

Time-Dependent Palm Probabilities and Queueing Applications

Brian H. Fralix, Germán Riaño, Richard F. Serfozo

July 17, 2007

Abstract

This study shows that time-dependent Palm probabilities of a non-stationary process are expressible as integrals of a certain stochastic intensity process. A consequence is a characterization of a Poisson process in terms of time-dependent Palm probabilities. These two results are analogous to results of Papangelou and Mecke, respectively, for stationary point processes. Included is a new proof of Watanabe's characterization of a Poisson process. Next, using stochastic intensities of time-dependent Palm probabilities, we present conditions under which the distribution of a stochastic process (e.g., a queueing process) at a fixed time is equal to its Palm probability distribution conditioned on a jump at that time. This result is a time-dependent analogue of an ASTA property (arrivals see time averages) for stationary processes. Another result is that, for an asymptotically stationary process, a Palm probability P_t conditioned on a point at a time t converges weakly to a Palm probability for a stationary process as $t \rightarrow \infty$. We also present formulas for time-dependent Palm probabilities of Markov processes, and Little laws for queueing systems that relate queue-length processes to time-dependent Palm probabilities of sojourn times of the items in the system.

1 Introduction

Consider a point process N on \mathfrak{R} that is stationary ($N(B)$ denotes the number of points in a set B , and the joint distribution of its increments is invariant under shifts in time). Let P^0 denote the Palm probability of N conditioned that N has a point at 0. Papangelou [19] proved that the absolute continuity $P^0 \ll P$ on \mathcal{F}_{0-} is equivalent to the existence of a stochastic intensity of N with respect to its σ -field history \mathcal{F}_t . He showed how the intensity is related to the Radon-Nikodym derivative dP^0/dP . This result can be used to prove Mecke's theorem [14] that the stationary process N is a Poisson process if and only if $P^0 = P$ on \mathcal{F}_{0-} .

This study presents analogous theorems and queueing applications for processes that need not be stationary. To describe the results, suppose that N is a point process on \mathfrak{R} that need not be stationary, and assume the mean measure

$\mu(B) = E[N(B)]$ is locally finite. Let P_t denote the Palm probability of N conditioned that it has a point at t . Such time-dependent Palm probabilities were introduced by Ryll-Nardzewski [22] and developed further by Kallenberg [12]. Also, see Nieuwenhuis [18] and the queueing studies of Rolski [21] and Riaño [20].

Our study sheds more light on the similarity of the structural relationship between P^0 and P for a stationary system and the relationship between P_t and P for non-stationary systems. The first result is a Papangelou-type theorem that says $P_t \ll P$ on \mathcal{F}_{t-} for μ -a.e. t if and only if N has a stochastic intensity $\lambda(t)$. In this case, $dP_t/dP = \lambda(t)/E[\lambda(t)]$ μ -a.e. t . We present a related result using a novel type of stochastic intensity, motivated by ideas in [15] and [6], that is weaker than the usual stochastic intensity.

Next, we prove that N is an \mathcal{F}_t -Poisson process with deterministic rate function λ if and only if $P_t = P$ on \mathcal{F}_{t-} for μ -a.e. t , and $\mu(B) = \int_B \lambda(t)dt$, $B \in \mathcal{B}$. Each of these statements is equivalent to Watanabe's [27] condition that $N(s, t] - \int_s^t \lambda(s)ds$, for $t > s$, is an \mathcal{F}_t -martingale for any fixed s .

Many of our results are for a stochastic process $X(t)$ on a general space that is related to N . This process satisfies *transient ASTA* at time t if

$$P_t(X(t-) \in \cdot) = P(X(t-) \in \cdot).$$

Using the Papangelou-type theorem, we show that if N has a $X(t)$ -stochastic intensity $\lambda(t)$ (in our weak sense), then transient ASTA at time t is equivalent to $E[\lambda(t)|X(t-)] = E[\lambda(t)]$ (a lack of bias condition). Hence transient ASTA is satisfied for all t if N is a Poisson process with a deterministic rate function. Our time-dependent ASTA results are analogous to those for stationary queueing processes such as in Brémaud [5], and for limiting averages of non-stationary processes such as in Melamed and Whitt [15, 16].

A summary of the rest of the study is as follows. Section 6 describes conditions under which a Palm probability $P_t(A)$ is the limit of a probability conditioned that n has a point in $[t, v]$, where $v \downarrow t$. Section 7 shows that when X and N are asymptotically stationary, the Palm probability P_t converges as $t \rightarrow \infty$ to a Palm probability for a stationary system. Section 8 gives formulas for Palm probabilities for Markov and Semi-Regenerative processes; this involves generalizations of Lévy's classical formula for expectations of functions of Markov processes. Finally, Section 9 presents time-dependent Little laws for queueing systems in which the distribution of a queue-length process is related to the Palm probability distribution of the sojourn times of items in the system.

2 Preliminaries

We will use the following notation throughout the paper. Let $N = \{N(B) : B \in \mathcal{B}\}$ denote a point process on \mathfrak{R} defined on a probability space (Ω, \mathcal{F}, P) , where $N(B)$ is the number of points in B and \mathcal{B} is the family of Borel sets in

\mathfrak{R} . That is,

$$N(B) = \sum_n \mathbf{1}(T_n \in B), \quad B \in \mathcal{B},$$

where $\dots \leq T_{-2} \leq T_{-1} \leq T_0 \leq 0 < T_1 \leq T_2 \leq \dots$ are the point locations. Also, $N(s, t]$ denotes the number of points in the interval $(s, t]$. Assume the mean measure $\mu(B) = E[N(B)]$ is σ -finite.

The main focus of our study are time-dependent Palm probabilities of N ; see [22, 12]. For the following definition, assume that Ω is a complete, separable metric space, and \mathcal{F} is its associated Borel sets. Then there exists a (μ -a.e. unique) probability kernel $P_t(A)$ such that

$$E\left[N(B)\mathbf{1}(\omega \in A)\right] = \int_B P_t(A)\mu(dt), \quad A \in \mathcal{F}, B \in \mathcal{B}. \quad (1)$$

This is proved in [9, 12]. Here $\mathbf{1}(\cdot)$ is the indicator function that is 1 or 0 according as the statement (\cdot) is true or false.

Definition 1 The collection P_t for μ -a.e. t defined by (1) is the family of time-dependent *Palm probabilities* induced by the point process N . The expectation under P_t is denoted by E_t .

A major tool for dealing with Palm probabilities is the following Campbell-Mecke formula, which is a consequence of (1).

Theorem 2 For any measurable $f : \mathfrak{R} \times \Omega \rightarrow \mathfrak{R}_+$,

$$\int_{\Omega} \int_{\mathfrak{R}} f(t, \omega) N(dt) P(d\omega) = \int_{\mathfrak{R}} \int_{\Omega} f(t, \omega) P_t(d\omega) \mu(dt). \quad (2)$$

Motivated by the following result, $P_t(A)$ is called the probability of A , given that N has a point at time t (an event that may have probability 0).

Proposition 3 If N is a simple point process (its points are distinct a.s.), then

$$P_t(N(\{t\}) = 1) = 1, \quad \mu\text{-a.e. } t.$$

Proof This follows since by (2), for $B \in \mathcal{B}$,

$$\begin{aligned} \int_B P_t(N(\{t\}) = 1)\mu(dt) &= E\left[\int_B \mathbf{1}(N(\{t\}) = 1)N(dt)\right] \\ &= E\left[\sum_n \mathbf{1}(T_n \in B)\right] = \mu(B). \end{aligned}$$

■

Note that P_t is defined on the underlying probability space (Ω, \mathcal{F}, P) , which may be the home for random elements other than N . On this probability space, suppose that $X = \{X(t) : t \in \mathfrak{R}\}$ is a measurable stochastic process that takes values in a complete, separable metric space \mathbb{E} with paths in the set $D(\mathfrak{R})$ of

functions on \mathbb{E} that are right-continuous and have left-hand limits. The process X is a vehicle for expressing time-dependent Palm probabilities for events in \mathcal{F} that may or may not depend on N .

