important particular planar case. For tests of the distributional hypothesis see Stoyan, Kendall & Mecke (1995) and Molchanov (1997a). Molchanov (1997b) is a good survey on statistics for random sets in general.

The idea of the *minimum contrast method* is to consider a contrast functional of the type

$$\Delta(\theta) = \int_a^b (\hat{f}(x) - f(x, \theta))^2 \mathrm{d}x \ .$$

Here $f(\cdot, \theta)$ is a function describing a characteristic of the Boolean model depending on some parameter θ; $\hat{f}$ is the empirical counterpart of this function estimated from the data. The idea is to minimize $\Delta(\theta)$ over the set of parameters θ leading to an optimal choice $\hat{\theta}$, which is called the minimum contrast estimator for θ. For f the covariance C can be chosen, see Diggle (1981). This is helpful probably only in the cases of spherical or Poisson polygonal grains, since otherwise the formulae for C are very complicated. A more practical choice for f is a contact distribution function as in Serra (1982). Heinrich (1993) gave a mathematical foundation of this estimation method applied to random sets and investigated the asymptotic behaviour of minimum-contrast estimators when the sampling window becomes large.

Quite natural is the *method of intensities*, which is a variant of the moment method of statistics. It is based on formulae as the following three ones:

$$A_A = 1 - \exp(-\lambda \overline{A}) \ , \tag{2}$$

$$L_A = \lambda(1 - A_A)\overline{U} \ , \tag{3}$$

$$N_A^+ = \lambda(1 - A_A) \ . \tag{4}$$

Here A_A is the area fraction, i.e. the same as p in section 3. L_A denotes the specific boundary length, i.e. the mean boundary length of the Boolean model per unit area. N_A^+ is the intensity of a point process, namely that of lower convex tangent points. Those are tangent points of the Boolean model corresponding to horizontal tangent points with the property that the neighbourhood of the points in the Boolean model is *above* the tangent lines. Finally, $\overline{A}$ and $\overline{U}$ are mean area and perimeter of the typical grain K_0.

Sometimes the last equation is replaced by an equation for the specific connectivity (or Euler–Poincaré) number. In the context of the specific connectivity number, the idea of tangent point

counting was developed, see Cahn (1967), DeHoff (1967) and Haas, Matheron & Serra (1967); for the Boolean model, convexity points were first considered by Matheron (1975).

The method of intensities has a long history. For the isotropic case it first appeared in Santaló (1976), p. 284, and then in Kellerer (1983), who used the specific connectivity number. In the form given here it was used in Bindrich & Stoyan (1990), while the stationary (non-isotropic) case was considered in Weil (1988).

Already the method of intensities uses information (namely L_A and N_A^+) of *accompanying* or *associated stochastic processes*. These are the fibre process of boundary curves of the Boolean model and the point process of lower convex tangent points. Both processes contain so much information that they can be used to estimate the intensity λ and the distribution of the typical grain K_0; in some sense they can be considered to be "sufficient statistics". This is clear for the fibre process. For the point process of lower convex tangent points, this was shown by Molchanov (1995); this point process plus the covariance are sufficient to determine those characteristics. The associated processes are described by the usual parameters for such processes, which are re-expressed in terms of the characteristics of the Boolean model. Estimators of the process characteristics are then transformed into estimators for parameters or characteristics of the Boolean model.

For the point process of convex tangent points, the intensity N_A^+ satisfies equation (4). This leads to a natural and elegant estimator for the intensity λ, namely

$$\hat{\lambda} = \frac{\hat{N}_A^+}{1 - \hat{p}}\,, \tag{5}$$

where $\hat{N}_A^+$ and $\hat{p}$ are the natural estimators of N_A^+ and p, namely the number of lower convex tangent points in the window W and the area of Ξ in W, both divided by the window area $A(W)$. The quantity $A(W)^{1/2}(\hat{\lambda} - \lambda)$ follows for large W a Gaussian distribution with zero mean and variance $\lambda/(1 - p)$, see Molchanov & Stoyan (1994).

The estimator given by (5) is consistent, but it is not unbiased. Still worse, its expectation is infinite! This shows that statistics for the Boolean model is a really non-trivial problem.

The tangent point approach yields still more. In Molchanov & Stoyan (1994), also the second order product density of the point process of convex tangent points is given. It depends on the shape and size of the typical grain. This offers a possibility to determine grain characteristics by analysis of this point process. Heinrich & Werner (1998) studied in detail the asymptotic properties of kernel-type estimators for the diameter distribution of spherical grains as suggested by Molchanov (1995). Molchanov (1995) suggested also to study marked point processes corresponding to m different tangent directions (not only the horizontal one which leads to N_A^+). Clearly, with increasing m the marked point approach becomes similar to the fibre process approach.

The fibre process approach has to date not been completely exploited. Typically, only one-dimensional distributions have been considered: the distribution of the normal direction and the distribution of the curvature in a 'typical' fibre point (i.e. in a 'typical' boundary point of the Boolean model). The directional information can be obtained in several ways, namely by linear contact distributions corresponding to different directions or by the rose of intersections (the intensities of the point processes of intersection points of test lines and fibres), see Molchanov (1995, 1996a) and Weil (1995, 1997). Then, equations such as

$$L_A(\Gamma) = \lambda(1 - p)\mathrm{E}L(K_0, \Gamma)$$

are used. Here $L_A(\Gamma)$ is the length density of the sub-fibre process of fibre pieces with normal directions in $\Gamma \subset [0, \pi]$. $\mathrm{E}L(K_0, \Gamma)$ is the mean of the boundary length of K_0 where the normal direction is in Γ.

Since

$$\mathrm{E}(L(K_0, \Gamma)) = L(\mathrm{E}(K_0), \Gamma)\,,$$

Figure 3. *A simulated sample of a Boolean model. The author believes that simple inspection of the figure suffices to determine the typical grain.*

$L(\mathrm{E}(K_0), \Gamma)$ can be estimated. If for a planar convex set K the measure $L(K, \Gamma)$ is given, K can be reconstructed, of course only up to location. Thus $\mathrm{E}(K_0)$ can be estimated, see Weil (1995). However, this whole procedure is rather complicated and probably numerically unstable. Furthermore, in the important isotropic case the directional approach fails completely. Then the directional distributions are uniform and $\mathrm{E}(K_0)$ is a disc. The directional approach tells much about directions, but in case of, or close to, isotropy, it tells little on shape.

Other typical characteristics of fibres such as their curvatures have been considered in Hall (1988), p. 321–333, Jeulin (1998) and Molchanov & Heinrich (1995). Jeulin's book contains results which were already applied in a statistical study of powders, while the latter paper gives also extensions to general germ-grain models and contains an asymptotic analysis of empirical curvature measures under suitable mixing conditions. It is possible to determine the distribution of curvature in a typical point of the fibre process. That is of course the same as the curvature distribution function in a typical point of the boundary of the typical grain K_0. This leads to an estimation method for the diameter distribution in the case of circular grains.

In the author's opinion, both essentially one-dimensional approaches use too far-going forms of condensation of data. Perhaps, certain higher-dimensional distributions are more suitable for statistics of the Boolean model. Also, pair correlation functions of the fibre process or the curvature measure may be useful in this context.

However, consider Figure 3, which shows a sample of an isotropic Boolean model with quadratic grains. Many of the readers probably are able to reconstruct the typical grain from the image: It is

obvious that the grains are squares, and some of the square sides are completely visible. In contrast, the Aumann mean of that typical grain is a disc and the best that can be produced in this situation by the directional approaches of Molchanov and Weil is the radius of the Aumann mean disc. The curvature approach would produce nonsense. (The case of isotropic rectangles can be treated by means of the tangent point approach, see Molchanov, 1995.) Thus a qualitatively new approach is necessary which can use geometrical information going farther than merely to the neighbourhood of boundary points. It is a difficult question whether a naive approach as that for the squares can be improved and formalized.

6 Further Random Set Models

The last few years have seen the creation of new random set models which are flexible and more amenable than the Boolean model, in which random sets are constructed by randomly scattered independent grains. These models are designed to have some particular geometrical properties, such as, for example, some degree of connectivity or invariance with respect to morphological transformations. They are constructed using the Gibbs field approach, either as marked point processes or as lattice-based sets.

