

**UCLA**

**Department of Statistics Papers**

**Title**

Rescaling Marked Point Processes

**Permalink**

<https://escholarship.org/uc/item/58g1z40k>

**Authors**

Zaliapin

I. V.

Y. Y. Kagan

et al.

**Publication Date**

2011-10-25

# Rescaling Marked Point Processes

David Vere-Jones<sup>1</sup>, *Victoria University of Wellington and Statistics Research Associates Ltd, Wellington, New Zealand.*

Frederic Paik Schoenberg<sup>2</sup>, *University of California, Los Angeles*

## Abstract

In 1971, Meyer showed how one could use the compensator to rescale a multivariate point process, forming independent Poisson processes with intensity one. Meyer's result has been generalized to multi-dimensional point processes. Here, we explore generalization of Meyer's theorem to the case of marked point processes, where the mark space may be quite general. Assuming simplicity and the existence of a conditional intensity, we show that a marked point process can be transformed into a compound Poisson process with unit total rate and a fixed mark distribution.

Key words: residual analysis, random time change, Meyer's theorem, model evaluation, intensity, compensator, Poisson process.

<sup>1</sup> Department of Mathematics, Victoria University of Wellington, P.O. Box 196, Wellington, New Zealand. [dvj@mcs.vuw.ac.nz](mailto:dvj@mcs.vuw.ac.nz)

<sup>2</sup> Department of Statistics, 8142 Math-Science Building, University of California, Los Angeles, 90095-1554. [frederic@stat.ucla.edu](mailto:frederic@stat.ucla.edu)

## 1 Introduction.

Before other matters, both authors would like to express their appreciation to Daryl for his stimulating and forgiving company, and to wish him a long and fruitful continuation of his mathematical, musical, woodworking, and many other activities. In particular, it is a real pleasure for the first author to acknowledge his gratitude to Daryl for his hard work, good humour, generosity and continuing friendship throughout the development of (innumerable draft versions and now even two editions of) their joint text on point processes.

The topic of the present paper was broached more than thirty years ago in papers by Knight (1970) and Meyer (1971). Knight (1970) showed that for any orthogonal sequence of continuous local martingales, by rescaling time for each via its associated predictable process, one forms a multivariate sequence of independent standard Brownian motions. Then Meyer (1971) extended Knight's theorem to the case of point processes, showing that given a simple multivariate point process  $\{N_i; i = 1, 2, \dots, n\}$ , the multivariate point process obtained by rescaling each  $N_i$  according to its compensator is a sequence of independent Poisson processes, each having intensity one. Since then, alternative proofs and variations of this result have been given by Papangelou (1972), Brémaud (1972), Aalen and Hoem (1978), Kurtz (1980), and Brown and Nair (1988). Papangelou (1972) gave the following interpretation in the univariate case: "Roughly, moving in  $[0, \infty)$  so as to meet expected future points at a rate of one per time unit (given at each instant complete knowledge of the past), we meet them at the times of a Poisson process."

Immediately after the statement of his theorem, Meyer remarked "On peut se demander si ce résultat est vraiment plus qu'une curiosité mathématique!" The result has nevertheless

proven to be extremely useful for point process analysis, providing the basis for point process residual analysis which is a very powerful model evaluation tool (Ogata, 1988; Ogata, 2001; Schoenberg, 1997) and which has also been used for point process prediction (Ogata, 1988; Ogata, 1993).

Generalizations of Meyer's result to point processes on  $\mathbf{R}^d$  have been established by Merzbach and Nualart (1986), Nair (1990), and Schoenberg (1999). In each case, the method used was to focus on one dimension of the point process, and rescale each point along that dimension according to the conditional intensity.

Here, we explore the application of Meyer's theorem to temporal marked point processes, where the mark space may be quite general, but the mark distributions have densities with respect to some reference measure. We show that, in this situation, a horizontal rescaling analogous to that of Meyer (1971) results in a compound Poisson process with a ground intensity of one and a stationary mark distribution equal to the reference measure. We follow the informal treatment that appeared recently in Daley and Vere-Jones (2003, Chapter 7). Alternative formulations and extensions to more general state spaces are discussed briefly in the final section of the paper.

## 2 Preliminaries.

Given a measurable space  $(\mathcal{S}, \mathcal{B})$ , a *point process*  $N$  is a non-negative, integer-valued random measure defined on  $\mathcal{B}$ ; that is to say, a measurable mapping from a probability space  $(\Omega, \mathcal{E}, \mathcal{P})$  into the space  $\mathcal{N}_{\mathcal{S}}$  of integer-valued,  $\sigma$ -finite measures on  $(\mathcal{S}, \mathcal{B})$ . For any  $B \in \mathcal{B}$ ,  $N(B)$  represents the number of points in  $B$ . Following Daley and Vere-Jones (1988), we shall

assume that  $\mathcal{S}$  is a complete separable metric space, with  $\mathcal{B}$  the  $\sigma$ -algebra of Borel sets in  $\mathcal{S}$ . While in general  $N(B)$  is permitted to take the value  $+\infty$ , it is required to be a.s. finite for compact  $B$ .

For *marked point processes*, the space  $\mathcal{S}$  is a product space  $\mathcal{G} \otimes \mathcal{M}$ , where  $\mathcal{G}$  is called the *ground space* and  $\mathcal{M}$  the *mark space*. For instance, if  $\mathcal{S} = \mathcal{G} \otimes \mathcal{M}$  where  $\mathcal{G}$  is the real line or a portion thereof, we call the process  $N$  a *temporal marked point process* (hereafter abbreviated *t.m.p.p.*); if  $\mathcal{S} = \mathbf{R}^2 \otimes \mathcal{M}$ , then  $N$  is a *planar marked point process*, etc.