For now, we make no assumptions on the dependency between N and X . In some applications, X is the primary process and N represents times at which certain events of X occur such as the event that X has a jump (then N is a function of X). On the other hand, N may be the primary process and X may represent auxiliary events or functions of processes that interact with N .

The Campbell-Mecke formula for any real-valued process $Y(t)$ is

$$E\left[\int_{\mathfrak{R}} Y(t)N(dt)\right] = \int_{\mathfrak{R}} E_t[Y(t)]\mu(dt), \quad (3)$$

provided the expectations exist. For instance, $Y(t) = g(N, X, t)$ where g is a real-valued function (all functions in this study are assumed to be measurable). Then by the definition of a Radon-Nikodym derivative,

$$E_t[Y(t)] = \frac{E[Y(t)N(dt)]}{\mu(dt)}. \quad (4)$$

Next, we show that when the processes N and X are jointly stationary, the Palm probabilities P_t of events for these processes can be represented by a single Palm probability. Let S^t denote the time-shift operator such that

$$S^t N = \{N(B+t) : B \in \mathcal{B}\}, \quad S^t X = \{X(s+t) : s \in \mathfrak{R}\}.$$

The pair (N, X) is *stationary* if $(S^t N, S^t X) \stackrel{d}{=} (N, X)$, $t \in \mathfrak{R}$. In that case, the mean measure μ is a multiple of the Lebesgue measure.

For this stationary system, let P^0 denote the Palm probability conditioned that N has a point at 0; see for instance [1, 9, 12, 23]. Under this Palm probability, the distribution of (N, X) is given by

$$\int_B P^0((N, X) \in \cdot)\mu(dt) = E\left[\int_B \mathbf{1}((S^t N, S^t X) \in \cdot)N(dt)\right], \quad B \in \mathcal{B}. \quad (5)$$

In particular, using the derivative $\lambda = \mu'(t)$ and $B = (0, 1]$,

$$P^0((N, X) \in \cdot) = \lambda^{-1} E\left[\int_{(0,1]} \mathbf{1}((S^t N, S^t X) \in \cdot)N(dt)\right].$$

Furthermore, the P_t are all equal to P^0 in the following sense.

Proposition 4 *If (N, X) is stationary, then*

$$P_t(S^t(N, X) \in \cdot) = P^0((N, X) \in \cdot), \quad \mu\text{-a.e. } t.$$

Proof This follows since (2) and (5) yield

$$\begin{aligned} \int_B P_t(S^t(N, X) \in \cdot)\mu(dt) &= E\left[\int_B \mathbf{1}(S^t(N, X) \in \cdot)N(dt)\right] \\ &= \int_B P^0((N, X) \in \cdot)\mu(dt). \end{aligned}$$

■

3 Palm Probabilities Described by Stochastic Intensities

In this section, we present a Papangelou-type theorem relating time-dependent Palm probabilities of a point process to its stochastic intensity. In addition to the usual stochastic intensity based on the entire history, we consider an intensity based only on present information.

Suppose that N and X are the processes defined above with the additional property that they are adapted to a filtration \mathcal{F}_t , $t \in \mathfrak{R}$. Assume, for each rational number s , that \mathcal{F}_s is countably generated by a π -system \mathcal{C}_s (a collection of subsets that is closed under finite intersections — if two finite measures agree on a π -system, they agree on the σ -field generated by that π -system [11]). This assumption is automatically satisfied by the filtration generated by (X, N) .

We will use the following terminology for a measurable function $f : \mathfrak{R} \times \Omega \rightarrow \mathfrak{R}$. The f is \mathcal{F}_t -adapted if $f(t, \cdot)$ is \mathcal{F}_t -measurable, for each t . The f is \mathcal{F}_t -progressive if, for any $a < t$, the set $\{(s, \omega) \in [a, t] \times \Omega : f(s, \omega) \in A\} \in \mathcal{B}[a, t] \times \mathcal{F}_t$. Finally, f is \mathcal{F}_t -predictable if

$$\{(t, \omega) \in \mathfrak{R} \times \Omega : f(t, \omega) \in B\} \in \mathcal{P}(\mathcal{F}_t), \quad B \in \mathcal{B},$$

where $\mathcal{P}(\mathcal{F}_t)$ is the σ -field generated by the rectangles $(s, t] \times A$, for $s \leq t$, $A \in \mathcal{F}_s$.

We first consider the usual time-dependent randomized intensity for N based on the information contained in the \mathcal{F}_t .

Definition 5 An \mathcal{F}_t -progressive function $\lambda : \mathfrak{R} \times \Omega \rightarrow \mathfrak{R}_+$ is an \mathcal{F}_t -intensity of N (under P) if

$$E[N(a, b) | \mathcal{F}_a] = E\left[\int_{(a, b]} \lambda(t) dt \middle| \mathcal{F}_a\right], \quad (a, b] \in \mathcal{B}.$$

Here is a key property of an intensity [4]; it is analogous to the Campbell-Mecke formula.

Theorem 6 If λ is an \mathcal{F}_t -intensity of N and $Y(t)$ is a nonnegative predictable process, then

$$E\left[\int_{\mathfrak{R}_+} Y(t) N(dt)\right] = \int_{\mathfrak{R}_+} E[Y(t)\lambda(t)] dt. \quad (6)$$

We will now present a Papangelou-type theorem that gives necessary and sufficient conditions for $P_t \ll P$ on \mathcal{F}_{t-} , for μ -a.e. t . This is non-stationary generalization of Papangelou's result, which states that for a stationary point process N on the real line, $P^0 \ll P$ on \mathcal{F}_{0-} if and only if the point process has an \mathcal{F}_t -intensity. Brémaud later extended this result to histories that aren't

necessarily the history induced by the point process. His proof includes a lemma that all predictable processes on the line have a nice form, and he uses this form along with a stationarity invariance of P to prove the result.

Our approach is different. We show that when enough σ -fields are countably generated, a well-known martingale approximation of Radon-Nikodym derivatives yields an intensity that's \mathcal{F}_t -predictable. To be precise, we will assume that \mathcal{F}_t is generated by a countable collection of sets \mathcal{C}_t , for each rational t . It can be further assumed that each \mathcal{C}_t is a π -system, since the system is closed under only a finite number of intersections.

Theorem 7 *For the point process N with intensity measure μ , the following statements are equivalent:*

- (a) N has an \mathcal{F}_t -intensity.
- (b) $P_t \ll P$ on \mathcal{F}_{t-} , μ -a.e. t and $\mu \ll \mathcal{L}$ (Lebesgue measure) .

The following proof constructs an \mathcal{F}_t -intensity λ that is \mathcal{F}_t -predictable, when $P_t \ll P$ on \mathcal{F}_{t-} . The Palm probabilities are related to λ by

$$\frac{dP_t}{dP} = \frac{\lambda(t)}{E[\lambda(t)]}.$$

Consequently, for any real-valued process $Y(t)$,

$$E_t[Y(t)] = \frac{E[Y(t)\lambda(t)]}{E[\lambda(t)]}, \quad (7)$$

provided the expectations exist. This formula is a special case of (4) when N has a stochastic intensity.