There is an interesting form of generalisation of the Boolean model, namely the *weighted Boolean model* of Kendall (1997), which is closely related to *area-* and *quermass weighted point processes*, see Baddeley & van Lieshout (1995) and Baddeley, Kendall & van Lieshout (1998). In its simplest form a weighted Boolean model is defined as follows. It is given in a bounded window W, and its grains are discs of radius r. The germ process $\boldsymbol{x} = \{x_1, x_2, \dots\}$ is not a Poisson point process but has probability density f with respect to a Poisson process of intensity β (restricted to W) given by

$$f(\boldsymbol{x}) = \frac{\gamma^{-\nu_2(\bigcup\{b(x_i,r):\, x_i \in \boldsymbol{x}\})}}{Z(\beta,\gamma)},$$

where $Z(\beta, \gamma)$ is the normalizing constant. For $\gamma > 1$ this model was already introduced by Widom & Rowlinson (1970) in statistical physics as the "penetrable sphere model". This model can be generalised by considering densities with respect to a Poisson process independently marked by radii with a given distribution function. Then the density function f is a function both of points and radii. In the corresponding weighted Boolean model, the radius distribution may be different from the original mark distribution and the radii may even depend on location. Even more general models of this type have been considered in Møller (1996), where yet another particular example is the continuum random cluster model (Klein, 1982, Møller, 1994b, and Häggström, van Lieshout & Møller, 1996) obtained by replacing the area of the random set (given by the union of discs in the density f) by the number of maximal connected components in this random set.

Such models can be simulated in an elegant way (Kendall, 1996, Häggström, van Lieshout & Møller, 1996, and Kendall & Møller, 1997) and thus can be adapted to empirical structures. Using a specification in the sense of Preston (1976), it is possible to define analogous stationary models in $\mathbb{R}^d$. A simple particular case is the model of inpenetrable spheres studied in Mase (1985) and Stoyan (1989) and in the theory of marked Gibbs processes, see Goulard, Särkkä & Grabarnik (1996).

Similar models are lattice-based sets constructed by random fields. Some of them can be interpreted as Gibbs point processes on lattices. A classical model of this type is the Ising model. In order to avoid problems with identifiability of the germs in germ-grain models and to incorporate knowledge about the connected components in the random set, Møller & Waagepetersen (1998) proposed to use a new class of discretized random fields called *Markov connected component fields*; indeed the Ising model and many other Markov random field models are particular examples of such models.

Yet another new idea is offered by the *morphologically constrained random sets* of Sivarkumar & Goutsias (1997). These sets have the property that they are close to their openings or closings with

suitable structuring sets.

A further Poisson-process-based model is Matheron's *dead leaves model*, see Jeulin (1997). It produces tessellations of a nature quite different from Voronoi or line tessellations. It is helpful in the analysis of SEM images of powders and particle systems.

Finally, a large class of models must be mentioned, which still seem to be too difficult for some of the existing random set methods. These are *excursion sets* of random fields, i.e. sets defined by $\{x \in \mathbb{R}^d : Z(x) \geq z\}$ for a random field Z and a real number z, see Adler (1982). For usual volume fractions they produce patterns different to Boolean models, while for very high thresholds (and small volume fractions) they may have distributions similar to those of Boolean models. Since such models sometimes produce patterns similar to empirical ones, they are of some practical interest. For such models the covariance C can be calculated, see Cramér & Leadbetter (1967). But to date, only approximations of the linear contact distribution for some smooth Gaussian random fields can be given, see Nott & Wilson (1997a). These authors also study the problem of estimating parameters of the underlying random field $Z(x)$, $x \in R^d$, see Nott & Wilson (1997b) and some references therein.

7 Edges of Voronoi Tessellation

The Boolean model was also used in the context of modelling of crystallisation processes, probably first by Kolmogorov (1937). Consider a Poisson process system of germ points in which radial growth starts at time $t = 0$. The speed of growth α is the same for all germs. So at time t, each germ becomes a sphere of radius αt if enough space is available. The volume fraction of the corresponding Boolean model is

$$V_V(t) = 1 - \exp\left(-\lambda \frac{4}{3}\pi\alpha^3 t^3\right) \quad \text{for } t \geq 0\,. \tag{6}$$

One can assume that growth is stopped where two spheres come into contact, while otherwise growth into empty space still is possible. Figure 4 shows some steps of this growth process. Equation 6 and some generalisations are called Johnson–Mehl–Avrami equations.

By the way, several authors have studied the empty space regions for very large t, when the volume fraction of the Boolean model is close to one. Hall (1988) and Aldous (1989) have shown that the distribution of a typical empty region is asymptotically the same as that of a Poisson polygon (or polyhedron). This is the typical cell of a tessellation generated by a system of random lines (or planes) called Poisson line (plane) process, see also Section 9. A generalisation of Hall's result can be found in Molchanov (1996b). Chiu (1995) studied the probability of complete coverage of cubes. A paper by Erhardsson (1996) adds to the body of evidence that the union of all uncovered regions has a distribution similar to that of a Boolean model with Poisson polygonal grains. This may explain why such models turned out to be good models for systems of pores, see Serra (1982). (There are also other arguments for the use of this particular Boolean model, namely its polygonal nature and the form of the corresponding covariance.)

The growth process leads in a natural way to the Poisson Voronoi tessellation, see Okabe, Boots & Sugihara (1992), Møller (1994a), and Stoyan, Kendall & Mecke (1995). The cells of this tessellation are convex sets that are assigned to the corresponding germ points. They consist of those points which are closer to a given germ point than to all other germ points. In this context, the system of faces (in the spatial case) or edges (in the planar case) of the tessellation is considered as a random closed set. It is of course a stationary and isotropic random set, but the volume fraction is zero and thus also the covariance vanishes. However, the contact distributions are interesting and can be determined analytically. The linear contact distribution was already obtained by Gilbert (1962) (who did not yet know the name), while the spherical contact distribution was given in Muche & Stoyan (1992). The determination of these distributions is based essentially on geometrical properties of the Voronoi tessellation. The complement of the probability appearing in the definition of the contact

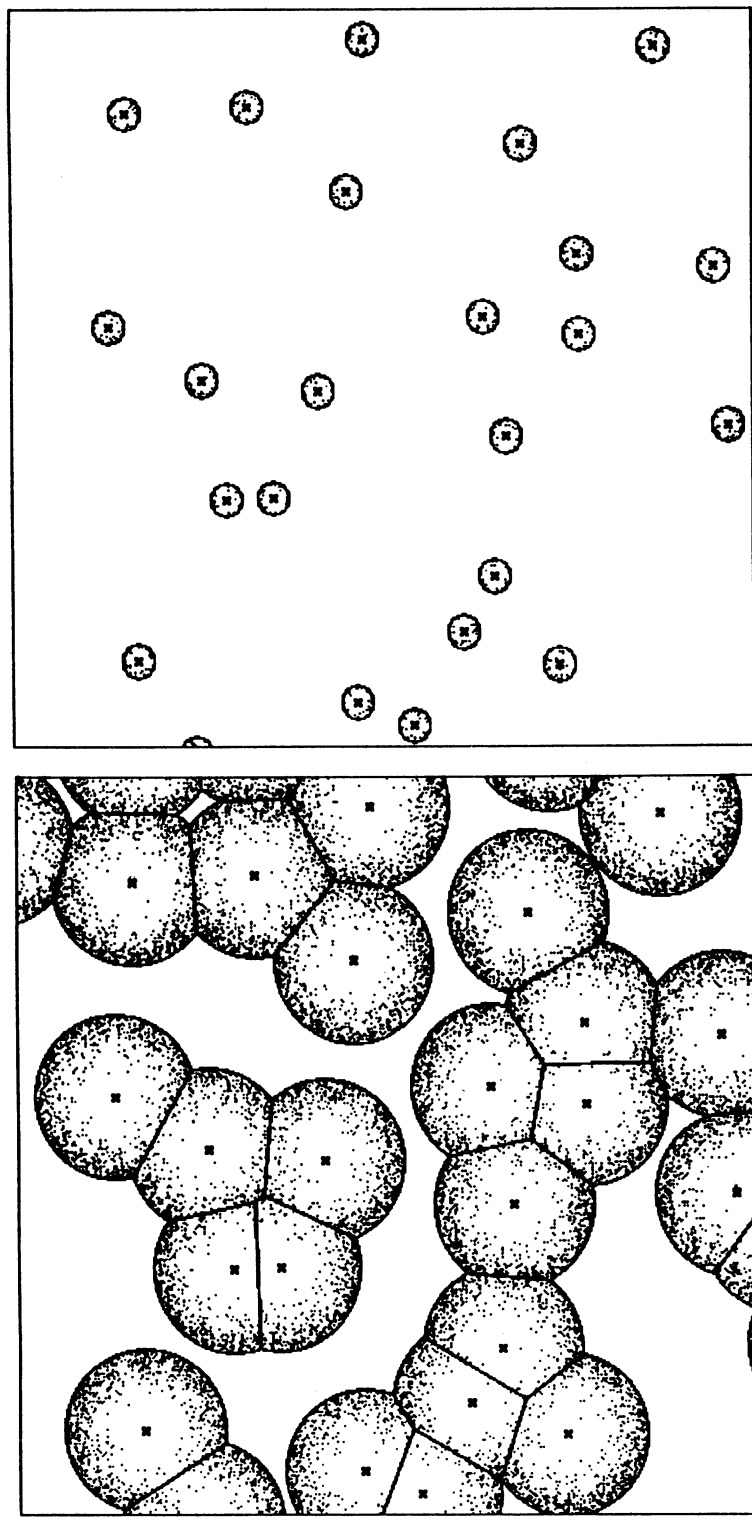

Figure 4. *Three steps of the growth process which lead eventually to the Voronoi tessellation.*