The essential distinction we shall make between a marked point process on  $\mathcal{G} \otimes \mathcal{M}$  and a general point process on  $\mathcal{G} \otimes \mathcal{M}$  is that for every compact set  $G \in \mathcal{B}_{\mathcal{G}}$ , the Borel sets in  $\mathcal{G}$ ,

$$N(G \otimes \mathcal{M}) < \infty \quad (\text{a.s.})$$

In the case of a t.m.p.p., for example, this implies that the process can be represented as a countable set of pairs  $(t_i, m_i)$  with the property that, almost surely, only a finite number of such pairs have time coordinates  $t_i$  falling within any bounded time interval  $(a, b]$ . Thus the time points by themselves constitute a well-defined point process, the *ground process*, on  $\mathbf{R}$ .

In what follows we focus on t.m.p.p.s where  $\mathcal{G} = [t_0, \infty]$  consists of all times not before some initial origin time,  $t_0$ . We suppose throughout that  $\mathcal{M}$  is a complete separable metric space, equipped with its own Borel  $\sigma$ -algebra  $\mathcal{B}_{\mathcal{M}}$  and reference measure  $\pi$ , which to facilitate the statement of results in Section 3 is taken to be a fixed probability measure.

A point process is called *simple* if, with probability 1, all its points are distinct. In the case of a t.m.p.p.  $N$ , simplicity means all the pairs  $(t_i, m_i)$  are distinct, while simplicity of the ground process requires that no two points of  $N$  occur simultaneously.

Recall that an  $F$ -predictable process  $Y(t, \omega)$  is a process on  $\mathbf{R} \times \Omega$  which is measurable

with respect to the predictable  $\sigma$ -algebra  $\Psi^{\mathcal{F}}$  generated by events of the form  $E \times (s, t]$ , with  $E \in \mathcal{F}_s$ . We extend this definition to the marked case by defining a *marked-predictable process* as a process  $Y(t, m, \omega)$  which is measurable with respect to the product  $\sigma$ -algebra  $\sigma(\Psi^{\mathcal{F}} \times \mathcal{B}_{\mathcal{M}})$  on  $\mathbf{R} \times \mathcal{M} \times \Omega$ . Equivalently, the marked-predictable processes are generated by cross-products of sets generated by left-continuous temporal stochastic processes with the Borel subsets of  $\mathcal{B}_{\mathcal{M}}$ .

Now suppose  $N$  is an ordinary (unmarked) temporal point process and  $\mathcal{F}$  is a filtration on  $\mathcal{S} \subseteq \mathcal{R}$  to which  $N$  is adapted. If  $N$  has first moment measure  $\mu$  with  $\mu(B) = E[N(B)] < \infty$  for bounded Borel sets  $B \in \mathcal{B}_{\mathcal{G}}$ , the  $\mathcal{F}$ -*compensator*  $A$  of  $N$  is defined as the unique  $\mathcal{F}$ -predictable process such that  $N - A$  is an  $\mathcal{F}$ -martingale. This concept can be extended to a t.m.p.p.  $N$  by defining  $A$  to be the collection, over every bounded  $M \in \mathcal{B}_{\mathcal{M}}$ , of compensators  $A(t, M)$  of the simple point processes  $N(t, M) \equiv N([t_0, t] \times M)$  (Jacod, 1975). In this case the defining equation of the compensator takes the form, for each  $M \in \mathcal{B}_{\mathcal{M}}$ ,

$$E[N(b, M) - N(a, M) | \mathcal{F}_a] = E[A(b, M) - A(a, M) | \mathcal{F}_a], \quad (1)$$

with  $A(t, M) \equiv A(t, M, \omega)$  predictable. For fixed  $M$ ,  $A(t, M)$  is a.s. monotonic increasing in  $t$ , and for fixed  $t$ ,  $A(t, M)$  defines a.s. a measure on  $(\mathcal{M}, \mathcal{B}_{\mathcal{M}})$ .

We say a t.m.p.p.  $N$  is *non-terminating* with respect to the reference measure  $\pi$  if with probability one,  $\lim A(t, M) = \infty$  as  $t \rightarrow \infty$  for all  $M \in \mathcal{B}_{\mathcal{M}}$  with  $\pi(M) > 0$ .

Let  $\ell$  denote Lebesgue measure on  $\mathcal{G}$ . If, for each  $m \in \mathcal{M}$  there exists an integrable, non-negative,  $\mathcal{F}$ -adapted process  $\lambda(t, m) \equiv \lambda(t, m, \omega)$  such that, with probability 1, for all  $t \in \mathbf{R}$  and  $M \in \mathcal{B}_{\mathcal{M}}$ ,

$$\int_M \int_{t_0}^t \lambda(s, m) \ell(ds) \pi(dm) = A(t, M), \quad (2)$$

then  $\lambda$  is called an  $\mathcal{F}$ -conditional intensity of  $N$ .

Conditions for the existence of a conditional intensity can be phrased in terms of the measure  $C$  induced by the point process on the product space  $\mathbf{R} \times \mathcal{M} \times \Omega$ , through the equations,

$$C((a, b] \times B \times G) = E\{I_G(\omega)N[(a, b] \times B]\} \quad (3)$$

where  $G \in \mathcal{E}$ ,  $B \in \mathcal{B}_{\mathcal{M}}$ ,  $t_0 \leq a < b$ . This measure represents a marked version of Campbell measure (see e.g. Section 13.2 of Daley and Vere-Jones, 1988).

The existence of a conditional intensity is equivalent to requiring this measure to be absolutely continuous with respect to the product measure  $\ell \times \pi \times \mathcal{P}$ . When this condition is satisfied, the Radon-Nikodym derivative  $\lambda(t, m, \omega)$  can be identified with a version of the  $\mathcal{F}$ -conditional intensity for the m.p.p.  $N$ .

