Survival Functions and Contact Distribution Functions for Inhomogeneous, Stochastic Geometric Marked Point Processes

Vincenzo Capasso and Elena Villa

Dept. of Mathematics, University of Milan, via Saldini 50, 20133 Milano, Italy email: vincenzo.capasso@unimi.it, elena.villa@mat.unimi.it

Abstract

Many real phenomena , including phase change, such as crystallization processes, tumor growth, forest growth, etc. may be modelled as stochastic birth-and-growth processes, in which crystals develop from points (nuclei) that are born at random both in space and time. In this paper we revisit these processes by classical methods of survival analysis with specific reference to the role played by the survival function and the corresponding hazard rate with respect to capture of a point by the so called crystalline phase. General expressions for the hazard and survival functions associated with a point are provided. Known results for Poisson type processes follow as particular cases. Further a link between hazard functions and contact distribution function of stochastic geometry is also obtained.

Keywords: Contact distribution function; Crystallization; Hazard function; Stochastic birth-and-growth processes; Survival analysis AMS Classification 2000: 60D05,28A75

1 Introduction and Basic Notations

A birth-and-growth process is composed of two processes, birth (nucleation) and subsequent growth of spatial cells (crystals), which are, in general, stochastic both in time and space. Thus, for each fixed time t, the crystallized region, denoted by Θ^t , is a random closed set and so it may be studied by using tools of stochastic geometry. On the other hand the process evolves in time; following the classical approach based on Doob-Meyer decomposition theorem, dynamical aspects are described in terms of its compensator with respect to the natural history of the process, which includes the (random) spatial occupation due to the growth of crystals. The aim of this paper is to analyze these processes by the methods of survival analysis, thus extending to the spatially heterogeneous case partial results known for homogeneous Poisson nucleation processes (see e.g. [3],[5]). In particular, the survival function S(t,x) of a point x at time t, representing the probability that x is not yet crystallized at time t, and its link with the so-called hazard function h(t,x), representing the rate of capture of the point x, are analyzed in Section 2. In Section 3 we relate the hazard function to the well known contact distribution function of stochastic geometry (see, e.g., [14]).

Definition 1 A marked point process (MPP) Φ on a complete separable metric space (c.s.m.s.) \mathcal{X} with marks in \mathbf{K} c.s.m.s. is a point process on $\mathcal{X} \times \mathbf{K}$ with the property that the marginal process $\{\Phi(B \times \mathbf{K}) : B \in \mathcal{B}(\mathcal{X})\}$ is itself a point process. (See, e.g., [6]).

So the nucleation process can be modelled by a MPP Ψ on \mathbb{R}^+ with marks in a measurable space $\mathbf{K} \subset \mathbb{R}^d$:

$$\Psi = \sum_{n=1}^{\infty} \varepsilon_{(T_n, X_n)},$$

where the random variables X_n and T_n are, respectively, the spatial location in \mathbf{K} and the birth time of the nth nucleus.

For any $A \in \mathcal{B}(\mathbb{R}^+)$ and $B \in \mathcal{B}(\mathbf{K})$ (where \mathcal{B} is the Borel σ -algebra) we have

$$\Psi(A \times B) = \sum_{n} \mathbf{1}_{A}(T_{n}) \mathbf{1}_{B}(X_{n})$$

$$= \#\{\text{nuclei born in } B \text{ during time } A\}.$$

Denoted by $\mathcal{F} = \{\mathcal{F}_t\}$ the internal history of the process Ψ , where

$$\mathcal{F}_t = \sigma\{\Psi([0, s] \times B) : 0 < s \le t, B \in \mathcal{B}(\mathbf{K})\},\$$

it is possible to prove (see [9],[11]) that there exists a random measure ν on $\mathbb{R}^+ \times \mathbf{K}$, said compensator of Ψ , such that $\forall B \in \mathcal{B}(\mathbf{K})$,

- i) the process $\nu(t, B) = \nu([0, t] \times B)$ is \mathcal{F}_t -predictable;
- ii) the compensated process

$$M(t,B) = \Psi([0,t] \times B) - \nu([0,t] \times B)$$

is a zero-mean martingale.

On the other hand, denoted by

$$\tilde{\nu}(dt) = \nu(dt \times \mathbf{K})$$

the compensator of the marginal process $\tilde{\Psi}(\cdot) = \Psi(\cdot \times \mathbf{K})$, there exists [11] a stochastic kernel k from $\Omega \times \mathbb{R}^+$ to \mathbf{K} such that

$$\nu(dt \times dx) = k(t, dx)\tilde{\nu}(dt). \tag{1}$$

Remark 1 In many applications it is supposed that further nuclei cannot be born in an already crystallized zone. When we want to emphasize this, we will write

$$\nu(dt \times dx) = k(t, dx)\tilde{\nu}(dt) = k_0(t, dx)\tilde{\nu}(dt)(1 - \mathbf{1}_{\Theta^{t-}}(x)),$$

where $\nu_0(dt \times dx) = k_0(t, dx)\tilde{\nu}(dt)$ is the compensator of the process Ψ_0 , called the *free-process*, in which nuclei can be born anywhere.

In either cases the compensator of the process, be it free or not, can be factorized as in (1), so we will consider a generic process Ψ , unless otherwise specified.

Definition 2 The intensity measure of Ψ is the (σ -finite) measure Λ on $\mathbb{R}^+ \times \mathbf{K}$ given by

$$\Lambda(A \times B) := \mathbf{E}(\Psi(A \times B)).$$

We observe that from i) and ii) it follows that

$$\nu(dt \times dx) = \mathbf{E}(\nu(dt \times dx) | \mathcal{F}_{t-})$$

$$= \mathbf{E}(\Psi(dt \times dx) | \mathcal{F}_{t-}) - \mathbf{E}(M(dt \times dx) | \mathcal{F}_{t-})$$

$$= \mathbf{E}(\Psi(dt \times dx) | \mathcal{F}_{t-}), \qquad (2)$$

and so

$$\mathbf{E}(\nu(dt \times dx)) = \mathbf{E}[\mathbf{E}(\Psi(dt \times dx) \mid \mathcal{F}_{t-})] = \Lambda(dt \times dx); \tag{3}$$

moreover, since **K** is a c.s.m.s. and Λ is a σ -finite measure, it is possible to factorize Λ in the following way [11]:

$$\Lambda(dt \times dx) = \tilde{\Lambda}(dt)Q(t, dx),\tag{4}$$

where $\tilde{\Lambda}$ is the intensity measure of the marginal process $\tilde{\Psi}$ and, $\forall t \in \mathbb{R}^+$, $Q(t, \cdot)$ is a probability measure on \mathbf{K} , called the *mark distribution* at time t.