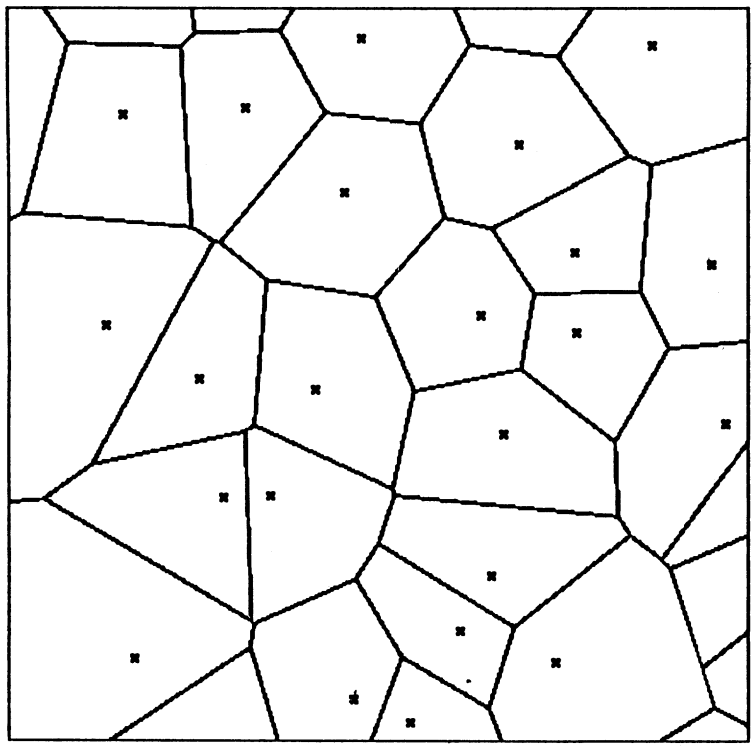

Figure 4 (continued). *Three steps of the growth process which lead eventually to the Voronoi tessellation.*

distribution is equal to the probability that the origin cell (that cell of the tessellation which contains the origin o) contains the test set $b(o, r)$ and $s(o, r)$, respectively.

In Muche & Stoyan (1992) also a relation between the linear contact distribution function H_l and the chord length distribution function was exploited. This has led to numerically useful expressions for the chord length density function of the Voronoi tessellation. (Here the chords are just the chords of the cells generated by intersection of the tessellation with a deterministic or independent random line.) The chord length density functions depend on the dimension d of the space. The functions for $d \leq 4$ and $d > 4$ differ qualitatively: while for the smaller d's the function takes its maximum for positive distances, for the greater d's the maximum is at 0. In some sense this shows the increasing variability of the Voronoi tessellation with increasing space dimension d. For more general Voronoi tessellations generated by (non-Poisson) stationary point processes, chord length and contact distribution functions were calculated in Heinrich (1996).

The crystallisation process mentioned in the beginning of this section can be generalised. The speed of growth may be time-dependent, the germs may appear subsequently according to a Poisson process with time-dependent nucleation rate. This process leads eventually to the Johnson–Mehl tessellation, the cells of which are not necessarily convex, see Møller (1992). The time-dependent volume fraction can still be obtained by means of the Boolean model formulae. It is a base for statistical estimation of the nucleation and growth rates, see, for example, Capasso, Micheletti, De Giosa & Mininni (1996). Other forms of generalisation start with non-Poisson point processes, see Uebele & Hermann (1996), and use growth velocities depending on local configurations, see Kelton (1993).

8 The Stienen Model and a Series of Random Systems of Non-overlapping Spheres

The *Stienen model* was originally suggested by the Dutch engineer J. Stienen in his PhD dissertation. It found the interest of the author and his colleagues since (i) it is a random closed set which is a union of non-overlapping spheres and (ii) some of its distributional characteristics can be calculated analytically. The following explanation of the model may help to understand that there are indeed situations in physics and materials science in which the model may be helpful. (In general, random systems of non-overlapping spheres are very important in both of these areas.) Take a Poisson process of germ points and start a uniform radial growth process in each of the points. The growth process of a point stops when the corresponding sphere has first contact with one of the faces of the Voronoi cell corresponding to that germ point. This rule can be interpreted physically. The sphere growth may result from diffusion of material to the germ points which was originally homogeneously distributed in space. The space in each Voronoi cell in some sense defines the amount of material available for the corresponding germ point. Another, more mathematical, definition is as follows. Take again the points of a planar Poisson process. Every point is the centre of a sphere the diameter of which is the distance to the nearest neighbour in the point process. Both definitions make it easy to see that the spheres do not overlap.

For the Stienen model the volume fraction p is easy to obtain, by means of the formula

$$p = \lambda \overline{V}$$

where $\overline{V}$ is the mean volume of the 'typical' Stienen sphere. Clearly the same formula is true for all stationary systems of non-overlapping grains, but in the case of the Stienen model it is easy to calculate the mean volume. This uses the fact that the diameter of the typical sphere has the same distribution as the nearest neighbour distance of the Poisson point process, which is

$$D(r) = 1 - \exp(-\lambda b_d r^d) \qquad \text{for } r \geq 0\,,$$

where b_d is the volume of the unit sphere in $\mathbb{R}^d$. This leads to $\overline{V} = (1/2)^d/\lambda$ and $p = (1/2)^d$. Also the analytical determination of the covariance is possible. This is shown in the paper Schlather & Stoyan (1997), where complicated techniques of the theory of marked point processes are applied. While in the one-dimensional case an analytical formula can be given, for higher dimensions this seems to be hopeless. Complicated integrals appear that are tractable only with numerical methods. For certain values of r the covariance $C(r)$ takes values which are (a little) smaller than p^2, which is the value of $C(r)$ for $r \to \infty$. For a Boolean model with convex grains, $C(r)$ is always greater than p^2.

Other set-theoretic characteristics, such as the linear and spherical contact distributions, are still unknown for this model.

The intersection of a three-dimensional Stienen model with a deterministic plane yields a system of circles of intersection. Clearly also these circles do not mutually intersect one another. It is possible to give the distribution function of the diameters of these circles and to determine the intensity of the planar stationary and isotropic point process of the centres of the circles. Still more is possible: also the pair correlation function of this point process can be determined, an analogue to the correlation function for random fields, see Stoyan (1990). This is possible by means of a "second-order version" of the solution of the Wicksell problem, see Stoyan, von Wolfersdorf & Ohser (1990).

The Stienen model is of course not the stationary and isotropic system with the largest non-overlapping spheres with centres in the points of a Poisson point process. The explanation of the Stienen model above already suggests a new model, the *lily-pond model* or *Poisson germ model with maximally non-overlapping spheres* (PMS model). The growth process above can of course be continued. Assume that the radial growth speed is the same for all spheres and growth is stopped for a sphere only then, when it comes in contact with another sphere. This system of spheres contains pairs of spheres which already belong to the Stienen model. All other spheres have radii larger than

the corresponding Stienen spheres.

This model first appeared in the context of percolation, see Häggström & Meester (1996). It is more complicated than the Stienen model. Even the determination of the volume fraction p seems to be rather difficult. Daley, Stoyan & Stoyan (1997) could show that the volume fraction of the PMS model is bounded by 0.2448; by simulation they found the value 0.186.

Systems of non-overlapping spheres with greater volume fractions can be obtained only when the system of centres is not a Poisson point process.

Such a system is the *simple sequential inhibition model* (SSI model, for short), see Diggle (1983). This model was originally defined for bounded regions. In the SSI model, spheres are placed sequentially and randomly in a given region. If a new sphere is placed so that it intersects a sphere already in place, then the new sphere is rejected. The placing process of spheres is usually stopped when it is impossible to place any new sphere. In this way for spheres of fixed radii a value of volume fraction $p = 0.38$ is possible in a "very large" region. As pointed out to me by J. Møller at a recent meeting in Oberwolfach, an infinite stationary extension can be obtained by means of a homogeneous Poisson process on the product space of space and time. Then, a sphere of radius R is associated with each of the "arrivals" of this Poisson process. If an arrival (x_i, t_i) is deleted when its sphere overlaps the sphere of another arrival (x_j, t_j) with $t_i > t_j$, then as time tends to infinity the retained spheres give a complete packing of the space. Equivalently, this can be interpreted as a spatial birth-and-death process as in Stoyan, Kendall & Mecke (1995), p. 185. The corresponding birth rate satisfies

$$b(x, \varphi) = 1 - \mathbf{1}_{\varphi \oplus b(o,R)}(x) ,$$

while the death rate vanishes.