If we are concerned with predictability, we restrict the event  $G$  in (??) to  $\mathcal{F}_a$ , and require the marked Campbell measure to be absolutely continuous with respect to the product measure  $\ell \times \pi \times \mathcal{P}$  on the  $\sigma$ -algebra  $\Psi^{\mathcal{F}} \times \mathcal{B}_{\mathcal{M}}$ . The density then has a version which is measurable with respect to this  $\sigma$ -algebra; that is to say, when the absolute continuity condition is satisfied on  $\Psi^{\mathcal{F}} \times \mathcal{B}_{\mathcal{M}}$ , we may suppose the conditional intensity to be itself a marked-predictable function. Even if we start with a version of the conditional intensity that is not predictable, it induces a measure on  $\Psi^{\mathcal{F}} \times \mathcal{B}_{\mathcal{M}}$  which is absolutely continuous with respect to the product measure, and so gives rise to a predictable version of the conditional intensity. Thus, there is no loss of generality in assuming that the conditional intensity that we are dealing with is predictable.

It is well known that if an  $\mathcal{F}$ -conditional intensity  $\lambda$  exists, a version of it obeys, almost

surely and for almost all fixed  $t \in \mathcal{G}$  and  $M \in \mathcal{B}_{\mathcal{M}}$ ,

$$\int_M \lambda(t, m) \pi(dm) = \lim_{\Delta t \downarrow 0} E [N(t + \Delta, M) - N(t, M)] / \Delta t, \quad (4)$$

where we have again dropped the explicit dependence on  $\omega$  for brevity (see e.g. Papangelou, 1972; chapter 2.4 of Karr, 1986). Provided  $\lambda$  exists, one may also define the compensator-type quantity

$$a(t, m) = \int_{t_0}^t \lambda(s, m) \ell(ds). \quad (5)$$

We have then, for  $M \in \mathcal{B}_{\mathcal{M}}$ ,

$$A(t, M) = \int_M a(t, m) \pi(dm). \quad (6)$$

If, for  $\pi$ -almost all  $m$ ,  $a(t, m) \rightarrow \infty$ , we shall say that  $N$  is  $\pi$ -essentially non-terminating; clearly this implies that  $N$  is non-terminating as defined below (??).

Frequently in applications one takes for  $\mathcal{F}$  the filtration  $\mathcal{H}$  generated by the history of the process  $N$  from a given origin time  $t_0$ , usually set at 0. We shall call this the *internal history* of  $N$ .

A fundamental result for this set-up is that the conditional intensity uniquely determines the finite-dimensional distributions of any simple point process (e.g. Chapter 13 of Daley and Vere-Jones, 1988). If, in particular,  $N$  is a simple, ordinary (unmarked) point process, and the  $\mathcal{H}$ -conditional intensity  $\lambda$  of  $N$  is deterministic - that is, a fixed, non-random function, which does not depend on the history - then  $N$  is a *Poisson process*. The name stems from the fact that for such a process, the number of points occurring in the time interval  $(a, b)$  is a Poisson random variable whose mean is the integral

$$E \{N[(a, b)]\} = \int_a^b \lambda(t) \ell(dt) \quad (7)$$

(see e.g. Chapter 2 of Daley and Vere-Jones, 1988). The assumption that the conditional intensity and the compensator are deterministic means that the past of the process has no influence on its present or future, and so the process has the lack of memory property characteristic of a Poisson process.

For marked point processes, if the ground process of  $N$  has conditional intensity  $\lambda_g(t)$ , then  $\lambda_g$  is called the *ground intensity* of  $N$ . When  $\lambda(t, m)$  exists, a version of the ground intensity is given by

$$\lambda_g(t) = \int_{\mathcal{M}} \lambda(t, m) \pi(dm). \quad (8)$$

Note that by definition the ground intensity obeys

$$\int_{t_0}^t \lambda_g(s) d\ell(s) = A(t, \mathcal{M}). \quad (9)$$

In this case, if  $\lambda(t, m)$  is deterministic, then  $N$  is a *compound Poisson process*, with the property that for every time interval  $(a, b]$  and mark set  $M \in \mathcal{B}_{\mathcal{M}}$ , the number of events in  $(a, b) \times M$  is Poisson with mean

$$E \{N[(a, b) \times M]\} = \int_a^b \int_M \lambda(t, m) \ell(dt) \pi(dm). \quad (10)$$

Note that we do not require here that the mark be integer-valued. If, for example,  $\lambda(t, m) = \lambda f(m)$ , where  $\lambda$  is a constant and  $\int_{\mathcal{M}} f(m) \pi(dm) = 1$ , then  $N$  is called a *stationary compound Poisson process*. The ground process is then a stationary Poisson process with rate  $\lambda$  and the marks are i.i.d. with density  $f$  with respect to the reference measure  $\pi$ .

More generally, we say  $N$  has *separable* conditional intensity  $\lambda$  if both the conditional intensity and ground intensity of  $N$  exist, and obey the relation

$$\lambda(t, m, \omega) = \lambda_g(t, \omega) f(m), \quad (11)$$

where  $f(m)$  is a fixed probability density with respect to  $\pi$ . This corresponds to the situation where the successive marks are chosen independently of the evolution of the process in time, although the size of past marks may influence the future ground intensity of the process. In the terminology of Daley and Vere-Jones (2003), such a process has ‘*unpredictable marks*’.