Definition 3 A point process Φ on \mathbb{R}^d is said simple if $\Phi(\{x\}) \leq 1 \ \forall x \in \mathbb{R}^d$.

Remark 2 If Ψ is a MPP with simple marginal process $\tilde{\Psi}$ and marks X_i , then, formally, by (1) and (2) it follows that

$$k(t,B)\tilde{\nu}(dt) = \mathbf{E}(\Psi(dt \times B)|\mathcal{F}_{t-})$$

$$= \mathbf{P}(\Psi(dt \times B) = 1|\mathcal{F}_{t-})$$

$$= \mathbf{P}(X_1 \in B|\tilde{\Psi}(dt) = 1, \mathcal{F}_{t-})\mathbf{E}(\tilde{\Psi}(dt)|\mathcal{F}_{t-})$$

$$= \mathbf{P}(X_1 \in B|\tilde{\Psi}(dt) = 1, \mathcal{F}_{t-})\tilde{\nu}(dt),$$

from which it is evident that k(t, dx) is in general stochastic. Observe that by (3) and (4) we have

$$Q(t, dx)\tilde{\Lambda}(dt) = \mathbf{E}(\nu(dt \times dx)) = \mathbf{E}(k(t, dx)\tilde{\nu}(dt));$$

so, if Ψ is a MPP with marks independent of the marginal process, then

$$\mathbf{E}(k(t, dx)\tilde{\nu}(dt)) = \mathbf{E}(k(t, dx))\mathbf{E}(\tilde{\nu}(dt)) = \mathbf{E}(k(t, dx))\tilde{\Lambda}(dt),$$

that is

$$Q(t, dx) = \mathbf{E}(k(t, dx)).$$

In particular, if Ψ is a position-dependent Q-marking of $\tilde{\Psi}$ (see [11]), then k(t, dx) is deterministic and coincides with Q(t, dx); while if Ψ is an independent Q-marking of $\tilde{\Psi}$ (i.e. marks X_i are IID random elements with distribution Q and independent of $\tilde{\Psi}$) we have k(dx) = Q(dx) [8].

By hypothesis we consider birth-and-growth processes Ψ such that the marginal process $\tilde{\Psi}$ is simple and the mark distribution $Q(t,\cdot)$ is absolutely continuous with respect to the d-dimensional Lebesgue measure ν_d on \mathbf{K} ; besides, we assume normal growth with a growth speed G of crystals continuous and in general space-and-time dependent such that $G(t,x) \geq G_0 > 0$ for all $t \in \mathbb{R} <^+, x \in \mathbf{K}$.

Denote by $\Theta_s^t(x)$ the crystal born at time s and point x and grown to time t. Then the occupied region at time t is given by

$$\Theta^t = \bigcup_{j: T_j \le t} \Theta^t_{T_j}(X_j).$$

Definition 4 The survival function of a point x at time t is the probability that the point x is not yet covered ("captured") by any crystal at time t:

$$S(t,x) := \mathbf{P}(x \notin \Theta^t).$$

2 Survival and Hazard Functions

In order to study the survival function of a point x, it is useful to introduce the nonnegative r.v. T_x representing the *capture time of* x; in this way we can relate it all to a typical problem of survival analysis and then study it in terms of MPP Ψ taking the geometric aspects of the crystallization process into account.

In fact, let T be a nonnegative r.v. with cumulative density function F; then T may be considered as a random failure time, the function $S(t) := \mathbf{P}(T > t)$ is called the *survival function of* T and

$$\mathbf{P}(T \in [t, t + \Delta t) \,|\, T \ge t) = \frac{\mathbf{P}(t \le T < t + \Delta t)}{\mathbf{P}(T \ge t)} = \frac{F(t + \Delta t) - F(t)}{1 - F(t)}.$$

If T is a continuous r.v. with pdf f, then the following limit

$$h(t) := \lim_{\Delta t \downarrow 0} \frac{\mathbf{P}(T \in [t, t + \Delta t) \mid T \ge t)}{\Delta t}$$
 (5)

(with 0/0 := 0) exists and $h(t) = \frac{f(t)}{1 - F(t)}$, i.e. $h(t) = -\frac{d}{dt} \ln S(t)$.

Definition 5 h(t) is called the hazard function of T, while

$$\mathcal{H}(t) := -\int_0^t \frac{S(ds)}{S(s-)}$$

is known as the cumulative (or integrated) hazard function of T.

Remark 3 If T is a continuous r.v., then $\mathcal{H}(ds) = h(s)ds$.

If T is a discrete r.v. we can define a discrete hazard function $h(t) := \mathbf{P}(T = t \mid T \ge t)$, from which it follows $\mathcal{H}(t) = \sum_{s \le t} h(s)$.

We introduce now the so-called product-integral and its relation with hazard and survival functions. (For a complete and elementary treatment of the basic theory of the product-integral see [7] and [1]).

Definition 6 Let $X:[0,\infty)\to\mathbb{R}$ be a cadlag function of locally bounded variation. We define

$$Y(t) = \prod (1 + X(ds))$$

the product-integral of X over intervals of the form [0,t], as the following function:

$$Y(t) = \prod_{s \in [0,t]} (1 + X(ds)) := \lim_{\max|t_i - t_{i-1}| \to 0} \prod_i (1 - X(t_i) - X(t_{i-1})),$$

where $0 = t_0 < t_1 < \cdots < t_n = t$ is a partition of [0, t].

If X is a step-function, the product-integral becomes a finite product over the jump times of X, thus

$$Y = \prod (1 + \Delta X)$$

where $\Delta X = X - X_-$; if X is continuous, the product-integral is just the ordinary exponential $\prod (1 + dX) = e^X$.