M. Schlather has suggested a further definition, which may be useful for simulation. It uses the idea of dependent thinning which leads to Matern's second hard core process, see Stoyan, Kendall & Mecke (1995), p. 163. Take a Poisson process of intensity one in $\mathbb{R}^d \times [0, \infty)$ consisting of points (x, t). The Matern thinning rule applied to this process works as follows: A point (x, t) produces a point $x \in \mathbb{R}^d$ of the hard core process if there is no other point (x', t') with

$$\|x - x'\| < 2R \qquad \text{and} \qquad t' < t . \tag{7}$$

Then construct for all points retained cylinders centred at x, of radius $2R$ and infinite height. Delete all points in the cylinders and reconsider the points outside. Again a point (x, t) is retained if it satisfies (7) for all (x', t') outside of the cylinders, etc. If a sample of the stationary SSI model in the bounded region W is simulated following this idea, it is sufficient to generate points (x, t) with x in $W \oplus b(o, 2R)$.

A fascinating and complicated problem are *random dense* or *close packings of spheres* which is, until now, defined only operationally. In nature, there exist structures which seem to follow this model. Similar structures can be constructed by various physical experiments (see, for example, Frost, Schön & Salamon, 1993). One of them is filling a large container with hard metallic spheres and then shaking and pressing the system. Using such "natural" packings, random close packings can be characterized and explored. For example, it is known that the fraction of volume occupied by the spheres is about 0.64 or, as some people believe, $2/\pi = 0.6366$. But to give a rigorous definition is difficult. A sentence such as "a random close packed structure is defined to be a random packing of objects with packing fraction so large that any further increase in density will result in the formation of ordered substructures" (Frost, Schön & Salamon, 1993) describes some aspects of the problem but will not make any mathematician or statistician happy. Reading the classical paper in *Nature* 239, October 27, 1974, pages 488–489, is still recommendable.

Random close packings of spheres can be also simulated, and this is the way that has yielded most distributional information about such structures. Of course, only bounded packings can be simulated.

But by giving the volume fraction and the pair correlation function, it is implicitly assumed that there is *some* stationary and isotropic random close sphere packing. To date, no deeper mathematical work has been accomplished.

Various simulation procedures have been used. Early variants simulated the process of filling a container with hard spheres according to gravity. An alternative is a "Palm" approach in which a sphere is fixed at the origin and other spheres come from random directions and are packed stepwise around the centre. So far, no shaking or pressing is modeled in those two approaches. Better algorithms seem to be those by Jodrey & Tory (1984) and Stillinger, see Frost, Schön & Salamon (1993), and similar procedures, see Bargiel & Tory (1993). They begin with a system of spheres in which overlappings are permitted. During the application of the algorithm these overlappings are reduced stepwise and in parallel the structure is pressed.

A further model for stationary systems of non-overlapping spheres of fixed radius R is the *Poisson hard core process*. The sphere centres form a stationary and isotropic Gibbs point process with a pair potential $\phi(r)$ which is infinite for $r < 2R$ and vanishes otherwise. (For a complete definition more space than available here is necessary; see Stoyan, Kendall & Mecke, 1995.) The volume fraction of this model and the intensity of the point process of sphere centres depend on a further model parameter, the so-called chemical activity α. The smaller (i.e., since α is negative, the larger the absolute value) α is, the greater is the volume fraction. However, it is not clear whether or not the volume fraction is a continuous function of α or a function having one or more jumps.

S. Mase explained to me at a recent meeting in Oberwolfach that the maximum value of volume fraction is $\pi/\sqrt{18} \approx 0.74$, the volume fraction of the dense regular sphere packing. This is obtained in the limit as $\alpha \to -\infty$.

It is an open question which distributional properties the point process of sphere centres has in the case of a volume fraction of, say, 0.64. Is it similar to the centres point process corresponding to a random close packing of spheres?

Recall that stationary (thus infinite) Gibbs point processes and random sets have been discussed. Here, volume fraction and intensity are unknown characteristics, which depend on α. In the literature, also *finite* Gibbs point processes are frequently studied, which have a finite and known number of points in a given bounded region.

9 Particle Statistics

Particle statistics is an important branch of engineering statistics. Particles such as dust and powder particles or sand and rock grains play an important role in many technological processes. Their shape and size has influence on processes and properties such as agglomeration, aggregation, cohesivity, combustion, packing, permeability, porosity, etc. Therefore, there is a large literature, which perhaps should be better known to the statistical community. Two recent books (Hawkins, 1993, Stoyan & Stoyan, 1994) may serve as introductions. Therein and in most of the related papers, "particles" are planar images of three-dimensional particles, typically projections, as in Figure 5.

A classical method in particle statistics is describing the particles by one or more size parameters (for example, area or maximum Feret diameter) and a series of shape ratios. The perhaps most frequently used one is the circularity or area-perimeter ratio

$$f = \frac{4\pi A}{U^2},$$

where A is particle area and U particle perimeter. It takes the value 1 for a circular particle and otherwise values less than 1. The extent of deviation from 1 is considered as a measure of non-circularity. Many other shape ratios are considered in Stoyan & Stoyan (1994). They characterize, for example, deviations from convexity or from elliptic shape.

Each particle is then represented by a series of real numbers, and samples of particles are analysed

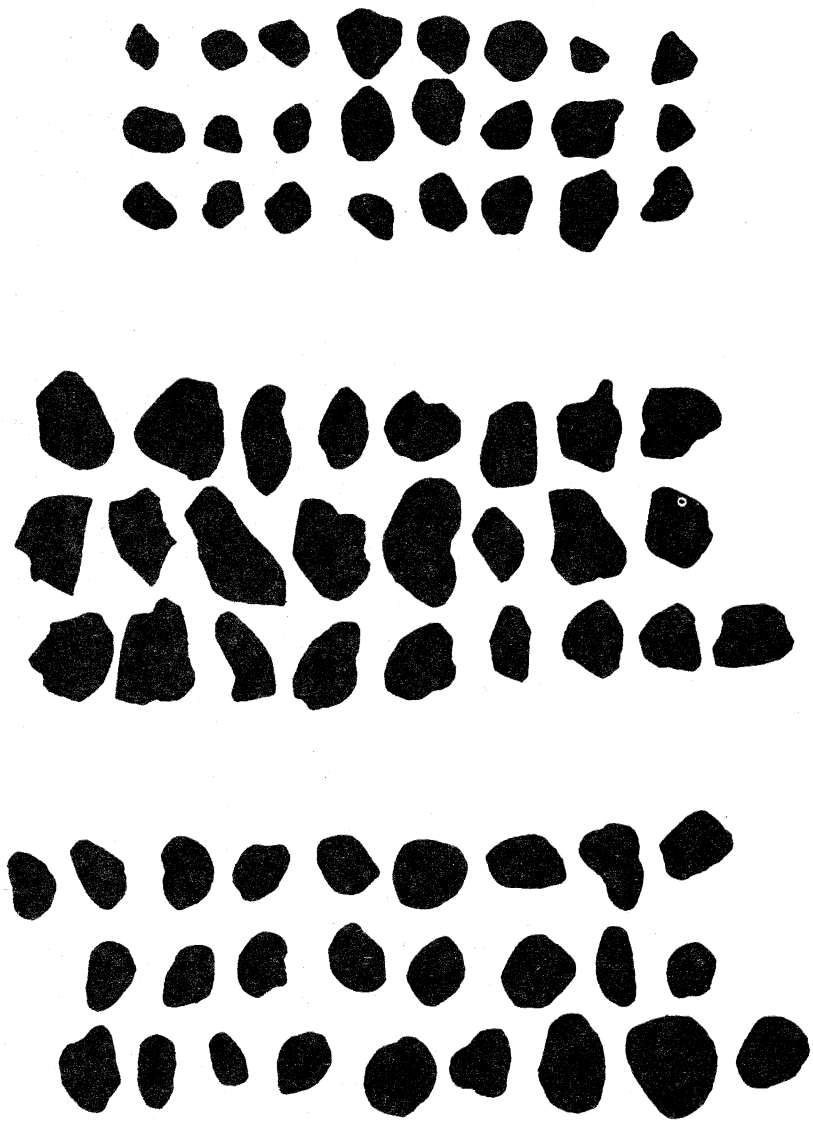

Figure 5. *Three samples of projected sand grains from Baltic Sea (above), River Selenchuk (Caucasus) (middle) and Desert Gobi.*

by means of methods of multivariate statistics. This approach often is a very powerful one, particularly when the "right" shape ratios are used.