In general, the ratio  $\lambda(t, m)/\lambda_g(t)$  might depend on  $t$ , and on the history of the process  $N$  up to time  $t$ , as well as potentially other information contained in  $\mathcal{F}_t$ . For a general (possibly non-separable) t.m.p.p.  $N$ , we say  $N$  has *conditional mark density*  $f \equiv f(t, m, \omega)$  if  $N$  has a conditional intensity  $\lambda$  and ground intensity  $\lambda_g$  obeying  $f(t, m, \omega) = \lambda(t, m, \omega)/\lambda_g(t, \omega)$  for  $\lambda_g(t, \omega) > 0$ ; when  $\lambda_g(t, \omega) = 0$  we set  $f(t, m, \omega) = 1$ .

We note that the definitions for t.m.p.p.s extend to the case where  $G = \mathbf{R}^d$  instead of  $\mathbf{R}$ , just as the definitions of unmarked point processes extend from  $\mathbf{R}$  to  $\mathbf{R}^d$ ; see e.g. Nair (1990) or Schoenberg (1999). In such extensions some care needs to be taken in selecting the family of  $\sigma$ -algebras to mimic the role of the history when  $G$  represents time.

### 3 The Rescaling Theorem

In this section we investigate ways to rescale a t.m.p.p.  $N$  with conditional mark density  $f$ , so that the rescaled version of  $N$  is a compound Poisson process. We restrict ourselves to the case where rescaling is achieved by moving each point of  $N$  parallel to the time axis. That is, we refer to transformations where each point  $\{t, m\}$  of  $N$  gets shifted to a new time, but its mark remains  $m$ . We refer to this as *horizontal rescaling* to distinguish it from the case where the marks are rescaled but the time points are left invariant (*vertical rescaling*).

We recall that the reference measure  $\pi$  on  $(\mathcal{M}, \mathcal{B}_{\mathcal{M}})$  is taken to be a probability mea-

sure. When the process is separable, rescaling is a straightforward corollary of the rescaling theorem for the unmarked case.

**Theorem 3.1.** Suppose  $N$  is a non-terminating t.m.p.p. with simple ground process and a separable conditional intensity, as in (??). Rescale each point of  $N$  via  $\{t, m\} \longrightarrow \{A(t, \mathcal{M}), m\}$ . Then the rescaled point process  $\tilde{N}$  is a stationary compound Poisson process with unit ground rate and stationary mark density  $f(m)$  relative to  $\pi$ . Equivalently, if we transform each point via  $\{t, m\} \longrightarrow \{A(t, \mathcal{M})f(m), m\}$ , the rescaled process is still a stationary compound Poisson process with unit ground rate, but the stationary mark distribution is now given by  $\pi$ .

**Proof.** From the 1-d random time change theorem (Meyer, 1971) it follows that the marginal point process  $N(t)$  gets transformed into a Poisson process with ground rate  $\tilde{\lambda}_g(\tau) = \int_{\mathcal{M}} f(m)\pi(dm) = 1$ , and as the marks  $m$  do not get transformed and are independent of  $N$ , the first result follows immediately. For the second result, note that the transformed ground process still has conditional intensity  $\tilde{\lambda}_g(\tau) = \int_{\mathcal{M}} \pi(dm) = 1$  and therefore still corresponds to a unit rate Poisson process. In the second case, points with different marks get rescaled differently, whereas in the first case they all get rescaled in the same way.  $\square$

Now we suppose that  $N$  has simple ground process and a conditional intensity  $\lambda$ , but we do not assume that  $\lambda$  is separable. Throughout the discussion  $a(t, m)$  refers to the compensator-type function (??).

For a general t.m.p.p., assuming the conditional intensity exists, we define a family of time transformations, one for each mark value, through the equations

$$(t, m) \rightarrow (a(t, m), m). \tag{12}$$

Under these transformations, the original point process  $N$ , consisting of the family of points  $\{t_i, m_i\}$ , is transformed into the point process  $\tilde{N}$ , consisting of the family  $\{\tau_i, m_i\} \equiv \{a(t_i, m_i), m_i\}$ . Consistently with this definition, we define the *rescaled*  $\sigma$ -field  $\tilde{\mathcal{F}}_\tau$  to be the  $\sigma$ -field generated by events of the form

$$\{\tilde{N}[(\alpha, \beta) \times M] = k\}, \quad (0 \leq \alpha < \beta, M \in \mathcal{M}, k \in \{0, 1, 2, \dots\}). \quad (13)$$

Observe that, by construction,  $\tilde{N}$  is  $\tilde{\mathcal{F}}$ -adapted.

The following is the main result of this section. To our embarrassment, it came to our attention just before the paper was due to be printed that our original proof rested on a false lemma. In fact, the truth of the Theorem in the case of discrete marks (multivariate point process) follows from well-known results, which can then be used to approximate the continuous case. In the amended discussion below we outline our original argument, point out the error, and sketch a possible resolution, perforce leaving a more adequate discussion to Schoenberg and Vere-Jones (2003).

**Theorem 3.2.** Suppose that the marked point process  $N$  is  $\mathcal{F}$ -adapted and  $\pi$ -essentially non-terminating, with simple ground process and  $\mathcal{F}$ -conditional intensity  $\lambda(t, m)$ . Let  $\tilde{N}$  denote the rescaled point process defined by the mapping  $(t, m) \longrightarrow (a(t, m), m)$ , so that  $\tilde{N}$  is  $\{t, m\}$ , and  $\tilde{N}$  is  $\tilde{\mathcal{F}}$ -adapted, where  $\tilde{\mathcal{F}}$  is defined as in (??). Then  $(\tilde{N}, \tilde{\mathcal{F}})$  is a stationary compound Poisson process with unit ground intensity and stationary mark density  $\pi$ .

**Proof.** As in earlier treatments (Meyer, 1971; Papangelou, 1972; Brémaud, 1972; Aalen and Hoem, 1978; Kurtz, 1980; Merzbach and Nualart, 1986; Brown and Nair, 1988; Nair, 1990; Schoenberg, 1999), our proof hinged on the fact that the compensator determines the process uniquely. Since this in turn requires the ground process to be simple, we first observe

that this property holds for the transformed time process  $\tilde{N}$ .