Theorem 1 [1] $Y = \prod (1 + dX)$ exists and is a cadlag function of locally bounded variation. It is the unique solution to the integral equation

$$Y(t) = 1 + \int_{s \in [0,t]} Y(s-)X(ds).$$

From Definition 5 and Theorem 1, we have that, for a r.v. $T \geq 0$,

$$S(t) = \prod_{[0,t]} (1 - d\mathcal{H}).$$

Let us now consider the capture time T_x of point x and observe that

$$\mathbf{P}(T_x > t) = \mathbf{P}(x \notin \Theta^t) = S(t, x).$$

So $S(\cdot, x)$ is the survival function of the r.v. T_x and, in order to apply what we just said in terms of the MPP Ψ , we must go back to the "causes" of capture of point x. To this end we need to introduce the concept of causal cone (see e.g.[3],[10]).

Definition 7 The causal cone C(t,x) of a point x at time t is the space-time region in which at least one nucleation has to take place so that the point x is covered by crystals at time t:

$$C(t, x) := \{(s, y) \in [0, t] \times \mathbf{K} : x \in \Theta_s^t(y)\}.$$

We denote by $S_x(s,t)$ the section of the causal cone C(t,x) at time s < t,

$$S_x(s,t) := \{ y \in \mathbf{K} : (s,y) \in C(t,x) \} = \{ y \in \mathbf{K} : x \in \Theta_s^t(y) \}.$$

Under the assumption that the growth model is deterministic and for all t' > t is $\Theta_s^t(x) \subset \Theta_s^{t'}(x)$, we suppose further that the causal cone C(t,x) is well-defined for any x, t, and that the sections $S_x(s,t)$ are such that dim $\partial S_x(s,t) < d$ for any x,s,t.

Remark 4 From the definition of C(t, x) it easily follows that

$$S(t, x) = \mathbf{P}(\Psi(\mathcal{C}(t, x)) = 0).$$

Moreover, from the simplicity of the marginal process $\tilde{\Psi}$, we have, formally,

$$\begin{split} Q(t,B)\tilde{\Lambda}(dt) &= \Lambda(dt \times B) \\ &= \mathbf{E}(\Psi(dt \times B)) \\ &= \mathbf{P}(\Psi(dt \times B) = 1) \\ &= \mathbf{P}(X_1 \in B \,|\, \tilde{\Psi}(dt) = 1) \mathbf{P}(\tilde{\Psi}(dt) = 1) \\ &= \mathbf{P}(X_1 \in B \,|\, \tilde{\Psi}(dt) = 1) \tilde{\Lambda}(dt) \end{split}$$

Therefore the mark distribution Q(t, B) represents the probability that a nucleus $X_1 \in B$, given that it was born during [t, t+dt). The hypothesis of absolute continuity of Q implies that nuclei cannot be born in a set B with dim B < d; in particular $Q(s, \partial \mathcal{S}_x(s, t)) = 0 \ \forall s < t$ and so T_x is a continuous r.v..

In fact, by absurd let T_x not be a continuous r.v.; then there exists $t \in \mathbb{R}^+$ such that $\mathbf{P}(T_x = t) > 0$. Observe that

$$\mathbf{P}(T_x = t) > 0 \Leftrightarrow \mathbf{P}(\Psi(\partial \mathcal{C}(t, x)) \neq 0) > 0 \Leftrightarrow \mathbf{E}(\Psi(\partial \mathcal{C}(t, x))) > 0.$$

But,

$$\mathbf{E}(\Psi(\partial\mathcal{C}(t,x))) = \int_0^t \int_{\partial\mathcal{S}_x(s,t)} Q(s,dy) \tilde{\Lambda}(ds) = 0.$$

Since in the continuous case the product-integral coincides with the ordinary exponential, we have

$$S(t,x) = \exp\left\{-\int_0^t h(s,x)ds\right\},\,$$

where

$$h(t,x) = \lim_{\Delta t \downarrow 0} \frac{\mathbf{P}(T_x \in [t, t + \Delta t) \mid T_x \ge t)}{\Delta t}$$

is the hazard function of T_x , as defined by (5).

By the continuity of T_x ,

$$h(t,x) = \lim_{\Delta t \downarrow 0} \frac{\mathbf{P}(T_x \in (t, t + \Delta t] \mid T_x > t)}{\Delta t}$$

$$= \lim_{\Delta t \downarrow 0} \frac{\mathbf{P}(x \in (\Theta^{t + \Delta t} \setminus \Theta^t) \mid x \notin \Theta^t)}{\Delta t}$$

$$= \lim_{\Delta t \downarrow 0} \frac{\mathbf{P}(x \in \Theta^{t + \Delta t} \mid x \notin \Theta^t)}{\Delta t}.$$

Definition 8 For all $x \in \mathbf{K}$, the function $h(\cdot, x)$ given by

$$h(t,x) := \lim_{\Delta t \downarrow 0} \frac{\mathbf{P}(x \in \Theta^{t+\Delta t} \,|\, x \not\in \Theta^t)}{\Delta t}$$

is called the hazard function associated with point x.

Note now that, formally,

$$\mathcal{H}(ds, x) = h(s, x)ds = \mathbf{P}(x \in \Theta^{s+ds} \mid x \notin \Theta^s)$$

$$= \mathbf{P}(\Psi[\mathcal{C}(s + ds, x) \setminus \mathcal{C}(s, x)] > 0 \mid \Psi(\mathcal{C}(s, x)) = 0). \quad (6)$$

Unfortunately, Eq. (6) does not provide an explicit dependence of h upon the kinetic parameters of the process.

In fact, since $C(s + ds, x) \setminus C(s, x) \not\subseteq (s, s + ds] \times \mathbf{K}$, it is not possible simplify (6) by using the hypothesis of simplicity of $\tilde{\Psi}$.

We show now that, by defining a suitable r.v., it is possible to write again the survival function as a product-integral for any fixed t, but so that the properties of Ψ are taken into account.

Definition 9 For any fixed $t \in \mathbb{R}^+$, let $N_{t,x}$ be the positive r.v. defined in the following way:

$$N_{t,x} := \left\{ \begin{array}{cc} s & \text{if the first nucleus in } \mathcal{C}(t,x) \text{ was born at time } s \\ +\infty & \text{if no nucleus was born in } \mathcal{C}(t,x) \end{array} \right.$$

We observe that:

- since $N_{t,x}$ is the birth time of the first nucleus in C(t,x), it is well defined also for processes where new nuclei cannot be born in an already crystallized zone;
- $N_{t,x}$ is a discrete r.v. \iff the intensity measure $\tilde{\Lambda}$ of $\tilde{\Psi}$ is discrete.