Nevertheless, the author believes that a set-theoretic approach may have its advantages, perhaps not as an alternative but as a supplement or a method of visualisation.

A difficult problem of a set-theoretic approach in particle statistics is that the particles are not sets; they do not have an *a priori* location and orientation. For example, the locations and orientations of the particles in Figure 9 are arbitrary and irrelevant. Each particle could be considered a representative of an equivalence class of congruent sets.

The approach to be described now uses the following idea: the particles are given reasonable locations and orientations; so they are sets; for the obtained sample of sets, statistical methods for random compact sets are applied; in particular, sample means can be calculated.

The paper Stoyan & Molchanov (1997) gives the theoretical foundations of this approach; see also Stoyan (1997). The idea is to describe compact sets by real-valued functions, for example, by their indicator function or their support function. The latter is of course adapted to the case of convex particles; however, in some applications it may be acceptable to use convex hulls of non-convex particles. These functions are then considered to be elements of suitable L^2-spaces, and thus powerful Hilbert space methods are available. The main aim is the determination of a set-valued sample mean. It is defined in the sense of Fréchet for equivalence classes of Hilbert space elements with respect to isometries. As shown by Stoyan & Molchanov (1997) it is important to have the particles in *optimal relative position*, i.e. the sum of their squared pairwise distances has to be minimal. In this case, a representative of the equivalence classes mean can be obtained simply by averaging the functions describing the particles. This is analogous to ideas of Procrustean analysis (Ziezold, 1994, Stoyan & Stoyan, 1994, and Dryden & Mardia, 1998).

For an important particular case the optimal relative positions can be obtained in a very elegant way. This is the case of convex particles and support functions used for their description. Here the particles have to be shifted so that they are located with their *Steiner points* at the origin. The Steiner point S_Ξ of a convex set Ξ with support function s is given by

$$S_\Xi = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}_\phi s(\phi) \mathrm{d}\phi \ .$$

Here e_ϕ is the unit vector with angle ϕ between the corresponding line and x-axis, $0 \leq \phi \leq 2\pi$. Note that the support function can be defined as

$$s(\phi) = \max \langle e_\phi, x \rangle \quad \text{for } x \in \Xi$$

where $\langle \cdot, \cdot \rangle$ denotes the usual scalar product.

Having the Steiner points at the origin, it suffices to find the optimal orientations of the sets by minimizing the sum

$$\Delta = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \int_0^{2\pi} (s_i(\phi) - s_j(\phi + \alpha))^2 \mathrm{d}\phi \quad \text{for } 0 \leq \alpha \leq 2\pi \ . \tag{8}$$

Here s_i and s_j are support functions corresponding to convex sets having the same shape as the i-th and j-th element in the sample but rotated around the origin.

The sample mean is then the convex set having as support function the average of the n support functions of the particles in optimal relative position. This is the same as the Aumann mean of the sets (in optimal relative position). In general the Aumann mean of a random compact set is given by

$$\mathrm{E}(\Xi) = \{\mathrm{E}\xi : \xi \text{ is a selection point of } \Xi\} \ .$$

A selection point is a random vector (or point) belonging with probability one to Ξ. In the case of convex Ξ_i, $E(\Xi)$ can be obtained by

$$E(\Xi) = \frac{1}{n}(\Xi_1 \oplus \cdots \oplus \Xi_n) . \tag{9}$$

Unfortunately, in the case of indicator function approach there is not such an elegant solution. There is no unique location point in a particle with the property of the Steiner point that is given only by the geometry of the particle alone. Rather, it can be easily shown by examples that a fixed particle may have different location points when considered together with different other particles. Nevertheless, in analogy to the case of convex sets, Stoyan & Molchanov (1997) suggested to give, in heuristic averaging algorithms, the centers of gravity a role similar to that of the Steiner point. That means that the particles are first shifted so that their centres of gravity are in the origin. Then they are rotated so that a quantity analogous to Δ given by equation (8) but defined for indicator functions becomes a minimum. This optimization problem can be solved by an analogue of the Gover algorithm (see Stoyan & Stoyan, 1994, p. 149), which leads to the optimal orientations by iteration.

Then, when a sample of sets is given which are "close together", further statistical analysis makes sense. One can compute means, medians, variances and maximum and minimum elements in the sample. Here, the empirical covering function $\hat{p}$ plays an important role. It is defined by

$$\hat{p}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\Xi_i}(x)$$

i.e. $\hat{p}(x)$ is equal to that proportion of sets Ξ_i which contain the point x; Ξ_i is the i-th particle with centre of gravity in the origin and having optimal orientation.

There are good reasons to take the function $\hat{p}$ as an empirical mean of the sample, see Stoyan & Stoyan (1994). Also an interpretation as a fuzzy set is possible. A variance can be calculated by

$$s^2 = \frac{1}{n}\sum_{i=1}^{n} \int (\hat{p}(x) - \mathbf{1}_{\Xi_i}(x))^2 \mathrm{d}x .$$

A set-valued sample mean is the Vorob'ev mean $E_V \Xi$, see Stoyan & Stoyan (1994). It is the compact set that has as boundary that contour line of the function $\hat{p}(x)$ that bounds an area equal to the mean of the areas of the particles in the sample. A median can be given analogously: here the contour line is that given by $\hat{p}(x) = 0.5$.

Figure 6 shows the three estimated empirical covering functions for the sand grains, obtained after shifting to origin and optimal rotations. The figures show very well the fluctuations in shape and size of the sand grains. The roundness of the desert and sea grains becomes clearly visible, and also the size differences of the three samples is obvious. Figure 7 shows approximations for the Aumann means or mean bodies. They were calculated by means of formula (9), ignoring the fact that the particles are not convex. Also, these figures show well the shape and size differences of the three samples. Further statistical procedures and more statistical results for the sand grains are given in Stoyan & Stoyan (1994) and Stoyan (1997).

An alternative approach to particle statistics is the use of landmark methods, see Dryden & Mardia (1998). Here, each particle is described by k points on its outline, either "important" points (e.g. such of high curvature) or regularly distributed ones. Then these sets of "landmarks" are the statistical data in a Procrustean approach, they are translated, rotated, dilated and permutated. In spite of the possibility of using permutations, this approach is perhaps not as natural here as it is in biological applications where the landmarks (such as nose tip and chin tip if human face profiles are analysed) have a biological meaning (and thus are not permutated).

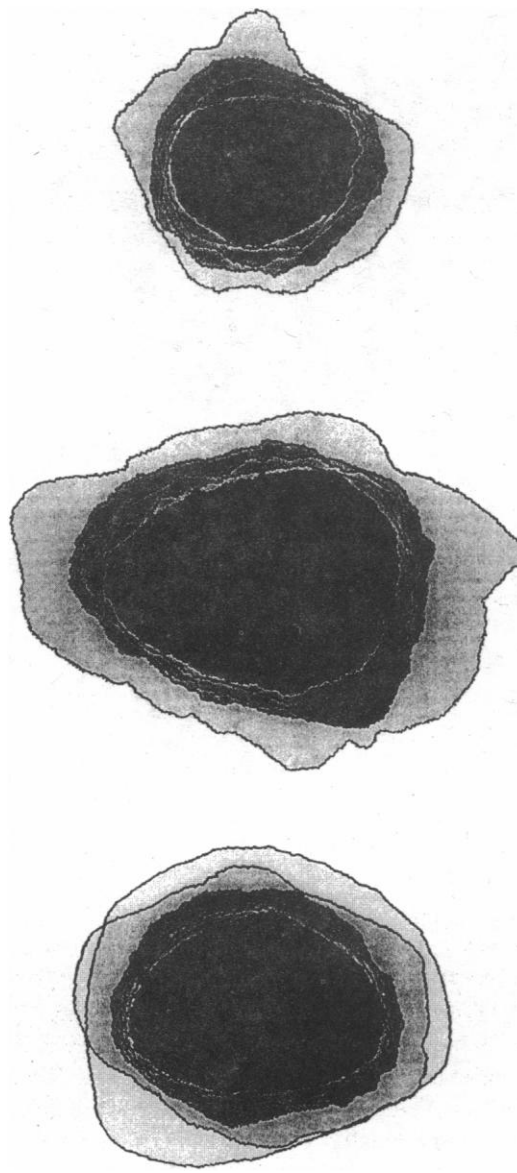

Figure 6. *Empirical covering functions for the three sand grain samples.*

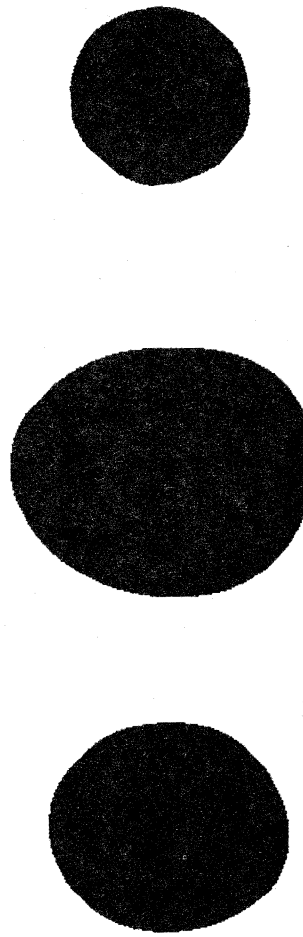

Figure 7. *Aumann means for the three sand grain samples.*

10 Models for Random Compact Sets

For the statistical analysis of particles, stochastic models may be of great value. Still little work has been done in this field, and the models to be described in the following have still an *ad hoc* flavour and lack a mathematical characterization.