Simplicity of the ground process for  $\tilde{N}$  means that the probability is zero that  $\tilde{N}$  has any two points  $(\tau, m_1)$  and  $(\tau, m_2)$  with the same first coordinate. Consider the event where two such points exist, i.e. where  $N$  has two distinct points  $(t_1, m_1)$  and  $(t_2, m_2)$  such that  $a(t_1, m_1) = a(t_2, m_2)$ . Assume without loss of generality that  $t_1 < t_2$ . (The case  $t_1 = t_2$  can be excluded by our assumption that  $N$ 's ground process is simple.) Let  $z = (t_1, m_1)$ , and let  $D_z = \{(t, m) : t > t_1; a(t, m) = a(t_1, m_1)\}$ . Since  $a(t, m)$  is  $\mathcal{F}$ -predictable and  $\Psi_{t_1}^{\mathcal{F}} \subseteq \Psi_t^{\mathcal{F}}$  for  $t \geq t_1$ ,  $D_z$  is an  $\mathcal{F}$ -predictable set, i.e. the indicator of this set  $\mathbf{1}_{D_z}$  is an  $\mathcal{F}$ -predictable process. Hence the expected number of points in  $D_z$  is given by  $E \int \mathbf{1}_{D_z} dN = E \int \mathbf{1}_{D_z} dA = 0$ , and since  $\mathbf{1}_{D_z}$  is non-negative this implies that  $N(D_z) = \int \mathbf{1}_{D_z} dN = 0$  a.s. Since  $N$  contains only countably many points, so does  $\tilde{N}$ . As  $D_z$  is a null set (i.e.  $P\{N(D_z) > 0\} = 0$ ), so is the countable union  $\cup_z D_z$ , i.e.  $P\{N(\cup_z D_z) > 0\} = 0$ . Thus  $\tilde{N}$  is simple.

Consider now an unmarked, simple point process  $N$ , with transformed process  $\tilde{N}(\tau) = N(A^{-1}(\tau))$ . A proof of the main theorem in this case can be based on the following observation. Let  $Y(\tau, \omega)$  be bounded and  $\tilde{\mathcal{F}}$ -predictable. Then  $Y(A(t), \omega)$  is  $\mathcal{F}$ -predictable. Indeed, if  $Y$  has the form of the indicator function of a basic predictable set  $I_E(\omega) \times I_{(\alpha, \beta]}(\tau)$ , for  $E \in \mathcal{F}_\alpha$ , then  $S := A^{-1}(\alpha) = \inf\{t : A(t) = \alpha\}$  and  $T := A^{-1}(\beta)$  are  $\mathcal{F}$ -stopping times, since under our assumptions,  $A(t)$  is even a continuous function. For this special case, the predictability of  $Y$  follows from the result (presumably well-known) that if  $S, T$  are two  $\mathcal{F}$ -stopping times with  $S < T$  (a.s.), and if  $E \in \mathcal{F}_S$ , then

$$Y(t, \omega) = I_E(\omega) I_{(S, T]}(t) \tag{14}$$

is  $\mathcal{F}$ -predictable. The case of general predictable  $Y$  follows from straightforward extension

arguments.

The theorem itself, still for the unmarked case, now follows from the following equalities

$$E \left[ \int Y(\tau) \tilde{N}(d\tau) \right] = E \left[ \int Y(A(t)) N(dt) \right] \quad (15)$$

$$= E \left[ \int Y(A(t)) dA(t) \right] \quad (16)$$

$$= E \left[ \int Y(s) ds \right] \quad (17)$$

$$= E \left[ \int Y(\tau) d\tau \right]. \quad (18)$$

Step (??) follows from the definition of the transformed process. Step (??) follows from the martingale property of the original process as in (??), and the observation above that  $Y(A(t))$  is  $\mathcal{F}$ -mark-predictable. Step (??) is the substitution  $s = A(t)$ . It requires the assumption of the Theorem that  $A(t)$  increases without bound for each  $m$  in the support of  $\pi$ . Step (??) is just the replacement of  $s$  by  $\tau$ . Taken together, they imply that the  $\tilde{\mathcal{F}}$ -compensator for  $\tilde{N}$  is the deterministic process

$$\tilde{A}(\tau) = \tau$$

and hence that  $\tilde{\lambda}(\tau, m) \equiv 1$  is an  $\tilde{\mathcal{F}}$ -conditional intensity for  $\tilde{N}$ . Thus  $\tilde{N}$  is a stationary Poisson process.

The corresponding equalities for a marked point process take the form

$$E \left[ \int Y(\tau, m) \tilde{N}(d\tau \times dm) \right] = E \left[ \int Y(a(t, m), m) N(dt \times dm) \right] \quad (19)$$

$$= E \left[ \int Y(a(t, m), m) a(dt, m) \pi(dm) \right] \quad (20)$$

$$= E \left[ \int Y(s, m) ds \pi(dm) \right] \quad (21)$$

$$= E \left[ \int Y(\tau, m) d\tau \pi(dm) \right]. \quad (22)$$

for  $Y$  any marked  $\tilde{F}$ -predictable process. These are the equations which formed the basis of our original proof, and the resulting equation (??) appears to be still true, certainly for the discrete marked case, as the assertion follows from the known results for multivariate point processes. However, it is certainly *not* still true that  $Y(a(t, m), m)$  is  $\mathcal{F}$ -predictable when  $Y$  is marked  $\tilde{F}$ -predictable. This is readily seen even for a bivariate process, by taking the two components to increase at different rates, and choosing e.g.  $Y(t, m) = \tilde{N}_1(0, \tau) + \tilde{N}_2(0, \tau)$ . The resulting process  $Y(a(t, m), m)$  is not even  $\mathcal{F}$ -adapted. Thus the proof cannot be based, as we originally supposed, on a simple extension of the univariate argument, but requires a more careful analysis of the interweaving of the  $a(t, m)$  for different  $m$ .