In fact, $\mathbf{P}(N_{t,x}=s) \neq 0$ if and only if

$$\mathbf{P}(\{\Psi(\{s\} \times \mathcal{S}_x(s,t)) = 1\} \cap \{\Psi(\mathcal{C}(s-,\mathcal{S}_x(s-,t)) = 0\}).$$

By observing that

$$1 \ge \tilde{\Psi}(\{s\}) = \Psi(\{s\} \times \mathbf{K}) \ge \Psi(\{s\} \times \mathcal{S}_x(s,t)),$$

we can use the hypothesis that $\tilde{\Psi}$ is simple and write

$$\mathbf{P}(\Psi(\lbrace s\rbrace \times \mathcal{S}_x(s,t)) = 1) = \tilde{\Lambda}(\lbrace s\rbrace)Q(s,\mathcal{S}_x(s,t)).$$

Thus, by assuming $Q(s, \mathcal{S}_x(s,t)) \neq 0$, it follows that $\tilde{\Lambda}(\{s\}) \neq 0$. As a consequence, if $\tilde{\Lambda}$ is continuous, then $N_{t,x}$ is continuous on [0,t].

Since $S(t,x) = \mathbf{P}(\Psi(\mathcal{C}(t,x)) = 0) = \mathbf{P}(N_{t,x} = +\infty) = \mathbf{P}(N_{t,x} > t)$, we have that, for any fixed $t \in \mathbb{R}^+$,

$$S(t,x) = \prod_{s \in [0,t]} (1 - \mathcal{H}_{N_{t,x}}(ds,x)),$$

where $\mathcal{H}_{N_{t,x}}(ds,x)$) is the cumulative hazard function of $N_{t,x}$.

Now we distinguish between $\tilde{\Lambda}$ discrete and continuous.

1. Let $\tilde{\Lambda}$ be discrete, and so also $\tilde{\nu}$. Then

$$\mathbf{P}(N_{t,x} > t) = \prod_{s \in [0,t]} [1 - \mathbf{P}(N_{t,x} = s \mid N_{t,x} \ge s)].$$
 (7)

In terms of the nucleation process Ψ ,

$$\mathbf{P}(N_{t,x} = s \mid N \ge s) = \mathbf{P}(\Psi[\{s\} \times \mathcal{S}_x(s,t)] = 1 \mid \Psi[\mathcal{C}(s-,\mathcal{S}_x(s-,t))] = 0)$$

$$= \mathbf{E}(\Psi[\{s\} \times \mathcal{S}_x(s,t)] \mid \Psi[\mathcal{C}(s-,\mathcal{S}_x(s-,t))] = 0)$$

$$= \mathbf{E}(\nu[\{s\} \times \mathcal{S}_x(s,t)] \mid \Psi[\mathcal{C}(s-,\mathcal{S}_x(s-,t))] = 0).$$
(8)

Therefore,

$$S(t,x) = \prod_{s \in [0,t]} [1 - \mathbf{E}(\tilde{\nu}(\{s\})k(s, \mathcal{S}_x(s,t)) | \Psi[\mathcal{C}(s-, \mathcal{S}_x(s-,t))] = 0)]. \tag{9}$$

Remark 5 If Ψ is a process with *independent increments* with respect to its history, then the behavior of Ψ in an interval (s,t] does not depend on what happened up to time s for all $s,t \in \mathbb{R}^+$ with s < t. So the expression (8) is equal to $\mathbf{P}(\Psi[\{s\} \times \mathcal{S}_x(s,t)] = 1)$ and we obtain

$$S(t,x) = \prod_{s \in [0,t]} [1 - \tilde{\Lambda}(\{s\})Q(s, \mathcal{S}_x(s,t))].$$

2. Let $\tilde{\Lambda}$ be continuous, and so also $\tilde{\nu}$. Then

$$\mathbf{P}(N_{t,x} > t) = \prod_{[0,t]} [1 - \mathbf{P}(s \le N_{t,x} < s + ds \mid N_{t,x} \ge s)]$$

$$= \prod_{[0,t]} [1 - \mathbf{P}(s < N_{t,x} \le s + ds \mid N_{t,x} > s)]$$

$$= \prod_{[0,t]} [1 - \mathbf{P}(\Psi[(s,s + ds] \times \mathcal{S}_x(s,t)] > 0 \mid \Psi[\mathcal{C}(s,\mathcal{S}_x(s,t))] = 0)]$$

$$= \prod_{[0,t]} [1 - \mathbf{E}(\Psi[(s,s + ds] \times \mathcal{S}_x(s,t)] \mid \Psi[\mathcal{C}(s,\mathcal{S}_x(s,t))] = 0)]$$

$$= \prod_{[0,t]} [1 - \mathbf{E}(\tilde{\nu}(ds)k(s,\mathcal{S}_x(s,t)) \mid \Psi[\mathcal{C}(s,\mathcal{S}_x(s,t))] = 0)].$$

Since $N_{t,x}$ is continuous on [0,t], we obtain that, for any fixed t,

$$S(t,x) = \exp\left\{-\int_0^t \mathbf{E}(\tilde{\nu}(ds)k(s,\mathcal{S}_x(s,t)) \mid \Psi[\mathcal{C}(s,\mathcal{S}_x(s,t))] = 0)\right\}. \tag{11}$$

Remark 6 Let us suppose that the compensator ν of Ψ is of the form (see Remark 1)

$$\nu(dt \times dx) = k_0(t, dx)\tilde{\nu}(dt)(1 - \mathbf{1}_{\Theta^{t-}}(x)).$$

It is clear that the probabilities in (7) and (10) are the same probabilities we have when we consider the free process Ψ_0 with compensator $\nu_0(dt \times dx) = k_0(t, dx)\tilde{\nu}(dt)$. So, there is no distinction between Ψ and Ψ_0 in computing the survival function.

Please note that $\nu_0(dt \times dx)$ is still assumed in general to be \mathcal{F}_{t-} -measurable (see Remark 2).

Therefore expressions (9) and (11) represent the survival function of a point x at time t in the case of a nucleation process with a compensator that in general can be stochastic.