Proof (a) \Rightarrow (b). If N has an \mathcal{F}_t -intensity λ , then clearly $\mu(B) = \int_B E[\lambda(t)]dt$, and so $\mu \ll \mathcal{L}$. Next, choose $A \in \mathcal{F}$ and define $t_A = \inf\{t : A \in \mathcal{F}_t\}$ and

$$Y^A(t, \omega) = \mathbf{1}(\omega \in A)\mathbf{1}(t > t_A).$$

It is clear that Y^A is predictable. Therefore, by the Campbell-Mecke formula and (6),

$$\begin{aligned} \int_B E_t[Y^A(t)]\mu(dt) &= E\left[\int_B Y^A(t)N(dt)\right] \\ &= \int_B E[Y^A(t)\lambda(t)]dt, \quad B \in \mathcal{B}. \end{aligned}$$

From this and $\mu(dt) = E[\lambda(t)]dt$, it follows that, for any set A that belongs to one of the π -systems \mathcal{C}_r , where r is rational,

$$E_t[Y^A(t)] = \frac{E[Y^A(t)\lambda(t)]}{E[\lambda(t)]}, \quad \mu\text{-a.e. } t.$$

Fix a t in the set of points that satisfy this equality for all sets contained in $\cup_r \mathcal{C}_r$. In particular, the equality holds for all sets $A \in \mathcal{C}_r$, for any rational number $r < t$. Therefore, by a well-known monotone class argument

$$P_t(A) = \int_A \frac{\lambda(t)}{E[\lambda(t)]} dP, \quad A \in \mathcal{F}_r.$$

By using the same type of monotone class argument, we see that this expression is true for all $A \in \mathcal{F}_{t-}$. This proves $P_t \ll P$ on \mathcal{F}_{t-} , μ -a.e. t , which finishes the proof that (a) \Rightarrow (b).

(b) \Rightarrow (a) Assume (b) is true, and let h denote the density of μ . Then by the Campbell-Mecke formula,

$$\int_A N(a, b] dP = E[N(a, b] \mathbf{1}(\omega \in A)] = \int_{(a, b]} P_t(A) h(t) dt, \quad A \in \mathcal{F}_a. \quad (8)$$

Suppose for now that there is a \mathcal{F}_t -progressive nonnegative function $f(t, \omega)$ such that

$$P_t(A) = \int_A f(t, \omega) P(d\omega). \quad (9)$$

Then from (8),

$$\begin{aligned} \int_A N(a, b] dP &= \int_A \int_{(a, b]} f(t, \omega) h(t) dt P(d\omega) \\ &= \int_A E\left[\int_{(a, b]} f(t, \omega) h(t) dt \middle| \mathcal{F}_a\right] P(d\omega). \end{aligned}$$

Hence, N has a stochastic intensity.

It remains to define a \mathcal{F}_t -progressive nonnegative function $f(t, \omega)$ that satisfies (9). Consider the function

$$\begin{aligned} f(t, \omega) &= \limsup_{n \rightarrow \infty} \sum_{m=1}^{r_{0,n}} \frac{P_t(B_{0,m}^n)}{P(B_{0,m}^n)} \mathbf{1}(\omega \in B_{0,m}^n) \mathbf{1}(t \in [0, \frac{1}{2^n}]) \\ &\quad + \sum_{k=1}^{2^{2n}} \sum_{m=1}^{r_{k,n}} \frac{P_t(B_{k,m}^n)}{P(B_{k,m}^n)} \mathbf{1}(\omega \in B_{k,m}^n) \mathbf{1}(t \in (\frac{k}{2^n}, \frac{k+1}{2^n}]), \end{aligned}$$

where $\{B_{k,m}^n\}_m$ is the $(k, n)^{th}$ finite partition of Ω that consists of $\mathcal{F}_{\frac{k}{2^n}}$ -measurable sets, and $r_{k,n}$ represents the number of sets in the $(k, n)^{th}$ partition that have positive P -measure. We assume that for any fixed dyadic rational $\frac{k_0}{2^{n_0}}$, the sequence of partitions $\{B_{k_0 2^p, m}^{n_0+p}\}_m$ becomes finer and finer as p increases, and it also generates $\mathcal{F}_{\frac{k_0}{2^{n_0}}}$. We further refine our partitions so that $\{B_{k,m}^n\}_m$ becomes finer and finer as k increases (for fixed n). Notice that f is \mathcal{F}_t -predictable, which implies that it is also \mathcal{F}_t -progressive. Finally, we can conclude from a martingale approximation of Radon-Nikodym derivatives (see Application (VIII) of Section

9.5 of [8]) that for each t , $f(t)$ is the Radon-Nikodym derivative of P_t with respect to P on \mathcal{F}_{t-} , since for any fixed t , the resulting sequence of partitions must generate \mathcal{F}_{t-} . To see this, notice that for any fixed t , the corresponding sequence of partitions generate \mathcal{F}_s , for every dyadic rational $s < t$, which implies that the sequence also generates \mathcal{F}_{t-} . Thus, f satisfies (9). \blacksquare

The preceding proof shows that, for *any* \mathcal{F}_t -intensity λ of N ,

$$P_t(A) = \int_A \frac{\lambda(t)}{E[\lambda(t)]} dP, \quad A \in \mathcal{F}_{t-}$$

for all $t \in N_\lambda$ such that $\mu(N_\lambda) = 0$ (note the dependence of N_λ on the intensity λ). However, this may not necessarily be the Radon-Nikodym derivative, since the integrand may not necessarily be \mathcal{F}_{t-} -measurable. What we have shown is that *there exists* an \mathcal{F}_t -intensity that satisfies this Radon-Nikodym property.

The rest of this section contains another version of Theorem 7 based on another type of stochastic intensity, which is used in the ASTA results presented shortly.

Definition 8 A measurable process $\lambda(t)$ is an $X(t)$ -intensity for N if

$$E[N(a, b) | X(a)] = E\left[\int_{(a, b]} \lambda(s) ds \mid X(a)\right], \quad a < b.$$

Clearly, an \mathcal{F}_t -intensity of N is also an $X(t)$ -intensity of N , provided $X(t)$ is \mathcal{F}_t -measurable. What's nice about an $X(t)$ -intensity is that it only requires information about the marginal distributions of X under P_t , instead of the entire history of X up to time t .

The proof of Theorem 7 did not make explicit use of the entire history of the processes, and so here we are able to use a similar proof for an $X(t)$ -intensity.

Theorem 9 If N has an $X(t)$ -intensity $\lambda(t)$, then $P_t \ll P$ on $\sigma(X(t-))$, and $\mu \ll \mathcal{L}$. In particular, $\mu(\cdot) = \int_{(\cdot)} E[\lambda(s)] ds$ and

$$P_t(A) = \int_A \frac{\lambda(t)}{E[\lambda(t)]} dP, \quad A \in \sigma(X(t-)), \quad \mu\text{-a.e.t.} \quad (10)$$

Proof For simplicity, we prove this for real-valued X . Consider a countable collection of continuous functions $f_{n,s} : \mathfrak{R} \rightarrow [0, 1]$, for $n \in \{1, 2, \dots\}$ and rational $s \in \mathfrak{R}$, defined by

$$f_{n,s}(t) = \begin{cases} 1 & t \leq s - 1/n, \\ -n(t - s) & s - 1/n < t < s, \\ 0 & t \geq s. \end{cases}$$

Clearly $\lim_{n \rightarrow \infty} f_{n,s}(t) = \mathbf{1}(t < s)$.