Some interesting models exist for random polygons. A natural choice is to take 'typical' cells of planar stationary tessellations. A rigorous definition of 'typical' in this context uses the Palm theory for point processes, see Stoyan, Kendall & Mecke (1995). Heuristically, one can say that the typical cell is obtained from the whole tessellation by a sampling mechanism giving each cell the same chance to be sampled. (Matheron uses in this context the word "number law".)

A polygon of this type is the *Poisson Voronoi cell*, the typical cell of the Poisson Voronoi tessellation. Some facts about its distribution can be found in Møller (1994a), Okabe, Boots & Sugihara (1992), Stoyan, Kendall & Mecke (1995) and in Stoyan & Stoyan (1994). Another is the *Poisson polygon* studied by Matheron (1975) and Miles (1972). This is the typical cell of the Poisson line tessellation, i.e. the tessellation which results from a stationary and isotropic Poisson

line process.

These two random polygons are quite different: The mean numbers of vertices are 6 (Voronoi) and 4 (line), the circularity ratio calculated from *mean* area and *mean* perimeter is $\frac{\pi}{4} = 0.7854$ and 1, respectively. The author's impression is that these models are only of limited value for particle statistics; both depend only on one size parameter and thus are not flexible. In order to make the Poisson polygon more realistic, Cauwe (1973) has introduced the Poisson twin flat model.

A great variety of models is offered by convex hulls of random finite point sets. Since there are various possibilities for the distributions of the points, there should be a great variability of the polygon distribution, too. While many papers have been published on asymptotic properties of convex hulls of many points, little is known about the exact distributional properties of convex hulls with a few (4 to 10) points. In this context also the union-stable sets studied in Molchanov (1993) may be of value, see also Molchanov & Stoyan (1996) and Molchanov (1997b), which also gives a good survey of limit theorems for random sets in general.

The models discussed so far are perhaps not so realistic for particle statistics. Frequently particles are available only as projected three-dimensional particles, while the models are of a pure two-dimensional nature. Therefore, it could be useful to study planar projections of three-dimensional polyhedra. (In the case of convex hulls it suffices to determine the distribution of the projection of the generating point set and then to use the planar theory.) Integral geometry helps to find the mean values $\overline{A}$ and $\overline{U}$ of area and perimeter of the projected sets if the direction of projection is uniform by means of

$$\begin{aligned} \overline{U} &= \pi \overline{\overline{b}} \\ \overline{A} &= \frac{1}{4} \overline{S} \end{aligned}$$

where $\overline{\overline{b}}$ and $\overline{S}$ are mean average breadth and mean surface area of the three-dimensional polyhedron. Also the mean number n_2 of vertices of the projection polygon can be calculated. It satisfies the relation

$$n_2 = \mathrm{E}(N_3) - 2\mathrm{E}(\sum_{i=1}^{N_3} \Phi_i)$$

where N_3 is the (random) number of vertices of the three-dimensional polyhedron and Φ_i the relative angle at the i-th vertex, see Miles (1969). For deterministic polyhedra, the angles can be found by elementary geometrical methods, while, for the typical cell of a stationary tessellation, point process methods yield the mean angle sum, see Radecke (1980). The paper Stoyan & Stoyan (1996) contains values for the Platonian bodies as well as for the Poisson Voronoi cell and the Poisson polyhedron. The mean number of vertices of the projections of the latter polyhedra are 6 and 13.474, while the mean numbers for the original polyhedra are 8 and 27.01, respectively. The result for the Poisson polyhedron is a particular case of a more general result: the mean number of vertices of the projection is 6 for all typical cells of stationary plane tessellations.

However, the sand grains of Figure 5 are not polygons, they have rounded boundaries. Good models for such cases seem not to exist to date. Some *ad hoc* ideas are the following. One could start from a random polygon Π and consider the parallel set $\Pi \oplus b(o, r)$ that has circular parts in its boundary. A further model is the *shot noise circle*. In analogy to the shot noise process on the real line (see Daley & Vere-Jones, 1988) a shot noise process is considered on a circle. The random impulses are placed at the outside of the circular line and produce so outer bumps and eventually an irregular non-convex figure. Finally, the *Brownian circle* can be mentioned as an irregular non-polygonal figure, which is, however, not a model for sand grains. It is an isotropic starshaped set C centred at the origin. The distance $r(\phi)$ from o to the boundary of C in the direction of the ray starting in o and

having direction ϕ with positive x-axis is random and satisfies

$$r(\phi) = r + B_{\phi+\xi} \quad \text{for } 0 \leq \phi < 2\pi \ .$$

Here r is a positive constant and $\{B_t\}$ is the Brownian bridge on $[0, 2\pi]$, with $0 \leq t \leq 2\pi$. Furthermore, ξ is an independent and uniform random variable on $[0, 2\pi]$ and addition of indices is understood as modulo 2π. This set was used in Stoyan & Lippmann (1993) for modelling a figure which was before studied by Cressie & Hulting (1992): the irregular area covered by a human breast cancer grown on a flat dish covered with a nutrient medium. The Brownian circle is an alternative to Cressie's cancer model, being easier to simulate and producing figures more similar to the data, but offering no biological explanation. The problem of modelling this cancer is a particular case of the general problem of modelling spreading-out phenomena, which are also observed in fire research, see Vorob'ev (1996). Until now the most models for such phenomena are lattice-based.

11 A Point for Future Work

The models of stochastic geometry are sometimes critized by engineers and physicists. They say that these models are typically of a pure geometrical nature. In contrast, the irregular geometrical structures of nature and technology are formed according to physical or biological processes and principles. Mathematicians and statisticians should contact physicists, biologists or engineers or should try to learn enough on physics and biology to be able to build better models. There are of course first steps in this direction, for example the use of Gibbs processes. It may be that this is not so attractive for mathematicians, since they know or feel that realistic stochastic-geometric models are rather complicated and only lead to computer-intensive simulations. But as recent work in the field of Markov chain Monte Carlo methods shows, it is also a mathematically interesting problem to study and construct good simulation algorithms and combine this with more classical inference procedures, see e.g. Geyer (1997), Møller (1997) and Kendall (1997).

12 Acknowledgment

The author thanks L. Heinrich, D. Jeulin and I. Molchanov for careful reading of an earlier version of this paper and for many suggestions and additions. An intensive discussion with J. Møller improved the paper very much. Some colleagues helped with hints to references and questions and comments during the conferences in Vienna and Fontainebleau. E. Arjas and J. Rothe helped improving the English style of the paper. Last but not least the author thanks his wife for many discussions related to this paper; in particular, Helga has suggested the maximum Poisson spheres model.

References

Adler, R.J. (1982). *The Geometry of Random Fields*. New York: J. Wiley & Sons.

Aldous, D. (1989). *Probability Approximations via the Poisson Clumping Heuristic*. New York: Springer-Verlag.

Aubin, J.-P. & Frankowska, H. (1990). *Set-Valued Analysis*. Basel: Birkhäuser.

Baddeley, A. & Molchanov, I.S. (1997). On the expected measure of a random set. In: *Advances in Theory and Applications of Random Sets*. Proc. Int. Symp. Oct 9–11, 1996, Fontainebleau (ed. D. Jeulin). Singapore: World Scientific Publ. Comp. pp. 3–20.

Baddeley, A. & van Lieshout, M.N.M. (1995). Area-interaction point processes. *Ann. Inst. Statist. Math.* **47**, 601–619.

Baddeley, A., Kendall, W.S. & van Lieshout, M.C. (1998). Quermass-interaction models. *Adv. Appl. Probab.* **30**.

Bandemer, H. & Näther, W. (1992). *Fuzzy Data Analysis*. Dordrecht: Kluwer.

Bargiel, M. & Tory, E.M. (1993). Packing fraction and measures of disorder of ultradense packings of equal spheres. I. Nearly ordered packing. *Adv. Powder Technol.* **4**, 79–101.