If additional we assume that the (free) process Φ has independent increments and $\tilde{\Lambda}$ is continuous, by (11) it follows that

$$S(t,x) = e^{-\int_0^t \tilde{\Lambda}(ds)Q(\mathcal{S}_x(s,t))} = e^{-\Lambda(\mathcal{C}(t,x))}$$
(12)

and, consequently,

$$h(t,x) = \frac{\partial}{\partial t} \Lambda(\mathcal{C}(t,x)), \tag{13}$$

if the derivative exists.

So we obtain the same results as when Ψ is a nucleation process of the Poisson type (see [5],[12],[13]). In fact [11], the independence of increments implies that the compensator of Ψ is deterministic and coincides with the intensity measure Λ . Since a point process with a deterministic and continuous compensator is a Poisson process, then Ψ is a marked Poisson process.

Note that, in this case, Eq. (12) could be obtained as a direct consequence of the following theorem [11]:

Theorem 2 Assume that $(\mathbf{X}, \mathcal{X})$ is a c.s.m.s. and let Λ be a measure on $\mathbb{R}^+ \times \mathbf{X}$ such that $\Lambda(\cdot \times \mathbf{X})$ is continuous and locally bounded. Suppose that Φ is a marked Poisson process with mark space \mathbf{X} and intensity measure Λ . Then Φ (considered as a random measure on $\mathbb{R}^+ \times \mathbf{X}$) is a Poisson process with intensity measure Λ .

We now introduce the concept of *individual* hazard function and show that it coincides with the hazard function h(t,x) when the free process Ψ_0 is a marked Poisson process. This result was already proved in [4]; here we give a more immediate proof by using the theorem above.

Definition 10 The individual hazard function $h_1(t,x)$ is the capture rate of point x at time t by a single crystal, i.e.

$$h_1(t,x) := \lim_{\Delta t \mid 0} \frac{\mathbf{P}(\exists ! j \in \mathbb{N} : x \in (\Theta_j^{t+\Delta t} \setminus \Theta^t) \mid x \notin \Theta^t)}{\Delta t}.$$

Proposition 1 Let Ψ be a nucleation process such that Ψ_0 is a marked Poisson process with intensity measure Λ . Then

$$h(t,x) = h_1(t,x)$$

for all $x \in \mathbf{K}$ and for all $t \in \mathbb{R}^+$.

Proof. Let $h_0(t,x)$ and $h_{0,1}(t,x)$ be respectively the hazard function and the individual hazard function associated with the free process Ψ_0 . Since $(\mathcal{C}(t+dt,x)\setminus\mathcal{C}(t,x))\cap\mathcal{C}(t,x)=\emptyset$ and Ψ_0 (Poisson process on $\mathbb{R}^+\times\mathbf{K}$) has independent increments, we have

$$h(t,x)dt = h_0(t,x)dt$$

$$= 1 - \mathbf{P}(\Psi_0[\mathcal{C}(t+dt,x) \setminus \mathcal{C}(t,x)] = 0 \mid \Psi_0(\mathcal{C}(t,x)) = 0)$$

$$= 1 - \mathbf{P}(\Psi_0[\mathcal{C}(t+dt,x) \setminus \mathcal{C}(t,x)] = 0)$$

$$= 1 - e^{-\Lambda(\mathcal{C}(t+dt,x) \setminus \mathcal{C}(t,x))}$$

$$= \Lambda(\mathcal{C}(t+dt,x) \setminus \mathcal{C}(t,x)),$$

as in (13).

Denote by $h_k(t, x)$ and $h_{0,k}(t, x)$ the capture rate (for Ψ and Ψ_0 respectively) of x at time t by exactly k crystals. Thus we have

$$h_{0,k}(t,x)dt = \mathbf{P}(\Psi_0[\mathcal{C}(t+dt,x) \setminus \mathcal{C}(t,x)] = k \mid \Psi_0(\mathcal{C}(t,x)) = 0)$$

$$= \mathbf{P}(\Psi_0[\mathcal{C}(t+dt,x) \setminus \mathcal{C}(t,x)] = k)$$

$$= \frac{[\Lambda(\mathcal{C}(t+dt,x) \setminus \mathcal{C}(t,x))]^k}{k!} e^{-\Lambda(\mathcal{C}(t+dt,x) \setminus \mathcal{C}(t,x))}$$

$$= \frac{[h(t,x)dt]^k}{k!} (1 - h(t,x)dt) = O(h(t,x)dt)^k$$

Observe that:

for k = 1,

$$\mathbf{P}(\Psi[\mathcal{C}(t+dt,x)\setminus\mathcal{C}(t,x)] = 0 \mid \Psi(\mathcal{C}(t,x)) = 0) = \mathbf{P}(\Psi_0[\mathcal{C}(t+dt,x)\setminus\mathcal{C}(t,x)] = 0 \mid \Psi_0(\mathcal{C}(t,x)) = 0),$$

since the first nucleus is free to be born in $C(t+dt,x) \setminus C(t,x)$. So

$$h_1(t,x) = h_{0,1}(t,x) = \lim_{dt \downarrow 0} \frac{h(t,x)dt}{dt} - \frac{h(t,x)^2 dt^2}{dt} = h(t,x);$$

for $k \geq 2$,

$$\{\omega: \Psi[\mathcal{C}(t+dt,x)\setminus \mathcal{C}(t,x)] = k)\} \subseteq \{\omega: \Psi_0[\mathcal{C}(t+dt,x)\setminus \mathcal{C}(t,x)] = k)\}.$$
 So, $h_k(t,x) \le h_{0,k}(t,x) = 0$.

3 A Relation between the Hazard Function and the Local Spherical Contact Distribution Function

In the previous section we saw that the survival function S(t,x) can be obtained by the knowledge of the hazard function h(t,x) of capture time T_x of a point x.

From the survival analysis theory, we know that a natural estimator of the survival function is the Kaplan-Meier estimator, and it turns out to be the product-integral of the empirical cumulative hazard function (known as Nelson-Aalen estimator). So, if $T_{x,1}, \ldots, T_{x,n}$ is an IID sample of T_x , the Kaplan-Meier estimator $\hat{S} = \prod (1 - d\hat{\mathcal{H}})$ of S(t,x) is given by

$$\hat{S}(x,t) = \prod_{s \le t} \left(1 - \frac{\#\{i : T_{x,i} = s\}}{\#\{i : T_{x,i} \ge s\}} \right).$$

Observe that if we have got n realizations of Ψ , and so of Θ , then each $T_{x,i}$ is the capture time of the point x in the ith realization Θ_i . Thus $T_{x,1}, \ldots, T_{x,n}$ are IID as T_x .