Define $Y_{n,s,u}(t, \omega) = f_{n,s}(X(u)) \mathbf{1}(t > u)$. Then for any $B \in \mathcal{B}$,

$$\begin{aligned}
\int_B E_t[Y_{n,s,u}(t, \omega)]\mu(dt) &= E\left[\int_B Y_{n,s,u}(t)N(dt)\right] \\
&= E[f_{n,s}(X(u))N(B \cap (u, \infty))] \\
&= E\left[\int_B f_{n,s}(X(u))\lambda(t)\mathbf{1}(t > u)dt\right].
\end{aligned}$$

Therefore, for each integer n , and each rational pair (s, u) ,

$$E_t[f_{n,s}(X(u))\mathbf{1}(t > u)] = \frac{E[f_{n,s}(X(u))\lambda(t)]\mathbf{1}(t > u)}{E[\lambda(t)]}.$$

Fix a t that satisfies this equality for all choices of n, s, u above. Then, if t_n is a sequence that approaches t from below, the dominated convergence theorem implies

$$E_t[f_{n,s}(X(t-))] = \frac{E[f_{n,s}(X(t-))\lambda(t)]}{E[\lambda(t)]}.$$

Taking limits as $n \rightarrow \infty$, another application of the dominated convergence theorem yields

$$P_t(X(t-) < s) = \frac{E[\mathbf{1}(X(t-) < s)\lambda(t)]}{E[\lambda(t)]}$$

for any rational s , which proves the result. Our use of the dominated convergence theorem is valid, in spite of the change of measure, because the limits exist for each fixed ω , due to the fact that our sample paths are in the space $D(\mathfrak{R})$. ■

There are other types of intensities in the literature. Melamed and Whitt [15] define a *conditional intensity* of N as a process $\lambda^*(t)$ that satisfies

$$\lim_{h \downarrow 0} E[N(t, t+h)|X(t)] = \lambda^*(t).$$

It is not immediately clear how this relates to an $X(t)$ -intensity. However, if an $X(t)$ -intensity $\lambda(t)$ is right-continuous, bounded, and such that $\lambda(t)$ is $\sigma(X(t))$ -measurable, then clearly $\lambda(t) = \lambda^*(t)$. The analysis in [15] assumes uniform integrability to move limits inside of expectations when needed. Conditional intensities are also defined in [16] as $E[\lambda(t)|X(t)]$, where $\lambda(t)$ is the \mathcal{F}_t -intensity of N .

In [6], a conditional intensity is defined as a process $\mu(X(t))$ that satisfies

$$\lambda_N E_N^0[f(X(0))] = E[f(X(0))\mu(X(0))].$$

When N has an \mathcal{F}_t -intensity $\lambda(t)$, this is just $\mu(X(t)) = E[\lambda(t)|X(t)]$. To prove that such intensities can exist without assuming existence of a stochastic intensity, they consider an example where X only has jumps at points of N , and they use the inversion formula to show that this type of intensity exists. An $X(t)$ -intensity satisfies the same equality found in Papangelou's result, and it looks similar to the one found in [15].

4 Characterization of Poisson Processes

We now apply our non-stationary version of Papangelou's theorem for time-dependent Palm probabilities to characterize Poisson processes.

Using the notation above, the point process N on \mathfrak{R} is an \mathcal{F}_t -Poisson process with mean measure μ if, for any $a \leq b$, and $k \geq 0$,

$$P(N(a, b] = k | \mathcal{F}_a) = e^{-\mu(a, b]} (\mu(a, b])^k / k!.$$

In case $\mu(a, b] = \int_a^b \lambda(t) dt$, we say N is an \mathcal{F}_t -Poisson process with rate function $\lambda(t)$.

Mecke showed that a stationary point process N is \mathcal{F}_t -Poisson if and only if $P^0 = P$ on \mathcal{F}_{0-} . This is an immediate consequence of Brémaud's version of Papangelou's theorem. The following result is an analogue for time-dependent Palm probabilities. It also contains Watanabe's well-known martingale characterization of Poisson processes, which is statement (b).

Theorem 10 *The following statements are equivalent for a locally integrable function $\lambda : \mathfrak{R} \rightarrow \mathfrak{R}$.*

- (a) N is an \mathcal{F}_t -Poisson process with rate function λ .
- (b) $N(s, t] - \int_s^t \lambda(s) ds$, for $t > s$, is an \mathcal{F}_t -martingale for any fixed s .
- (c) $P_t = P$ on \mathcal{F}_{t-} for μ -a.e. t , and $\mu(B) = \int_B \lambda(t) dt$, $B \in \mathcal{B}$.

Proof (a) \Rightarrow (b). If N is \mathcal{F}_t -Poisson with deterministic rate function $\lambda(t)$, then clearly (b) is true.

(b) \Rightarrow (c). The hypothesis tells us that $E[N(s, t] | \mathcal{F}_s] = \int_{(s, t]} \lambda(u) du$ for any $s < t$, and from this we conclude that N has a deterministic \mathcal{F}_t -intensity $\lambda(t)$. From Theorem 7, we conclude (c).

(c) \Rightarrow (a). To prove (a), it suffices by properties of conditional expectations to show that, for each $A \in \mathcal{F}_s$, $s \in \mathfrak{R}$ and $n \geq 0$,

$$P(A, N(s, t] = n | \mathcal{F}_s) = P(A) \left(\int_{(s, t]} \lambda(u) du \right)^n e^{-\int_{(s, t]} \lambda(u) du} / n!, \quad t \geq s. \quad (11)$$

We first consider the case $n = 0$. Since $P(N[t, t] = 0) = 1$ for all t ,

$$P(A, N[s, t] \geq 1) = E[\mathbf{1}_A \int_{(s, t]} \mathbf{1}(N[s, u] = 0) N(du)]$$

$$\begin{aligned}
&= \int_{(s,t]} P_u(A, N[s, u] = 0) \mu(du) \\
&= \int_{(s,t]} P(A, N[s, u] = 0) \lambda(u) du
\end{aligned}$$

so we end up with the integral equation

$$P(A) - P(A, N[s, t] = 0) = \int_{(s,t]} P(A, N[s, u] = 0) \lambda(u) du.$$

From what is known about solutions to integral equations (see, e.g. [10]), we conclude that the solution is of the form (11) with $n = 0$.

To prove (11) for $n \geq 1$, all that is needed is to generate an infinite system of integral equations by applying the Campbell-Mecke formula in the same way as mentioned above to $P(A, N[s, t] \geq n + 1)$, for $n \geq 1$. In other words, for $n \geq 1$,

$$P(A) - P(A, N[s, t] \leq n) = \int_s^t P(A, N[s, u] = n) \lambda(u) du.$$

But from this, we obtain

$$\begin{aligned}
P(A, N[s, t] = n) &= \int_s^t P(A, N[s, u] = n - 1) \lambda(u) du \\
&\quad - \int_s^t P(A, N[s, u] = n) \lambda(u) du.
\end{aligned}$$

After using induction on n , and the result in [10], we see that this system is equivalent to the system found when one wants to prove that N is a nonhomogeneous Poisson process with rate function $\lambda(t)$ if and only if N has independent increments, $P(N(t, t + h] = 1) = \lambda(t)h + o(h)$, and $P(N(t, t + h] \geq 2) = o(h)$. Therefore, we obtain (11) for $n \geq 1$, and this completes the proof. \blacksquare

Note that the proof above also provides yet another proof of Watanabe's result that (a) \Leftrightarrow (b). At no point do we use the fact that (b) \Rightarrow (a). We should however point out that a similar type of argument is given in [1] to prove Watanabe's result, where they use a similar trick with stochastic integrals to compute the Laplace transform of $N(a, b]$ (the difference in the proofs involves the use of the Campbell-Mecke formula).

5 Arrivals See Time Averages

A basic issue for a queueing process is to determine the distribution of the queue length at the time of a customer arrival. For many queues, the distribution in equilibrium is the same as the distribution of the process at any arbitrary time (regardless of an arrival). That is, arrivals see time averages (ASTA). In this

section, we address the ASTA issue for the \mathcal{F}_t -adapted processes N and X on the real line. The focus is on the distribution of $X(t-)$ when N has a point at time t . Although these are abstract processes, we interpret a point of N as the time of an “arrival” of an event related to X .

Here is our main result.

Theorem 11 *If N has an $X(t)$ -intensity λ , then the following statements are equivalent.*

- (a) Transient ASTA: $P_t(X(t-) \in \cdot) = P(X(t-) \in \cdot)$, μ -a.e. t .
- (b) Lack of Bias: $E[\lambda(t)|X(t-)] = E[\lambda(t)]$, μ -a.e. t .

Hence if N is an \mathcal{F}_t -Poisson process with deterministic rate function $\lambda(t)$, then transient ASTA holds.