Bindrich, U. & Stoyan, D. (1990). Stereology for pores in wheat bread: statistical analyses for the Boolean model by serial sections. *J. Microsc.* **162**, 231–239.

Cahn, J.W. (1967). The significance of average mean curvature and its determination by quantitative metallurgy. *Trans. Amer. Inst. Metall. Eng.* **239**, 610–616.

Capasso, V., Micheletti, A., De Giosa, M. & Mininni, R. (1996). Stochastic modelling and statistics of polymer crystallization processes. *Surv. Math. Industry* **6**, 109–132.
Cauwe, Ph. (1973). *Morphologie mathématique et fragmentation.* PhD Thesis, University of Louvain.
Chiu, S.N. (1995). Limit theorems for the time of completion of Johnson–Mehl tessellations. Adv. Appl. Probab. **27**, 889–910.
Choquet, G. (1953, 1954). Theory of capacities. *Ann. Inst. Fourier* **V**, 131–295.
Cramér, H. & Leadbetter, M.R. (1967). *Stationary and Related Stochastic Processes.* New York: J. Wiley & Sons.
Cressie, N. & Hulting, F.L. (1992). A spatial statistical analysis of tumor growth. *J. Amer. Statist. Assoc.* **87**, 272–283.
Daley, D.J., Stoyan, D. & Stoyan, H. (1997). The volume fraction of a Poisson germ model with maximally non-overlapping spherical grains. (In preparation)
Daley, D.J. & Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes.* New York: Springer-Verlag.
DeHoff, R.T. (1967). The quantitative estimation of mean surface curvature. *Trans. Amer. Inst. Metall. Eng.* **239**, 617–621.
Diggle, P.J. (1981). Binary mosaics and the spatial pattern of heather. *Biometrika* **64**, 91–95.
Diggle, P.J. (1983). *Statistical Analysis of Point Processes.* London: Chapman & Hall.
Dryden, I.L. & Mardia, K.V. (1998). *Statistical Shape Analysis.* Chichester: J. Wiley & Sons.
Erhardsson, T. (1996). On the number of high excursions of a linear growth process. *Stoch. Proc. Appl.* **65**, 31–53.
Frost, R., Schön, J.C. & Salamon, P. (1993). Simulation of random close packed discs and spheres. *Computat. Materials Sc.* **1**, 343–350.
Geyer, C.J. (1997). Likelihood inference for spatial point processes. In: *Proc. Seminaire Européen de Statistique, Stochastic Geometry, Likelihood, and Computation* (ed. O. Barndorff-Nielsen, W.S. Kendall & M.N.N. van Lieshout). London: Chapman & Hall.
Gilbert, E.N. (1962). Random subdivisions of space into crystals. *Ann. Math. Statist.* **33**, 958–972.
Gill, R.D., van der Laan, J.C. & Robins, J.M. (1997). Coarsening at random: characterizations, conjectures and counter-examples. In: *Proc. First Seattle Conference on Biostatistics* (ed. D.Y. Lin). New York: Springer-Verlag.
Goodman, I.R. & Nguyen, H.T. (1985). *Uncertainty Models for Knowledge Based Systems.* Amsterdam: North Holland.
Goodman, I.R., Nguyen, H.T. & Walker, E.A. (1991). *Conditional Inference and Logic for Intelligent Systems: A Theory of Measure-Free Conditioning.* Amsterdam: North-Holland.
Goulard, M., Särkkä, A. & Grabarnik, P. (1996). Parameter estimation for marked Gibbs point processes through the maximum pseudo-likelihood method. *Scand. J. Statist.* **23**, 365–379.
Haas, A., Matheron, G. & Serra, J. (1967). Morphologie mathématique et granulométries en place I, II. *Ann. Mines* **11**, 736–753; **12**, 767–782.
Häggström, O. & Meester, R. (1996). Nearest neighbour and hard sphere models in continuum percolation. *Random Struct. Algorithms* **9**, 295–315.
Häggström, O., van Lieshout, M.N.M. & Møller, J. (1996). *Characterisation results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes.* R-96-2040, Department of Mathematics, Aalborg University. (To appear in *Bernoulli*).
Hall, P. (1988). *Introduction to the Theory of Coverage Processes.* New York: J. Wiley & Sons.
Hawkins, A.E. (1993). *The Shape of Powder-Particle Outlines.* Taunton: Research Studies Press and New York: J. Wiley & Sons.
Heinrich, L. (1992). On existence and mixing properties of germ-grain models. *Statistics* **23**, 271–286.
Heinrich, L. (1993). Asymptotic properties of minimum contrast estimators for parameters of Boolean models. *Metrika* **40**, 67–94.
Heinrich, L. (1996). Contact and chord length distribution functions of a stationary Voronoi tessellation. *Adv. Appl. Prob.*
Heinrich, L. & Molchanov, I.S. (1995). *Central limit theorem for a class of random measures associated with germ-grain models.* CWI Report BS-R9518, June 1995.
Heinrich, L. & Werner, M. (1998). Kernel estimation of the diameter distribution in Boolean models with spherical grains.
Hildenbrand, W. (1974). *Core and Equilibria of a Large Economy.* Princeton: Princeton University Press.
Jeulin, D. (1978). Description du réseau poreux macroscopique de cokes moulés par un modèle morphologique. Actes du Symposium "Analyse quantitative des microstructures en sciences des matériaux, biologie et médecine" (Caen 1977), Special Issues of Practical Metallography, n°. 8. Stuttgart: Riederer Verlag, pp. 253–263.
Jeulin, D. (1997). Dead leaves models: From space tessellation to random functions. In: *Advances in Theory and Applications of Random Sets.* Proc. Int. Symp. Oct 9–11, 1996, Fontainebleau (ed. D. Jeulin). Singapore: World Scientific Publ. Comp. pp. 137–156.
Jeulin, D. (1998). *Morphological Models of Random Structures.* Boca Raton, Florida: CRC Press.
Jodrey, W.S. & Tory, E.M. (1985). Computer simulation of close random packing of equal spheres. *Phys. Rev.* A **32**, 2347–2351.
Kellerer, A.M. (1983). On the number of clumps resulting from the overlap of randomly placed figures in the plane. *J. Appl. Prob.* **20**, 126–135.
Kelton, K.F. (1993). Numerical model for isothermal and non-isothermal crystallization of liquids and glasses. *J. Non-Crystalline Solids.* **163**, 283–296.
Kendall, D.G. (1974). Foundations of a theory of random sets. In: *Stochastic Geometry* (eds. E.F. Harding & D.G. Kendall), pp. 322–376, London: J. Wiley & Sons.
Kendall, W.S. (1997). On some weighted Boolean models. In: *Advances in Theory and Applications of Random Sets.* Proc. Int. Symp. Oct 9–11, 1996, Fontainebleau (ed. D. Jeulin). Singapore: World Scientific Publ. Comp. pp. 105–120.
Kendall, W.S. & Møller, J. (1997). Perfect Metropolis-Hastings simulation of locally stable point processes. (In preparation).
Kingman, J.F.C. (1993). *Poisson Processes.* Oxford: Clarendon Press.
Klein, W. (1982). Potts-model formulation of continuum percolation. *Physical Review* B**26**, 2677–2678.
Kolmogoroff, A.N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Berlin: Springer-Verlag. Reprinted by Chelsea,