Problems arise when we have only one realization of Θ . In fact, even if S does not depend on x (spatial homogeneity), we can not consider n random points x_1, \ldots, x_n in the space and their respective life time T_1, \ldots, T_n , because the independence of these should be lost. For example, let d be the distance between two points x_1 and x_2 , and suppose that the growth speed G of crystals is radial and constant. If we observe that x_1 is covered at time \bar{t} , then $\mathbf{P}(T_2 \leq \bar{t} + d/G) = 1$, since at time $\bar{t} + d/G$ the point x_2 is certainly captured by the crystal in x_1 at time \bar{t} (for a similar discussion see [2]).

As we already said, for any fixed t, Θ^t is a random closed set. From stochastic geometry we know that contact distributions are important tools to describe certain aspects of random closed sets and can be easily estimated. Now we show how the hazard function is related to local spherical contact distribution function [8].

Definition 11 The local spherical contact distribution function H_S of an inhomogeneous random set Ξ is given by

$$H_S(r,x) := \mathbf{P}(x \in \Xi \oplus b(0,r) \mid x \notin \Xi),$$

where \oplus is Minkowski addition and b(0,r) is a d-dimensional ball of radius r centered at the origin.

Proposition 2 Let Ψ a nucleation process as in hypotheses on page 3 and K such that dim $K \geq 2$. Assume that the growth speed G of crystals is radial and constant, and, for any t, x fixed, denote by $\tilde{H}(\cdot, t, x)$ the following function

$$\tilde{H}(\tau, t, x) := H_{S \Theta^t}(G\tau, x),$$

where H_{S,Θ^t} is the local spherical contact distribution function of the random closed set Θ^t .

$$h(t,x) = \frac{\partial}{\partial \tau} \tilde{H}(\tau,t,x)|_{\tau=0} \quad \text{for a.e. } x.$$
 (14)

Proof. Write the random set $\Theta^{t+\Delta t}$ as union of the set Θ^t grown up to time $t+\Delta t$ and the crystals born during the interval $(t, t+\Delta t]$:

$$\begin{split} \Theta^{t+\Delta t} &= \bigcup_{\substack{j \\ T_j \leq t + \Delta t}} \Theta^{t+\Delta t}_{T_j}(X_j) \\ &= \bigcup_{\substack{T_j \leq t \\ T_j \leq t}} \Theta^{t+\Delta t}_{T_j}(X_j) \ \cup \bigcup_{\substack{t < T_j \leq t + \Delta t \\ t < T_j \leq t + \Delta t}} \Theta^{t+\Delta t}_{T_j}(X_j) \\ &= \Theta^t \oplus \Delta t G B^d \ \cup \bigcup_{\substack{t < T_j \leq t + \Delta t \\ t < T_j \leq t + \Delta t}} \Theta^{t+\Delta t}_{T_j}(X_j), \end{split}$$

where $B^d = b(0, 1)$.

Let $\tilde{\mathcal{C}}_x(t,t+\Delta t)$ be the subset of the causal cone $\mathcal{C}(t,x)$ given by

$$\tilde{\mathcal{C}}_x(t, t + \Delta t) := \{ (s, y) \in (t, t + \Delta t] \times \mathbf{K} \mid x \in \Theta_s^{t + \Delta t}(y) \}.$$

By this definition it follows that

$$\nu_{d+1}(\tilde{\mathcal{C}}_x(t,t+\Delta t)) = \int_t^{t+\Delta t} \left(\int_{b(x,(t+\Delta t-s)G)} dy \right) ds = O(\Delta t)^{d+1}$$

(where ν_{d+1} is the d+1-dimensional Lebesgue measure) and

$$\mathbf{P}(\Psi(\tilde{\mathcal{C}}_x(t,t+\Delta t)) > 0 \mid \Psi(\mathcal{C}(t,x)) = 0) = \mathbf{E}(\Psi(\tilde{\mathcal{C}}_x(t,t+\Delta t)) \mid \Psi(\mathcal{C}(t,x)) = 0).$$

So we have

$$\mathbf{P}(x \in \Theta^{t+\Delta t} \mid x \notin \Theta^{t}) = 1 - \mathbf{P}(x \notin (\Theta^{t} \oplus \Delta t G B^{d} \cup \bigcup_{t < T_{j} \leq t + \Delta t} \Theta^{t+\Delta t}(X_{j})) \mid x \notin \Theta^{t})$$

$$= 1 - \mathbf{P}(\{x \notin \Theta^{t} \oplus \Delta t G B^{d}\}) \cap \{x \notin \bigcup_{t < T_{j} \leq t + \Delta t} \Theta^{t+\Delta t}_{T_{j}}(X_{j})\} \mid x \notin \Theta^{t})$$

$$= 1 - \mathbf{P}(x \notin (\Theta^{t} \oplus \Delta t G B^{d}) \mid x \notin \Theta^{t}) \mathbf{P}(x \notin \bigcup_{t < T_{j} \leq t + \Delta t} \Theta^{t+\Delta t}_{T_{j}}(X_{j}) \mid x \notin \Theta^{t})$$

$$= 1 - [1 - H_{S,\Theta^{t}}(G \Delta t, x)] (\mathbf{P}(\Psi(\tilde{\mathcal{C}}_{x}(t, t + \Delta t)) = 0 \mid \Psi(\mathcal{C}(t, x)) = 0)$$

$$= 1 - [1 - \tilde{H}(\Delta t, t, x)] [1 - \mathbf{P}(\Psi(\tilde{\mathcal{C}}_{x}(t, t + \Delta t)) \mid \Psi(\mathcal{C}(t, x)) = 0)]$$

$$= 1 - [1 - \tilde{H}(\Delta t, t, x)] [1 - \mathbf{E}(\Psi(\tilde{\mathcal{C}}_{x}(t, t + \Delta t)) \mid \Psi(\mathcal{C}(t, x)) = 0)].$$

By replacing in the definition of h(t, x), we obtain

$$h(t,x) = \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \{ 1 - [1 - \tilde{H}(\Delta t, t, x)] \cdot [1 - \mathbf{E}(\Psi(\tilde{\mathcal{C}}_x(t, t + \Delta t)) \mid \Psi(\mathcal{C}(t, x)) = 0)] \}. \tag{15}$$