Proof First note that from (10) it follows that, for any bounded continuous function $f : \mathbb{E} \rightarrow \mathfrak{R}$,

$$E_t[f(X(t-))]E[\lambda(t)] = E[f(X(t-))\lambda(t)], \quad \mu\text{-a.e. } t. \quad (12)$$

Now if (a) holds, then (12) is true with E_t replaced by E . Moreover, the resulting statement also holds for all bounded measurable f (by a standard approximation argument), and hence (b) holds.

Conversely, if (b) holds, then (a) follows since by (12),

$$\begin{aligned} P_t(X(t-) \in \cdot)E[\lambda(t)] &= E[\mathbf{1}(X(t-) \in \cdot)E[\lambda(t)|X(t-)]] \\ &= P(X(t-) \in \cdot)E[\lambda(t)], \quad \mu\text{-a.e. } t. \end{aligned}$$

Finally, if N is an \mathcal{F}_t -Poisson process with deterministic rate function $\lambda(t)$, then (b) obviously holds, which implies the transient ASTA property. \blacksquare

Recall from Proposition 4 that $P_t(S^t(N, X) \in \cdot) = P^0((N, X) \in \cdot)$ for μ -a.e. t , when (N, X) is stationary. In light of this, the following ASTA result for stationary processes, which was proved in Bremaud [5] (assuming N has an \mathcal{F}_t -intensity), is an immediate consequence of Theorem 11.

Corollary 12 *If (N, X) is stationary and N has an $X(t)$ -intensity λ , then the following statements are equivalent.*

- (a) Stationary ASTA: $P^0(X(0-) \in \cdot) = P(X(0-) \in \cdot)$.
- (b) Lack of Bias: $E[\lambda(0)|X(0-)] = E[\lambda(0)]$.

Next, we describe a conditional ASTA property for non-stationary processes, which is analogous to that for stationary queueing processes with Poisson arrivals in van Doorn and Regterschot [26]. Their results involve analyzing sample path averages in the spirit of [28], and therefore they did not use Palm probabilities. The result is stated in the stationary case through the use of Palm probabilities in [1], and we will now show how these results carry over to the non-stationary setting.

As in [1], suppose N has an \mathcal{F}_t -intensity of the form $g(Y(t))$, where Y is a \mathcal{F}_t -predictable process that takes values in a space \mathbb{E}^* , and $g : \mathbb{E}^* \rightarrow \mathfrak{R}_+$. We

are interested in the points of N when the process Y is equal to a fixed value $x \in \mathbb{E}^*$. For simplicity, assume that $P(Y(t) = x) > 0$, for all $t > 0$. Then the points of N when $Y(t) = x$ are represented by the point process N_x defined by

$$N_x(A) = \int_A \mathbf{1}(Y(t) = x)N(dt).$$

Let P_t^x denote the Palm probabilities induced by the point process N_x .

Theorem 13 *Under the preceding conditions,*

$$P_t^x(X(t-) \in \cdot) = P(X(t-) \in \cdot | Y(t) = x). \quad (13)$$

Proof For any bounded \mathcal{F}_t -predictable process $C(t)$,

$$\begin{aligned} E \left[\int_{\mathfrak{R}} C(t)N_x(dt) \right] &= E \left[\int_{\mathfrak{R}} C(t)\mathbf{1}(Y(t) = x)N(dt) \right] \\ &= E \left[\int_{\mathfrak{R}} C(t)\mathbf{1}(Y(t) = x)g(x)dt \right]. \end{aligned}$$

Therefore, $\mathbf{1}(Y(t) = x)g(x)$ is an \mathcal{F}_t -intensity for N_x . Using this and (7),

$$\begin{aligned} P_t^x(X(t-) \in \cdot) &= \frac{E[\mathbf{1}(X(t-) \in \cdot, Y(t) = x)g(x)]}{g(x)P(Y(t) = x)} \\ &= P(X(t-) \in \cdot | Y(t) = x), \end{aligned}$$

which completes the proof. ■

Remark 14 The preceding result can be extended to the $X(t)$ -intensity setting as follows. Assume there exists a measurable function g such that $g(Y(t-))$ is an $(X(t), Y(t))$ -intensity of N . Then we can apply Theorem 9 to show that for $a < b$,

$$\begin{aligned} \int_{(a,b)} P_t^x(X(t-) \in \cdot)\mu(dt) &= E \int_{(a,b)} \mathbf{1}(X(t-) \in \cdot)\mathbf{1}(Y(t-) = x)N(dt) \\ &= E \int_{(a,b)} \mathbf{1}(X(t-) \in \cdot)\mathbf{1}(Y(t-) = x)g(x)dt \end{aligned}$$

so it follows as before that, for μ -a.e. t ,

$$P_t^x(X(t-) \in \cdot) = P(X(t-) \in \cdot | Y(t-) = x).$$

6 Palm Probabilities as Limits

This section characterizes Palm probabilities for the processes N and X as limits of conditional probabilities.

The fact that the Palm probability P_t is 1 that N has a point at time t suggests that Palm probabilities should have the limiting property (15). This type of limit is indeed true under the weak condition (14) on N . Similar limits are true for Palm probabilities of nicely behaved functions of X .

Theorem 15 *Suppose $\mu(\cdot) = \int_{(\cdot)} \lambda(t) dt$ and*

$$\lim_{v \downarrow t} (v-t)^{-1} E \left[N(t, v) \mathbf{1}(N(t, v) \geq 2) \right] = 0, \quad t \in \mathfrak{R}. \quad (14)$$

Then, for each $t \in \mathfrak{R}$ and $A \in \mathcal{F}$,

$$P_t(A) = \lim_{v \downarrow t} P(A | N(t, v) \geq 1). \quad (15)$$

Furthermore, if $f : \mathfrak{R} \times D(\mathfrak{R}) \rightarrow \mathfrak{R}$ is bounded and

$$\lim_{v \downarrow t} (v-t)^{-1} P(N(t, v) = 1, f(u, X) \neq f(t, X) \text{ for some } u \in (t, v]) = 0, \quad (16)$$

then

$$E_t[f(t, X)] = \lim_{v \downarrow t} E \left[f(t, X) \mid N(t, v) \geq 1 \right], \quad t \in \mathfrak{R}. \quad (17)$$

Proof Since $E[N(\cdot)] = \int_{(\cdot)} \lambda(t) dt$, we have for a.e. t

$$\lim_{v \downarrow t} (v-t)^{-1} E[N(t, v)] = \lambda(t).$$

In addition,

$$\lim_{v \downarrow t} (v-t)^{-1} P(N(t, v) \geq 1) = \lambda(t). \quad (18)$$

This follows since

$$E[N(t, v)] = P(N(t, v) \geq 1) + E \left[(N(t, v) - 1) \mathbf{1}(N(t, v) \geq 2) \right]$$

and assumption (14) implies that the last expectation is $o(v-t)$.

Next, note that by (3), for a.e. t

$$\begin{aligned} E_t[f(t, X)]\lambda(t) &= \lim_{v \downarrow t} (v-t)^{-1} \int_t^v E_u[f(u, X)]\lambda(u) du \\ &= \lim_{v \downarrow t} (v-t)^{-1} E \left[\int_t^v f(u, X) N(du) \right]. \end{aligned} \quad (19)$$

Consider the decomposition

$$\begin{aligned} E \left[\int_t^v f(u, X) N(du) \right] &= E \left[f(t, X) \mathbf{1}(N(t, v) = 1) \right] \\ &\quad + E \left[\mathbf{1}(N(t, v) = 1) \int_t^v [f(u, X) - f(t, X)] N(du) \right] \\ &\quad + E \left[\mathbf{1}(N(t, v) \geq 2) \int_t^v f(u, X) N(du) \right]. \end{aligned} \quad (20)$$

Since f is bounded, the last two expectations are $o(v - t)$ by assumptions (14) and (16), respectively. Also, the expectation on the right-hand side of (20) equals

$$\begin{aligned} & E\left[f(t, X)\mathbf{1}(N(t, v) \geq 1)\right] - E\left[f(t, X)\mathbf{1}(N(t, v) \geq 2)\right] \\ &= E\left[f(t, X)|N(t, v) \geq 1\right]P(N(t, v) \geq 1) + o(v - t). \end{aligned}$$

Applying these observations to (19) along with (18), we have

$$E_t[f(t, X)]\lambda(t) = \lim_{v \downarrow t} E\left[f(t, X)\mathbf{1}(N(t, v) \geq 1)\right]\lambda(t).$$

This proves (17). Also, (17) implies (15), since assumption (16) is not needed in the preceding argument when $f(t, X)$ does not depend on t . \blacksquare

Example 16 *Limiting ASTA*. Kleinrock [13] considered a queueing process with Poisson arrivals in which the queue length process $X(t)$ has a limiting distribution $p(n)$. His aim was to give a plausible argument that $p(n)$ was the long run fraction of time that an arrival to the system sees n customers in the system.