\*\*

A heuristic argument for the marked case may be outlined as follows.

Fix disjoint Borel mark sets  $M_1$  and  $M_2$  and transformed time intervals  $B_1$  and  $B_2$ .

Suppose  $F$  is a set in the  $\sigma$ -field generated by sets of the form  $\tilde{N}_{M_1}(B)$ , where  $B$  is in  $B_1$ .

Let  $Q_{12}$  denote the event that  $\sup_{m \in M_1, \tau \in B_1} a^{-1}(\tau, m) < \inf_{m \in M_2, \tau \in B_2} a^{-1}(\tau, m)$ .

Then for  $(\tau_1, \tau_2) \in B_2$ ,

$$E \left[ \int_{M_2} \int_{\tau_1}^{\tau_2} d\tilde{N}(\tau, m) | F, Q_{12} \right] \quad (23)$$

$$= E \left[ \int_{M_2} \int_{a^{-1}(\tau_1, m)}^{a^{-1}(\tau_2, m)} dN(t, m) | F, Q_{12} \right] \quad (24)$$

$$= E \left[ \int_{M_2} \int_{a^{-1}(\tau_1, m)}^{a^{-1}(\tau_2, m)} dA(t, m) | F, Q_{12} \right] \quad (25)$$

$$= E \left[ \int_{M_2} \int_{\tau_1}^{\tau_2} d\ell(t) d\pi(m) | F, Q_{12} \right]. \quad (26)$$

Note that the above is true for all  $\tau_1$  and  $\tau_2$  in  $B_2$ . It follows that conditioned on  $F$  and

$Q_{12}$ , the compensator  $\tilde{A}$  of  $\tilde{N}_{M_2}$  is equivalent on  $B_2$  to  $\ell \times \pi$ , regardless of  $F$ . Thus given  $Q_{12}$ ,  $\tilde{A}$  is independent of  $F$ ; i.e. conditional on  $Q_{12}$ ,  $\tilde{N}_{M_2}(B'_2)$  is independent of  $\tilde{N}_{M_1}(B'_1)$ , for all  $B'_1 \in B_1$  and  $B'_2 \in B_2$ .

Similarly, conditional on  $Q_{21}$ ,  $\tilde{N}_{M_2}(B'_2)$  is independent of  $\tilde{N}_{M_1}(B'_1)$ , for all  $B'_1 \in B_1$  and  $B'_2 \in B_2$ .

To complete the proof, two further steps are needed. First, by varying  $M_1$ ,  $M_2$ ,  $B_1$  and  $B_2$ , and using the simplicity of  $\tilde{N}$ , we show that whenever  $M_1$ ,  $M_2$  are disjoint,  $\tilde{N}_{M_2}(B_2)$  is in fact independent of  $\tilde{N}_{M_1}(B_1)$  for all Borel  $B_1$  and  $B_2$ .

Second, for any Borel  $M \in \mathcal{M}$ , any transformed time interval  $(\tau_1, \tau_2)$  and any  $F \in \tilde{\mathcal{F}}_{\tau_1}$ , we decompose  $F$  into parts  $F_m$  in  $\sigma(\{\tilde{N}[(\alpha, \beta) \times \{m\}] = k; 0 \leq \alpha \leq \beta \leq \tau_1; k = 1, 2, \dots\})$ , and a remainder, so that

$$E \left[ \int_M \int_{\tau_1}^{\tau_2} d\tilde{N}(\tau, m) | F \right] = \int_M E \left[ \int_{\tau_1}^{\tau_2} d\tilde{N}(\tau, m) | F_m \right] d\pi(m) \quad (27)$$

$$= \int_M E \left[ \int_{a^{-1}(\tau_1, m)}^{a^{-1}(\tau_2, m)} dN(t, m) | F_m \right] d\pi(m) \quad (28)$$

$$= \int_M E \left[ \int_{a^{-1}(\tau_1, m)}^{a^{-1}(\tau_2, m)} dA(t, m) | F_m \right] d\pi(m) \quad (29)$$

$$= \int_M E \left[ \int_{\tau_1}^{\tau_2} d\ell(t) \right] d\pi(m) \quad (30)$$

$$= \pi(M) \times \ell(\tau_1, \tau_2), \quad (31)$$

from which the desired result follows. The key to the proof is that since  $\tilde{N}(\tau, m)$  is independent of all other marks, conditioning on  $F$  amounts to conditioning on  $F_m$ , which is a predictable set just as in the unmarked case. \*\*

We hope to develop this argument more rigorously in Schoenberg and Vere-Jones (2003).

Theorem 3.2 subsumes many well-known results. In the case of discrete marks (a *multivariate* point process), for example, we have (cf. Meyer, 1971; Brown and Nair, 1988):

**Corollary 3.3.** Let  $N$  denote a non-terminating multivariate point process with components  $N_k$ ,  $k = 1, 2, \dots, K$  with reference measure  $\pi = \{\pi_1, \pi_2, \dots, \pi_K\}$  satisfying  $\pi_k > 0$ ,  $k = 1, 2, \dots, K$ ,  $\sum_1^K \pi_k = 1$ . Suppose  $N$  has a simple ground process and an  $\mathcal{F}$ -conditional intensity  $\lambda$ . Let  $\tilde{N}$  denote the rescaled point process defined to have a point at  $\{a(t, k), k\}$  iff. the  $k$ -th component of  $N$  contains a point at  $t$ . Then  $\tilde{N}$  is a stationary compound Poisson process with unit intensity and stationary mark distribution  $\pi$ . Alternatively, if the rescaling is performed so that  $\tilde{N}$  has a point at  $(a(t, k)\pi_k, k)$  whenever the original process has a point at  $(t, k)$  (or equivalently, if the reference measure  $\pi$  assigns unit mass to each component), then the resultant process consists of  $K$  independent, unit rate Poisson processes, one for each mark.