Now observe that:

$$\bullet \ \lim_{\Delta t \downarrow 0} \frac{\tilde{H}(\Delta t, t, x)}{\Delta t} = \lim_{\Delta t \downarrow 0} \frac{\tilde{H}(\Delta t, t, x) - \tilde{H}(0, t, x)}{\Delta t} = \frac{\partial}{\partial \tau} \tilde{H}(\tau, t, x)_{|_{\tau=0}}.$$

• since $\Psi(\tilde{\mathcal{C}}_x(t,t+\Delta t))$ is a nonnegative r.v.,

$$\begin{split} \lim_{\Delta t\downarrow 0} \frac{\mathbf{E}(\Psi(\tilde{\mathcal{C}}_x(t,t+\Delta t)) \mid \Psi(\mathcal{C}(t,x)) = 0)}{\Delta t} &= \lim_{\Delta t\downarrow 0} \frac{1}{\Delta t} \frac{\mathbf{E}(\Psi(\tilde{\mathcal{C}}_x(t,t+\Delta t)) \mathbf{1} \{\Psi(\mathcal{C}(t,x)) = 0\})}{\mathbf{P}(\Psi(\mathcal{C}(t,x)) = 0)} \\ &= \leq \frac{1}{\mathbf{P}(\Psi(\mathcal{C}(t,x)) = 0)} \lim_{\Delta t\downarrow 0} \frac{\mathbf{E}[\Psi(\tilde{\mathcal{C}}_x(t,t+\Delta t))]}{\Delta t} \\ &= \frac{1}{\mathbf{P}(\Psi(\mathcal{C}(t,x)) = 0)} \lim_{\Delta t\downarrow 0} \frac{\Lambda(\tilde{\mathcal{C}}_x(t,t+\Delta t))}{\Delta t}. \end{split}$$

We know that $\Lambda(ds \times dy) = Q(s, dy)\tilde{\Lambda}(ds)$, where $\tilde{\Lambda}$ is the intensity measure of the marginal process and Q is a stochastic kernel from \mathbb{R}^+ to **K**. Since for all $s \in \mathbb{R}^+$, $Q(s,\cdot)$ is absolutely continuous with respect to the d-dimensional Lebesgue measure, then, for a.e. $x \in \mathbf{K}$,

$$\Lambda(\tilde{\mathcal{C}}_x(t,t+\Delta t)) = \int_t^{t+\Delta t} \int_{b(x,(t+\Delta t-s)G)} Q(s,dy) \tilde{\Lambda}(ds) = O(\Delta t)^d \int_t^{t+\Delta t} \tilde{\Lambda}(ds).$$

(From now on it is understood that $x \in \mathbf{K}$ a.s.). Since certainly $\int_t^{t+\Delta t} \tilde{\Lambda}(ds) < \infty$, we have that $\Lambda(\tilde{\mathcal{C}}_x(t,t+\Delta t)) = O(\Delta t)^d$. Therefore, $\forall d \geq 2$

$$\lim_{\Delta t \downarrow 0} \frac{\Lambda(\tilde{\mathcal{C}}_x(t, t + \Delta t))}{\Delta t} = 0.$$

By (15) we obtain

$$h(t,x) = \frac{\partial}{\partial \tau} \tilde{H}(\tau,t,x)|_{\tau=0}$$

1. Expression (14) may be intuitively explained in this way: capture of point xduring interval $(t, t + \Delta t]$ can be determined both by the growth of Θ^t and by the birth of new crystals. These two causes are conditionally independent and the second one has a negligible weight because of the simplicity of $\tilde{\Psi}$ and the absolute continuity of mark distribution.

- 2. If $\Lambda \ll \nu_{d+1}$, i.e. $\tilde{\Lambda}$ absolutely continuous, then $\Lambda(\tilde{\mathcal{C}}_x(t,t+\Delta t)) = O(\Delta t)^{d+1}$ and (14) is true also for dim $\mathbf{K} = 1$.
- 3. We give now a simple example about the validity of Eq.(14). Let us suppose that nuclei are born in time in accordance with a homogeneous Poisson point process with intensity μ and uniformly in space. So, the mark distribution Q is the uniform distribution on **K** which we assume has finite measure K. Consequently,

$$\nu(dt \times dx) = \Lambda(dt \times dx) = \frac{\mu}{K} dt dx,$$

and the survival function, independent of x, is given by

$$S(t) = \exp\left\{-\frac{\mu}{K} \int_0^t b_d (t-s)^d G^d ds\right\},\,$$

where b_d is the volume of the unit ball in \mathbb{R}^d . Thus,

$$h(t) = -\frac{d}{dt} \ln S(t) = \frac{\mu}{K} \int_0^t db_d G^d(t-s)^{d-1} ds.$$
 (16)

Observe now that, for every fixed t, Θ^t is a homogeneous Boolean model with intensity

$$\begin{array}{lll} \lambda & = & \text{mean number of nuclei for unit of volume} \\ & = & \nu([0,t]\times[0,1]^d) = \mu t Q([0,1]^d) = \mu t \frac{1}{\mathbf{K}} \end{array}$$

and typical grain $Z_0 = b(0, G(t - U))$ with $U \sim U([0, t])$.

Since the spherical contact distribution function of a homogeneous Boolean model is given by

$$H_S(r) = 1 - \exp\left\{-\lambda \sum_{j=1}^d b_j r^j \mathbf{E}(V_{d-j}(Z_0))\right\},\,$$

where b_j is the volume of the unit ball in \mathbb{R}^j and $V_j(Z_0)$ is the jth intrinsic volume of Z_0 , we have that

$$\tilde{H}(\tau,t) = 1 - \exp\left\{-\frac{\mu t}{K} \sum_{j=1}^{d} b_j (G\tau)^j \mathbf{E}(V_{d-j}(Z_0))\right\}.$$

Since $b_1 = 2$ and $\mathbf{E}(V_{d-1}(Z_0)) = \frac{1}{2} \int_0^t db_d G^{d-1}(t-s)^{d-1} \frac{1}{t} ds$, it follows that

$$h(t) = \frac{\partial}{\partial \tau} \tilde{H}(\tau, t, x)|_{\tau=0} = \frac{\mu}{K} \int_0^t db_d G^d(t-s)^{d-1} ds,$$

which coincides with the hazard function (16) obtained from the survival function.