He begins by showing that

$$P(X(t-) = n) = \lim_{v \downarrow t} P(X(t-) = n | N(t, v) \geq 1), \quad (21)$$

and conjectured that this is the distribution of the queue length an arrival would see at time t . Then letting $t \rightarrow \infty$ in (21) and using $p(n) = \lim_{t \rightarrow \infty} P(X(t-) = n)$, he concludes that $p(n)$ should be the probability an arrival in equilibrium sees n customers in the system.

Because of our Theorem 15, we now know that his conjecture was correct. Indeed, by Theorem 15, whose assumptions are satisfied by N and X , the right-hand side of (21) is $P_t(X(t-) = n)$. Then letting $t \rightarrow \infty$ yields

$$p(n) = \lim_{t \rightarrow \infty} P_t(X(t-) = n),$$

which is what he wanted to prove.

7 Asymptotic Stationarity

We saw in Proposition 4 that if (N, X) is stationary, then P_t is equal to the stationary Palm probability P^0 in that

$$P_t(S^t(N, X) \in \cdot) = P^0((N, X) \in \cdot), \quad \mu\text{-a.e. } t.$$

This suggests that if (N, X) behaves in the distant future like a stationary process (\bar{N}, \bar{X}) , then $P_t(S^t(N, X) \in \cdot)$ should converge in some sense to $P^0((\bar{N}, \bar{X}) \in \cdot)$ as $t \rightarrow \infty$.

We will use the following type of long-run stationarity. Here \xrightarrow{w} and \xrightarrow{v} denote weak and vague convergence of measures [11].

Definition 17 The process (N, X) is *asymptotically stationary* with limit (\bar{N}, \bar{X}) if

$$\begin{aligned} S^t(N, X) &\xrightarrow{d} (\bar{N}, \bar{X}), \\ \mu_t(\cdot) = E[S^t N(\cdot)] &\xrightarrow{v} \bar{\mu}(\cdot) = E[S^t \bar{N}(\cdot)], \quad \text{as } t \rightarrow \infty. \end{aligned}$$

The first condition refers to the convergence of the process (N, X) over the “entire” real time axis shifted over a time that tends to ∞ . The second condition is an analogous convergence of the mean measure. The limit (\bar{N}, \bar{X}) is necessarily stationary, and so the mean measure $\bar{\mu}$ for \bar{N} is a constant multiple of Lebesgue measure. Many processes that have a limiting distribution are asymptotically stationary; examples include regenerative processes, ergodic Markov processes and processes that are functionals of stationary processes. These and other properties of asymptotic stationarity were developed by Szcotzka [24]; also see [7, 25].

The first result describes the convergence of P_t when N has a stochastic intensity.

Theorem 18 *Suppose N has an \mathcal{F}_t stochastic intensity $\lambda(t)$, and (N, X, λ) is asymptotically stationary with limit $(\bar{N}, \bar{X}, \bar{\lambda})$. Assume $Y(t)$ is an \mathcal{F}_t -predictable process of the form $Y(t) = g(S^t(N, X))$ for some continuous $g : D(\mathfrak{R})^2 \rightarrow \mathbb{E}$. Then*

$$P_t(Y(t) \in \cdot) \xrightarrow{w} P^0(g(\bar{N}, \bar{X}) \in \cdot), \quad \text{for } \mu\text{-a.e. } t \rightarrow \infty.$$

Proof Because of the asymptotic stationarity and the form of $Y(t)$, it follows that $(Y(t), \lambda(t)) \xrightarrow{d} (g(\bar{N}, \bar{X}), \bar{\lambda}(0))$. Then applying (7), we have as μ -a.e. $t \rightarrow \infty$,

$$\begin{aligned} P_t(Y(t) \in \cdot) &= \frac{E[\mathbf{1}(g(S^t(N, X)) \in \cdot) \lambda(t)]}{E[\lambda(t)]} \\ &\xrightarrow{w} \frac{E[\mathbf{1}(g(\bar{N}, \bar{X}) \in \cdot) \bar{\lambda}(0)]}{E[\bar{\lambda}(0)]} = \bar{P}^0(g(\bar{N}, \bar{X}) \in \cdot). \end{aligned}$$

■

Example 19 Let X denote the queue-length process of a $GI/G/c$ queue, and let N denote its renewal arrival process. For simplicity, suppose the inter-arrival time distribution has a density that’s directly Riemann integrable. Assume the joint process (X, N) regenerates at the beginning of each busy period so that the process is regenerative and has a limiting distribution. We know from [4] that $\lambda(t) = h(t - T_{N(t)})$ is an \mathcal{F}_t -intensity of N , where h is the hazard function of the interarrival times. By the renewal theorem it follows that (N, X, λ) is asymptotically stationary with a limit $(\bar{N}, \bar{X}, \bar{\lambda})$. Then Theorem 18 tells us that

$$P_t(X(t-) = k) \rightarrow P^0(\bar{X}(0-) = k) \quad \text{a.e. } t \rightarrow \infty. \quad (22)$$

In case N does not have a stochastic intensity, the Palm probabilities, which are Radon-Nikodym derivatives, are generally ill-behaved as t varies, and may not converge in a conventional sense. It is natural, however, to consider convergence in a local neighborhood of t as $t \rightarrow \infty$. To this end, we will use a local type of convergence that is a slight variation of that in [12].

Let C denote the set of continuous functions $f : \mathfrak{R} \rightarrow \mathfrak{R}_+$ with compact support. A *local f -mixture* $P_{t,f}$ of the measures P_t at t is defined by

$$P_{t,f}(\cdot) = \frac{\int_{\mathfrak{R}} P_{t+u}(S^{t+u}(N, X) \in \cdot) f(u) \mu_t(du)}{\int_{\mathfrak{R}} f(u) \mu_t(du)}, \quad f \in C,$$

where $\mu_t(B) = \mu(B + t)$. An f -mixture \bar{P}_f^0 of the stationary Palm measure \bar{P}^0 for (\bar{N}, \bar{X}) is

$$\bar{P}_f^0(\cdot) = \frac{\int_{\mathfrak{R}} \bar{P}_0(S^u(\bar{N}, \bar{X}) \in \cdot) f(u) du}{\int_{\mathfrak{R}} f(u) du}, \quad A \in \mathcal{F}, f \in C.$$

The convergence of P_t in (23) below is a local convergence of P_s for all s near t , as $t \rightarrow \infty$.