Suppose next that the mark space is the real line. Let  $F_\pi \equiv \int_{m' < m} \pi(dm')$  be the cumulative measure corresponding to the probability measure  $\pi$ , and suppose that  $F_\pi$  is continuous as a function of  $m$ . Then a simple rescaling of the mark space, taking  $m^* = F_\pi(m)$ , converts the stationary mark distribution for the transformed process into the uniform distribution on  $[0, 1]$ . Since a compound Poisson process with constant rate and uniform mark distribution can equally be interpreted as a 2-dimensional Poisson process on a strip, we obtain the following result.

**Corollary 3.4.** Suppose that the t.m.p.p.  $N$  has real marks, that the conditions of Theorem 3.2 hold, and that the reference probability measure  $\pi$  admits a continuous cumulative version  $F_\pi$ . Then the doubly transformed process  $\tilde{N}^*$ , defined to have a point at

$(a(t_i, m), F_\pi(m_i))$  when the original process has a point at  $(t_i, m_i)$ , is a 2-dimensional Poisson process with unit rate over the strip  $\mathbf{R} \times [0, 1]$ .

## 4 Discussion and extensions

The results of the previous Section generalize in several directions, of which we briefly mention a few. A more general review is in preparation (Schoenberg and Vere-Jones, 2003).

First, although the results of Section 3 extend those of Brown and Nair (1988) to more general mark spaces, it is an open question whether the absolute continuity condition (existence of a conditional intensity) is essential to the formulation and proof of the main theorem. Discontinuity of the compensator at a location  $u = (t, m)$  in the time-mark domain means that there is nonzero probability of a point occurring arbitrarily near  $u$ , conditional on the prior history of the process, and this in turn generally leads to similar discontinuity in the compensator of the transformed process. However, additional randomness may be introduced to smooth out such discontinuities; for instance, for the case of a simple point process  $N$  on the line with possibly discontinuous compensator, one may obtain a rescaled Poisson process of unit rate by introducing a sequence of independent uniform random variables, independent of the filtration  $\mathcal{F}$  to which  $N$  is adapted, as proven in Theorem 5.1 of Kallenberg (1990).

Relatively little attention has been paid to the situation where the marked point process  $N$  is not simple, or where its ground process is not simple. To see that results such as Theorem 3.2 are not then directly obtainable, consider  $N$  the combination of two point processes  $N_1$  and  $N_2$ , where  $N_1 \neq N_2$ . In other words let  $N$  be a marked point process

with mark space  $\{1, 2\}$ , where  $N$  has a point at  $(t, 2)$  iff. there is a point at  $(t, 1)$ . In this case  $a(t, 1) = a(t, 2)$ , and the transformation in Theorem 3.2 does not result in a stationary compound Poisson process since the rescaled point process inherits from  $N$  the property of having a non-simple ground process.

In some situations, a non-simple marked point process  $N$  may be simplified by changing the form of the mark space. For example, if  $N$  has either one or two points at times  $t_i$ , with possibly different marks  $m_i$  and  $m_i^*$ , then one may redefine the mark space, letting  $\mathcal{M}_2 = \mathcal{M} \cup \mathcal{M}^2$ , and  $N$  will be simple when viewed as a point process on  $\mathcal{G} \otimes \mathcal{M}_2$ . If the marks are necessarily identical,  $m_i = m_i^*$ , then it may be enough to take as the modified mark space  $\mathcal{M} \otimes \{1, 2\}$ . If  $N$  may have triple-points or point of higher valency, one may similarly simplify  $N$  by changing the mark space to  $\mathcal{M}_K = \bigcup_{k=1}^K \mathcal{M}_k$ , or  $\mathcal{M}_\infty = \bigcup_{k=1}^{\infty} \mathcal{M}_k$ .

In the previous Section, the independence and i.i.d. properties of a compound Poisson process were achieved by distorting the relative time positions of points with different marks. The extent of this distortion depends on how close the original process was to being separable, for in Theorem 3.1 we saw that in that case a suitable rescaling can be achieved such that the dislocation of each point is entirely independent of its mark. An alternative possibility is that the reduction to a Poisson process can be achieved by rescaling the marks (vertical rescaling). Merzbach and Nualart (1986) analyzed the case where  $\mathcal{M}$  and  $\mathcal{G}$  are the real half-line  $[0, \infty]$ , showing that the rescaling

$$(t, m) \longrightarrow (t, \int_0^m \lambda(t, m') dm') \quad (32)$$

transforms  $N$  into a stationary Poisson process with rate 1 on the planar quadrant  $\mathcal{S}$ , under certain conditions which were later relaxed by Nair (1990) and Schoenberg (1999). This

result is summarized in the following Theorem.

**Theorem 4.1.** Suppose  $N$  is a t.m.p.p. with  $\mathcal{M} = [0, \infty]$ . Suppose  $N$  has simple ground process and  $\mathcal{F}$ -conditional intensity  $\lambda$ , and that the ground intensity  $\int_0^\infty \lambda(t, m) dm = \infty$  for all  $t \in \mathcal{G}$ . Then the rescaling in (??) results in a stationary Poisson process with unit rate on  $\mathcal{S} = \mathcal{G} \times \mathcal{M}$ .