4. Let us suppose that G time dependent. In this case H is given by

$$\tilde{H}(\tau, t, x) := H_{S,\Theta^t} \left(\int_t^{t+\tau} G(s) ds, x \right),$$

and $S_x(s, t + \Delta t) = b(x, \int_s^{t+\Delta t} G(u) du)$. Again $\Lambda(\tilde{C}_x(t, t + \Delta t)) = O(\Delta t)^d$, so

$$h(t,x) = \frac{\partial}{\partial \tau} \tilde{H}(\tau,t,x)|_{\tau=0} = G(t) \frac{\partial}{\partial r} H_{S,\Theta^t}(r,x)|_{r=0}.$$

5. If G is space-and-time dependent, we can not define a function \tilde{H} similarly as in the previous cases, because Θ^t does not grow in a homogeneous way during a time interval $(t, t+\tau]$. But, as follows by the proof, the hazard function h(t, x) coincides with the capture rate of x determined by the growth of Θ^t . In fact, also in this case $\Lambda(\tilde{C}_x(t, t + \Delta t)) = O(\Delta t)^d$, since certainly $S_x(s, t + \Delta t) \subseteq b(x, (t + \Delta t - s)G_M)$, where $G_M = \max G(u, y)$ for $(u, y) \in [s, t + \Delta t] \times S_x(s, t + \Delta t)$. So, if we denote by $\Theta(t + \Delta t)$ the random closed set Θ^t grown up to time $t + \Delta t$, then

$$h(t,x) = \lim_{\Delta t \downarrow 0} \frac{\mathbf{P}(x \in \Theta(t + \Delta t) \mid x \notin \Theta^t)}{\Delta t}.$$
 (17)

By the continuity of G, for Δt sufficiently small, during interval $[t, t + \Delta t]$ we may consider $G = G(t, \cdot)$, so that it depends only on space.

Observe now that when the growth speed of crystals is dependent only on time or space (i.e. G = G(t) or G = G(x)), then

$$x \in \Theta_s^t(y) \iff y \in \Theta_s^t(x),$$
 (18)

while if G = G(t, x) it is not true in general.

Since every point $y \in \partial \Theta^t$ may be seen as nucleus of a crystal born at time t, then $x \in \Theta(t + \Delta t) \iff \exists y \in \partial \Theta^t$ such that $x \in \Theta_t^{t + \Delta t}(y)$. By (18) we have

$$x \in \Theta(t + \Delta t) \iff \Theta^t \cap \Theta_t^{t + \Delta t}(x) \neq \emptyset.$$

Continuity of G allows to replace $\Theta_t^{t+\Delta t}(x)$ with $b(x, G(t, x)\Delta t)$, thus

$$x \in \Theta(t + \Delta t) \iff \Theta^t \cap b(x, G(t, x)\Delta t) \neq \emptyset,$$

and by (17) we obtain

$$h(t,x) = \lim_{\Delta t \downarrow 0} \frac{\mathbf{P}(x \in \Theta^t \oplus G(t,x)\Delta t B^d \mid x \notin \Theta^t)}{\Delta t}$$

$$= \lim_{\Delta t \downarrow 0} \frac{H_{S,\Theta^t}(G(t,x)\Delta t,x)}{\Delta t}$$

$$= G(t,x) \frac{\partial}{\partial r} H_{S,\Theta^t}(r,x)|_{r=0}.$$

Acknowledgements It is a pleasure to acknowledge useful discussions with A.Micheletti of MIRIAM & Dept. of Mathematics of the University of Milan.

Financial support is acknowledged to the Italian MIUR/cofin. Programme "Stochastic Processes with Spatial Structure" and to the Italian CNR contract n. 98.03635.ST74

References

- [1] Andersen, P.K.; Borgan, O.; Gill, R.D.; Keiding, N. Statistical Models Based on Counting Processes, 1993; Springer, New York.
- [2] Baddeley, A.J. A crash course in stochastic geometry, in Stochastic Geometry. Likelihood and Computation (O.E. Barndorff-Nielsen et al. Eds.), 1999; Chapman&Hall/CRC, Boca Raton.
- [3] Capasso, V. Ed. Mathematical Modelling for Polymer Processing. Polymerization, Crystallization, Manufacturing. Mathematics in Industry, vol.2, 2003; Springer-Verlag, Heidelberg.
- [4] Capasso, V.; Micheletti, A.; Burger, M. Densities of n-facets of incomplete Johnson-Mehl tessellations generated by inhomogeneous birth-and-growth processes. *Quaderno n.38*, *Dip. di Matematica*, *Università di Milano*, 2001.

- [5] Capasso, V.; Salani, C. Stochastic birth-and-growth processes modelling crystallization of polymers with spacially heterogeneous parameters. *Nonlinear Analysis RWA*, 2000; 1, 485–498.
- [6] Daley, D.J.; Vere-Jones, D. An Introduction to the Theory of Point Processes, 1998; Springer, New York.
- [7] Gill, R.D.; Johansen, S. A survey of product-integration with a view toward application in survival analysis. *Ann. Statist.*, **1990**; 18, 1501–1555.
- [8] Hug, D.; Last, G.; Weil, W. A survey on contact distributions, 2002. Preprint.
- [9] Karr A.F. Point Processes and Their Statistical Inference, 1991; Marcel Dekker, New York.
- [10] Kolmogorov, A.N. On the statistical theory of the crystallization of metals. *Bull. Acad.Sci. USSR*, *Math.Ser.*, **1937**; 1, 355–359.
- [11] Last, G.; Brandt, A. Marked Point Processes on the Real Line. The Dynamic Approach, 1995; Springer, New York.
- [12] Møller, J. Random Johnson-Mehl tessellations. Adv. Appl. Prob., 1992; 24, 814-844.
- [13] Møller, J. Topics in Voronoi and Johnson-Mehl tessellations, in Stochastic Geometry. Likelihood and Computation (O.E. Barndorff-Nielsen et al. Eds.), 1999; Chapman&Hall/CRC, Boca Raton.
- [14] Stoyan, D.; Kendall, W.S.; Mecke, J. Stochastic Geometry and its Application, 1995; John Wiley & Sons, New York.