Theorem 20 *If (N, X) is asymptotically stationary with limit (\bar{N}, \bar{X}) , then*

$$P_{t,f}(\cdot) \xrightarrow{w} \bar{P}_f^0(\cdot), \quad \text{for } \mu\text{-a.e. } t \rightarrow \infty, \text{ for each } f \in C. \quad (23)$$

Proof By the definition of $P_{t,f}$ and (3),

$$P_{t,f}(\cdot) = \frac{E \left[\int_{\mathfrak{R}} \mathbf{1}(S^{t+u}(N, X) \in \cdot) f(u) S^t N(du) \right]}{\int_{\mathfrak{R}} f(u) \mu_t(du)}.$$

Now, under the asymptotic stationarity of (N, X) , we have

$$\begin{aligned} \frac{\int_{\mathfrak{R}} \mathbf{1}(S^{t+u}(N, X) \in \cdot) f(u) S^t N(du)}{\int_{\mathfrak{R}} f(u) \mu_t(du)} \\ \xrightarrow{d} \frac{\int_{\mathfrak{R}} \mathbf{1}(S^u(\bar{N}, \bar{X}) \in \cdot) f(u) \bar{N}(du)}{\int_{\mathfrak{R}} f(u) \mu(du)}. \end{aligned}$$

Applying this limit to the first display yields

$$P_{t,f}(\cdot) \xrightarrow{w} \frac{E \left[\int_{\mathfrak{R}} \mathbf{1}(S^u(\bar{N}, \bar{X}) \in \cdot) f(u) \bar{N}(du) \right]}{\int_{\mathfrak{R}} f(u) \bar{\mu}(du)} = \bar{P}_f^0(\cdot).$$

■

8 Markov and Semi-Regenerative Processes

We now present a general framework for deriving Palm probabilities for a Markov jump process X associated with a point process consisting of a subset of its jump times. This is followed by a factorization of Palm probabilities for semi-regenerative processes.

For the first part of this discussion, suppose that X is a Markov jump process on the time axis \mathfrak{R} with state space \mathbb{E} and transition-rate kernel $q(x, A)$. Let N denote the point process of its jump times T_n , $n \in \mathbb{Z}$. For simplicity, assume \mathcal{F}_t is the σ -field generated by $(X(s) : s \leq t)$. Also, for each n , define

$$X_n = X(T_n), \quad \xi_n = T_n - T_{n-1}, \quad \mathcal{F}_n = \mathcal{F}_{T_n}.$$

Being a Markov jump process means the sequence (X_n, ξ_n) is a Markov chain with transition probabilities

$$P(X_{n+1} \in A, \xi_{n+1} > t | \mathcal{F}_n) = q(X_n, A)e^{-tq(X_n)}, \quad (24)$$

where $q(x) = q(x, \mathbb{E})$.

In referring to sample paths of X , we sometimes write $X = (X_-(t), X_+(t))$, for $t \in \mathfrak{R}$, where

$$X_-(t) = \{X(s) : s < t\}, \quad X_+(t) = \{X(s) : s \geq t\},$$

which are the sample paths of X before t and $\geq t$, respectively. These random paths are in the respective Skorohod spaces $D(\mathfrak{R}_-)$ and $D(\mathfrak{R}_+)$, where $\mathfrak{R}_- = (-\infty, 0)$. We also write $f(t, X) = f(t, X_-(t), X_+(t))$. In addition, let $p(x, A)$ denote the ‘‘future-path’’ probability kernel from \mathbb{E} to $D(\mathfrak{R}_+)$ such that

$$P(X_+(T_n) \in A | \mathcal{F}_{n-1}, X_n) = p(X_n, A). \quad (25)$$

The main tool for deriving Palm probabilities for Markov processes is as follows. We will prove it after some discussion.

Theorem 21 *Under the preceding assumptions, for $f : \mathfrak{R} \times D(\mathfrak{R}) \rightarrow \mathfrak{R}_+$,*

$$\begin{aligned} E \left[\int_{\mathfrak{R}} f(t, X) N(dt) \right] & \quad (26) \\ &= \int_{\mathfrak{R}} \int_{D(\mathfrak{R}_+)} \int_{\mathbb{E}} E \left[f(t, X_-(t), z) q(X(t), dx) \right] p(x, dz) dt, \end{aligned}$$

provided the expectations are finite.

A special case is the classical *Lévy formula*: For $h : \mathfrak{R} \times \mathbb{E} \times \mathbb{E} \rightarrow \mathfrak{R}_+$,

$$E \left[\int_{\mathfrak{R}} h(t, X(t-), X(t)) N(dt) \right] = E \left[\int_{\mathfrak{R}} \int_{\mathbb{E}} h(t, X(t), x) q(X(t), dx) dt \right].$$

There are a variety of Palm probabilities of Markov processes associated with certain subsets of their jump times. As an example, for a Jackson network

process in equilibrium, the mean time for an item to move from one sector to another sector involves Palm probabilities for jump times at which an item in the network begins a journey between the two sectors [23]. Examples like this involve exploiting the structure of the special jump times, which may depend on the future as well as the past of the process. It is not practical to formulate Palm probabilities for all such contingencies. However, the following general procedure and example illustrate how one can use Theorem 21 to derive Palm probabilities for Markov processes.

Remark 22 *General Procedure.* A subprocess of jump times of X has the form

$$\tilde{N}(B) = \int_B Y(t)N(dt) = \sum_n Y(T_n)\mathbf{1}(T_n \in B), \quad B \in \mathcal{B},$$

where $Y(t)$ is a process that takes values 0 or 1. Let \tilde{P}_t and \tilde{E}_t denote the Palm probabilities and expectations associated with (X, \tilde{N}) . These probabilities are determined as follows.

For any bounded $f : \mathfrak{R}_+ \times D(\mathfrak{R}) \rightarrow \mathfrak{R}_+$, use Theorem 21 to determine functions g and h that satisfy

$$\begin{aligned} E\left[\int_B f(t, X)\tilde{N}(dt)\right] &= E\left[\int_{\mathfrak{R}} f(t, X)Y(t)N(dt)\right] = \int_B g(t)dt, \\ E[\tilde{N}(B)] &= E\left[\int_B Y(t)N(dt)\right] = \int_B h(t)dt, \quad B \in \mathcal{B}. \end{aligned}$$

Then it follows that

$$\tilde{E}_t[f(t, X)] = \frac{g(t)}{h(t)}, \quad \text{a.e. } t.$$

Example 23 Consider the jump times of X at which its states before and after the jump are in a fixed set $A \in \mathcal{E}^2$. These times are depicted by the point process

$$\tilde{N}(B) = \int_B \mathbf{1}\left((X(t-), X(t)) \in A\right)N(dt), \quad B \in \mathcal{B}.$$

For any bounded $f : \mathfrak{R}_+ \times D(\mathfrak{R}) \rightarrow \mathfrak{R}_+$, Theorem 21 yields

$$\begin{aligned} E[\tilde{N}(B)] &= E\left[\int_{\mathfrak{R}} \mathbf{1}\left((X(t-), X(t)) \in A, t \in B\right)N(dt)\right] \\ &= \int_B E\left[\int_{\mathbf{E}} \mathbf{1}\left((X(t-), x) \in A\right)q(X(t), dx)\right]dt \\ &= \int_B E[q(X(t), A_{X(t)})]dt, \\ E\left[\int_B f(t, X)\tilde{N}(dt)\right] &= E\left[\int_{\mathfrak{R}} f(t, X)\mathbf{1}\left((X(t-), X(t)) \in A, t \in B\right)N(dt)\right] \\ &= \int_B g(t)dt, \end{aligned}$$

where for each $y \in \mathbb{E}$, $A_y = \{x \in \mathbb{E} : (y, x) \in \mathcal{E}^2\}$, and

$$g(t) = E \left[\int_{A_{X(t)}} \left(\int_{D(\mathbb{R}_+)} f(t, X_-(t), z) p(y, dz) \right) q(X(t), dy) \right].$$

Then by the preceding General Procedure,

$$\tilde{E}_t[f(t, X)] = \frac{g(t)}{E[q(X(t), A_{X(t)})]}, \quad \text{a.e. } t \quad (27)$$

One can obtain additional properties of \tilde{N} in the preceding example by using the techniques above. For instance, $q(X(t-), A_{X(t-)})$ is an \mathcal{F}_t -stochastic intensity for \tilde{N} . This implies the classical result that $q(X(t-))$ is an \mathcal{F}_t -stochastic intensity of N . Indeed, $\tilde{N} = N$ when $A = \mathbb{E}^2$. Here is a limiting property of \tilde{N} .