If the mark space is a subset  $M = [m_0, m_1]$  of  $\mathbf{R}_+$  rather than the entire real half-line, and the ground intensity  $\int_0^\infty \lambda(t, m) dm < \infty$ , the result in Theorem 4.1 applies nevertheless, as described in Schoenberg (1999). The resulting rescaled process  $\tilde{N}$  is a stationary Poisson process with rate 1 with respect to Lebesgue measure on the randomly-shaped region  $T = \{(t, m); t \in \mathcal{G}; 0 \leq m \leq \int_{m_0}^{m_1} \lambda(t, m') dm'\}$ .

Finally, let us refer to a few questions where further research seems needed.

(i) As we have indicated above, rescaling a marked point process can be achieved in several different ways, e.g. horizontally, vertically, or possibly by some combination of both. If the aim is to detect some departure from a standard model, which methods work best when?

(ii) In Theorem 4.1, a special role is played by the form of the mark-space; how should the results of that theorem be generalized to more general mark spaces?

(iii) Throughout the paper, it has been assumed that the model is fully known; in practice, however, it will commonly depend on one or more estimated parameters. How can the results of this paper be accommodated to allow for the additional uncertainty introduced by the estimation? Schoenberg (2002) exhibits a simple point process example where, if the fitted model is used in place of the true model, the conditional intensity of the transformed time

process differs significantly from unity. Hjort (1990), Andersen et al (1993, Section VI.3.3) and Maglaperidze et al (1998) outline some alternative approaches to goodness-of-fit tests, based on ideas of Khmaladze(1981, 1988,1993), which can be used in such situations.

## Acknowledgements

We would like to express our appreciation to Jean Jacod for detecting the error in our original version, and furnishing a convincing counterexample. Painful though this experience has been, it has spared us the worse embarrassment of presenting Daryl with a wrong paper. The authors' collaboration on this topic was fueled by Workshop 11, " Point Process Modeling and Seismological Applications of Statistics," at the University of Minnesota, June 10-14, 2002, sponsored by the Institute for Mathematics and its Applications as part of its Thematic Year on Mathematics in the Geosciences. This material is based upon work supported by the National Science Foundation under Grant No. 9978318. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## References

- Aalen, O. and Hoem, J. (1978). Random time changes for multivariate counting processes. *Scand. Act. J.*, 81–101.
- Andersen, P.K., Borgan, O., Gill, R.D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York.
- Brémaud, P.M. (1972). A martingale approach to point processes. Memorandum ERL-M345, Electronics Research Laboratory, University of California, Berkeley, Berkeley, CA.
- Brown, T. and Nair, M. (1988). A simple proof of the multivariate random time change theorem for point processes. *J. Appl. Probab.* **25**, 210–214.
- Daley, D. and Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes*. Springer, NY.
- Daley, D. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes, 2nd ed., Vol. 1* Springer, NY and Berlin.
- Hjort, N.L. (1990). Goodness of fit tests in models for life history data based on cumulative hazard rates. *Ann. Statist.* **18**, 1221–1258.
- Jacod, J. (1975). Multivariate point processes: predictable projections, Radon-Nikodym derivatives, representation of martingales. *Z. Wahrsch. verw. Geb.* **31**, 235–253.
- Kallenberg, O. (1990). Random time change and an integral representation for marked stopping times. *Prob. Theory* **86**, 167–202.

- Karr, A.F. (1986). *Point Processes and their Statistical Inference*. Dekker, NY.
- Khamaladze, E.V. (1981). Martingale approach to the goodness of fit tests. *Theory Probab. Appl.* **26**, 246–265.
- Khamaladze, E.V. (1988). An innovation approach to goodness-of-fit tests in  $R^m$ . *Ann. Statist.* **16**, 1503–1516.
- Khamaladze, E.V. (1993). Goodness of fit problem and scanning innovation martingales. *Ann. Statist.* **21**, 798–829.
- Knight, F.B. (1970). An infinitesimal decomposition for a class of Markov processes. *Ann. Math. Statist.* **41**, 1510–1529.
- Kurtz, T.G. (1980). Representations of Markov processes as multiparameter time changes. *Ann. Probab.* **8**, 682–715.
- Maglaperidze, N.O., Tsigroshvili, Z.P., and van Pul, M. (1998). Goodness-of-fit tests for parametric hypothesis on the distribution of point processes. *Math. Methods of Statist.* **7**, 60–77.
- Merzbach, E. and Nualart, D. (1986). A characterization of the spatial Poisson process and changing time. *Ann. Probab.* **14**, 1380–1390.
- Meyer, P. (1971). Demonstration simplifiée d'un théorème de Knight. In *Séminaire de Probabilités V*, Univ. Strasbourg, Lecture Notes in Math. **191**, 191–195.
- Nair, M. (1990). Random space change for multiparameter point processes. *Ann. Probab.* **18**, 1222–1231.

- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *J. Amer. Statist. Assoc.*, **83**, 9-27.
- Ogata, Y. (2001). Increased probability of large earthquakes near aftershock regions with relative quiescence. *J. Geophys. Res.* **106**(B5), 8729–8744.
- Papangelou, F. (1972). Integrability of expected increments of point processes and a related random change of scale. *Trans. Amer. Math. Soc.* **165**, 483–506.
- Schoenberg, F. (1997). *Assessment of Multi-dimensional Point Processes*. Ph.D. Thesis, University of California, Berkeley.
- Schoenberg, F. (1999). Transforming spatial point processes into Poisson processes. *Stoch. Proc. Appl.* **81**, 155–164.
- Schoenberg, F.P. (2002). On rescaled Poisson processes and the Brownian bridge. *Ann. Inst. Statist. Math.* **54**(2), 445–457.
- Schoenberg, F.P. and Vere-Jones, D. (2003). Rescaling results for point processes. *in preparation*.