New York, 1946.
Kolmogoroff, A.N. (1937). Statistical theory of crystallization of metals. *Bull. Acad. Sci. USSR, Mat. Ser.* **1**, 355–359.
Kruse, R. & Meyer, K.D. (1987). *Statistics with Vague Data*. Dordrecht: D. Reidel.
Li, S.-K. (1994). *Random Sets and Set-Valued Martingales*. Guizhou: Guizhou Science and Technology Publ. Co.
Mahler, R.P.S. (1993). Random sets as a foundation for general data fusion. In: *Proceedings 6th Joint Service Data Fusion Symp., vol. 1, part 1*, Applied Physics Laboratory, The Johns Hopkins University, Laurel, MD, June 14–18, 1993, pp. 358–394.
Mahler, R.P.S. (1994). Global integrated data fusion. *7th Nat. Symp. Sensor Fusion*, Sundia Laboratories, Albuquerque, New Mexico, March 16–18, 1994.
Mase, S. (1985). On the possible form of size distributions of mutually non-intersecting discs. *J. Appl. Prob.* **23**, 646–659.
Matheron, G. (1967). *Eléments pour une théorie des mileux poreux*. Paris: Masson.
Matheron, G. (1975). *Random Sets and Integral Geometry*. New York: J. Wiley & Sons.
Miles, R.E. (1969). Poisson flats in Euclidean spaces. Part I: A finite number of random uniform flats. *Adv. Appl. Prob.* **1**, 211–237.
Miles, R.E. (1972). The random division of space. *Suppl. Adv. Appl. Probab.*, 243–266.
Molchanov, I.S. (1993). *Limit Theorems for Unions of Random Sets*. Lecture Notes in Mathematics 1561. Berlin: Springer-Verlag.
Molchanov, I.S. (1995). Statistics of the Boolean model: from the estimation of means to the estimation of distributions. *Adv. Appl. Prob.* **27**, 63–86.
Molchanov, I.S. (1996a). Set-valued estimators for mean bodies related to Boolean models. *Statistics* **28**, 43–56.
Molchanov, I.S. (1996b). A limit theorem for scaled vacancies in the Boolean model. *Stochastics, Stochastic Reports* **58**, 45–65.
Molchanov, I.S. (1997a). *Statistics of the Boolean Model for Practitioners and Mathematicians*. Chichester: J. Wiley & Sons.
Molchanov, I.S. (1997b). Random closed sets: results and problems. In: *Proc. Seminaire Européen de Statistique, Stochastic Geometry, Likelihood, and Computation* (ed. O. Barndorff-Nielsen, W.S. Kendall & M.N.M. van Lieshout). London: Chapman & Hall.
Molchanov, I.S. & Stoyan, D. (1994). Asymptotic properties of estimators for parameters of the Boolean model. *Adv. Appl. Prob.* **26**, 301–323.
Molchanov, I.S. & Stoyan, D. (1996). Statistics of compact sets and random polygons. *Commun. Statist. Stoch. Models* **12**, 199–214.
Møller, J. (1992). Random Johnson–Mehl tessellations. *Adv. Appl. Probab.* **24**, 814–844.
Møller, J. (1994a). *Lectures on Random Voronoi Tessellations*. Lecture Notes in Statistics **87**, Berlin, Heidelberg, New York: Springer-Verlag.
Møller, J. (1994b). Discussion contribution. *Scand. J. Statist.* **21**, 346–349.
Møller, J. (1997). Markov chain Monte Carlo and spatial point processes. In: *Proc. Seminaire Européen de Statistique, Stochastic Geometry, Likelihood, and Computation* (ed. O. Barndorff-Nielsen, W.S. Kendall & M.N.M. van Lieshout). London: Chapman & Hall.
Møller, J. & Waagepetersen, R. (1998). Markov connected component fields. *Adv. Appl. Prob. (SGSA)*, **30**.
Muche, L. & Stoyan, D. (1992). Contact and chord distributions of the Poisson-Voronoi tessellation. *J. Appl. Prob.* **29**, 467–471.
Norberg, T. (1984). Convergence and existence of random set functions. *Ann. Prob.* **12**, 726–732.
Norberg, T. (1992). On the existence of ordered couplings of random sets—with applications. *Israel J. Math.* **77**, 241–264.
Nott, D.J. & Wilson, R.J. (1997a). Size distributions for excursion sets. In: *Advances in Theory and Applications of Random Sets*. Proc. Int. Symp. Oct 9–11, 1996, Fontainebleau (ed. D. Jeulin). Singapore: World Scientific Publ. Comp. pp.175–196.
Nott, D.J. & Wilson, R.J. (1997b). Parameter estimation for excursion set texture models..
Okabe, A., Boots, S. & Sugihara, K. (1992). Spatial Tessellations—Concepts and Applications of Voronoi Diagrams. Chichester: J. Wiley & Sons.
Preston, C.J. (1976). *Random Fields*. Lecture Notes in Mathematics, Berlin, Heidelberg, New York: Springer-Verlag.
Radecke, W. (1980). Some mean value relations on stationary random mosaics in the space. *Math. Nachr.* **97**, 203–210.
Robbins, H.E. (1944, 1945). On the measure of a random set I, II. *Ann. Math. Statist.* **15**, 70–74; **16**, 342–347.
Salinetti, G. & Wets, J.-B. (1986). On the convergence in distribution of measurable multifunctions (random sets), normal integrands, stochastic processes and stochastic infima. *Math. Oper. Res.* **11**, 385–422.
Santaló, L.(1976). *Integral Geometry and Geometrical Probability*. Reading: Addison-Wesley.
Schlather, M. & Stoyan, D. (1997). Covariance of the Stienen model. In: *Advances in Theory and Applications of Random Sets*. Proc. Int. Symp. Oct 9–11, 1996, Fontainebleau (ed. D. Jeulin). Singapore: World Scientific Publ. Comp. pp. 157–174.
Serra, J. (1982). *Image Analysis and Mathematical Morphology*. London: Academic Press.
Sivarkumar, K. & Goutsias, J. (1997). On the morphological discretization of random closed sets. In: *Advances in Theory and Applications of Random Sets*. Proc. Int. Symp. Oct 9–11, 1996, Fontainebleau (ed. D. Jeulin). Singapore: World Scientific Publ. Comp. pp. 49–66.
Stoyan, D. (1989). Statistical inference for a Gibbs point process of mutually non-intersecting discs. *Biom. J.* **31**, 153–161.
Stoyan, D. (1990). Stereological formulae for a random system of non-overlapping spheres. *Statistics* **21**, 131–136.
Stoyan, D. (1997). Geometrical means, medians and variances for samples of particles. *Particle and Particle Systems Charact.* **14**, 30–34.
Stoyan, D. & Stoyan, H. (1996). Mean values for orthogonal projections of random convex polyhedra. Research Report, TU

Bergakademie Freiberg.
Stoyan, D. & Molchanov, I.S. (1997). Set-valued means of random particles. *Math. Imaging and Vision* **7**, 111–121.
Stoyan, D. & Lippmann, G. (1993). Models of stochastic geometry—a survey. *ZOR—Methods and Models of Opns. Res.* **38**, 235–260.
Stoyan, D., Kendall, W.S. & Mecke, J. (1995). *Stochastic Geometry and Its Applications*. Chichester: J. Wiley & Sons.
Stoyan, D., von Wolfersdorf, L. & Ohser, J. (1990). Stereological problems for spherical particles—second-order theory. *Math. Nachr.* **146**, 33–46.
Torquato, S. (1991). Random heterogeneous media: Microstructure and improved bounds on effective properties. *Appl. Mech. Rev.* **44**, 37–76.
Uebele, P. & Hermann, H. (1996). Computer simulation of crystallization kinetics with non-Poisson distributed nuclei. *Modelling Simul. Mater. Sci. Eng.* **4**, 203–214.
Vorob'ev, O. (1996). Random set models of fire spread. *Fire Technology* **32**, 137–173.
Wang, P. Z. (1985). *Fuzzy Sets and Falling Shadow of Random Sets*. Bejing: Bejing Normal University Press.
Weil, W. (1995). The estimation of mean shape and mean particle number in overlapping particle systems in the plane. *Adv. Appl. Prob.* **27**, 102–119.
Weil, W. (1988). Expectation formulas and isoperimetric properties for non-isotropic Boolean models. *J. Microsc.* **151**, 235–245.
Weil, W. (1997). The mean normal distribution of stationary random sets and particle processes. In: *Advances in Theory and Applications of Random Sets*. Proc. Int. Symp. Oct 9–11, 1996, Fontainebleau (ed. D. Jeulin). Singapore: World Scientific Publ. Comp. pp. 21–34.
Widom, J.S. & Rowlinson, B. (1970). New model for the study of liquid-vapour phase transition. *J. Chem. Phys.* **52**, 1670–1684.
Zähle, U. (1984). Random fractals generated by random cutouts. *Math. Nachr.* **116**, 27–52.
Zhang, W.-X., Wang, P.-W. & Gao, Y. (1996). *Set-valued Stochastic Processes*. Beijing: Science Publ. Co.
Ziezold, H. (1994). Mean figures and and mean shapes applied to biological figure and shape distributions in the plane. *Biom. J.* **36**, 491–510.

Résumé

Cet article est un survol de certains aspects d'intérêt pratique et actuel de la théorie des ensembles fermés aléatoires. Il débute par quelques remarques sur l'histoire de cette partie de la théorie des probabilités et sur le rôle important de Georges Matheron. Ensuite sont introduites les caractéristiques fondamentales de la distribution des ensembles fermés aléatoires. Le modèle très important appelé schéma Booléen sert d'exemple pour discuter des problèmes mathématiques et statistiques. Une série d'autres modèles est examinée, à savoir: les ensembles d'excursion de champs aléatoires, les systèmes d'arêtes de la partition de Poisson Voronoi, et plusieurs systèmes aléatoires de sphères sans recouvrement. Finalement, des idées de statistique des particules, comprenant des modèles d'ensembles aléatoires compacts, sont présentées.

[*Received January 1997, accepted July 1997*]