Remark 24 In the setting of Example 23, suppose that X is an ergodic Markov process and let (\bar{N}, \bar{X}) denote a stationary version of (N, X) with Palm probability \bar{P}^0 . Then

$$P_t(S^t X \in \cdot) \xrightarrow{w} \bar{P}^0(\bar{X} \in \cdot), \quad \text{for a.e. } t \rightarrow \infty$$

This follows by taking the limit of (27), with $f(t, X) = \mathbf{1}(S^t X \in \cdot)$, as $t \rightarrow \infty$.

We are now ready to prove Theorem 21. A key step in the proof uses the following property of exponential random variables. This follows by writing the expectations as integrals of the exponential density, and using $e^{-\lambda t} = \lambda \int_t^\infty e^{-\lambda u} du$ and interchanging integrals.

Proposition 25 *If ξ is an exponential random variable with rate λ , then for any $h : \mathbb{R}_+ \rightarrow \mathbb{R}$,*

$$E[h(\xi)] = \lambda E \left[\int_0^\xi h(u) du \right],$$

provided the expectations exist.

Proof of Theorem 21 First note that the Markov property (24) implies that X_n is a Markov chain and

$$P(X_n \in A | \mathcal{F}_{n-1}) = q(X_{n-1}, A) / q(X_{n-1}), \quad (28)$$

$$P(\xi_n > u | \mathcal{F}_{n-1}, X_n) = e^{-q(X_{n-1})u}. \quad (29)$$

Now, conditioning on $\mathcal{G}_n(x, z) = (\mathcal{F}_{n-1}, X_n = x, X_+(T_n) = z)$,

$$\begin{aligned} E \left[\int_{\mathbb{R}} f(t, X) N(dt) \right] &= \sum_n E[f(T_n, X)] \\ &= E \left[\sum_n \int_{\mathbb{E}} \int_{D(\mathbb{R}_+)} E[f(T_n, X) | \mathcal{G}_n(x, z)] \right. \\ &\quad \left. \times P(X_n \in dx | \mathcal{F}_{n-1}) P(X_+(T_n) \in dz | \mathcal{F}_{n-1}, X_n = x) \right]. \end{aligned} \quad (30)$$

Using $X = (X_-(T_n), X_+(T_n))$, $T_n = T_{n-1} + \xi_n$, expression (29), Proposition 25 and the change-of-variable $t = T_{n-1} + u$, we have

$$\begin{aligned} E[f(T_n, X)|\mathcal{G}_n(x, z)] &= E[f(T_{n-1} + \xi_n, X_-(T_{n-1} + \xi_n), z)|\mathcal{G}_n(x, z)] \\ &= q(X_{n-1})E\left[\int_0^{\xi_n} f(T_{n-1} + u, X_-(T_{n-1} + u), z)du|\mathcal{G}_n(x, z)\right] \\ &= q(X_{n-1})E\left[\int_{T_{n-1}}^{T_n} f(t, X_-(t), z)dt|\mathcal{G}_n(x, z)\right]. \end{aligned}$$

Also, by (28) and (25) the last line in (30) equals $q(X_{n-1}, dx)q(X_{n-1})^{-1}p(x, dz)$. Substituting this and the last display into (30) yields

$$\begin{aligned} E\left[\int_{\mathfrak{R}} f(t, X)N(dt)\right] \\ = E\left[\sum_n \int_{T_{n-1}}^{T_n} \int_{\mathbb{E}} \int_{D(\mathfrak{R}_+)} f(t, X_-(t), z)q(X_{n-1}, dx)p(x, dz)dt\right]. \end{aligned}$$

The last expression equals the right-hand side of (26) since $X(t) = X_{n-1}$ for $t \in (T_{n-1}, T_n)$. This completes the proof of (26). \blacksquare

We end this section with a factorization of Palm probabilities for semi-regenerative processes. Instead of X being Markovian, assume it is *semi-regenerative* with respect to N in the sense that, for any integer n ,

$$P(X(T_n)_+ \in A|\mathcal{F}_{T_n}) = p(X(T_n), A), \quad A \in \mathcal{E}, \quad (31)$$

where $p(x, A)$ is a probability kernel from \mathbb{E} to $D(\mathfrak{R}_+)$.

Proposition 26 *In the preceding context,*

$$P_t(X(t)_+ \in A) = \int_{\mathbb{E}} p(x, A)P_t(X(t) \in dx), \quad A \in \mathcal{E}, \mu\text{-a.e. } t.$$

Proof Using (3) twice and the semi-regenerative property, for $B \in \mathcal{B}$,

$$\begin{aligned} \int_B P_t(X(t)_+ \in A)\mu(dt) &= E\left[\int_B \mathbf{1}(X(t)_+ \in A)N(dt)\right] \\ &= E\left[\sum_n P(X(T_n)_+ \in A, T_n \in B|\mathcal{F}_{T_n})\right] \\ &= E\left[\sum_n p(X(T_n), A)\mathbf{1}(T_n \in B)\right] \\ &= \int_B E_t[p(X(t), A)]\mu(dt) \\ &= \int_B \int_{\mathbb{E}} p(x, A)P_t(X(t) \in dx)\mu(dt). \end{aligned}$$

The third equality follows since T_n is \mathcal{F}_{T_n} -measurable. \blacksquare

Example 27 *GI/M/1 Queue.* Suppose that X is the queue-length process of an $GI/M/1$ queue with FIFO discipline, and that N denotes the arrival process. Assume X is stable, and so it is regenerative. Let $W(t)$ denote the waiting time of the last customer to enter the system at or before time t . Then $P_t(W(t) \leq w)$ is the waiting time distribution of an arrival at time t . Using Proposition 26, the Laplace transform of this distribution is

$$\begin{aligned} E_t[e^{-\alpha W(t)}] &= \sum_{n=1}^{\infty} E[e^{-\alpha W(t)} | X(t) = n] P_t(X(t) = n) \\ &= \sum_{n=1}^{\infty} \phi(\alpha)^n P(X(t-) = n - 1) \\ &= \phi(\alpha) E \left[\phi(\alpha)^{X(t)} \right]. \end{aligned}$$

Here $\phi(\alpha)$ is the Laplace transform of the service times. From this expression, it follows that

$$E_t[W(t)] = \eta(E[X(t)] + 1), \quad \text{a.e. } t, \quad (32)$$

where η is the mean service time.

9 Time-Dependent Little Laws

For this discussion, consider a queueing system in which the point process N denotes the arrival times T_n , which are assumed to be distinct (N is simple). Let $X(t)$ denote the number of items in the system at time t , let W_n denote the sojourn time in the system of the n th job that arrives at time T_n , and let $W(t)$ denote the sojourn time of the last item that arrived before or at time t . If (N, X) is stationary, the mean queue length is related to the mean sojourn time by the well-known Little law

$$E[X(0)] = \lambda E^0[W(0)].$$

Here $\lambda = E[N(0, t)]$ is the rate of arrivals, and E^0 is the Palm expectation given an arrival at time 0.

In this section, we present several Little laws for relating the time-dependent queue length to the Palm probabilities of the waiting times. Bertsimas and Mourtzinou [2] obtain results like some of ours, using probabilities that can be intuitively interpreted as Palm probabilities and sample path arguments. This approach requires additional assumptions on the existence of relevant limits and assumes the mean measure of the arrival process has an intensity. The Palm calculus approach lays bare the relation between queue lengths and waiting times and yields results that are not amenable to sample path arguments.

For simplicity, we consider the queueing process on \mathfrak{R}_+ and assume $X(0) = 0$. The results are easily extendable to the case $X(0) > 0$. Keep in mind that there are no assumptions on the order in which the items are processed and the service times; such assumptions, however, would typically be needed for computations.

Here is a result that is well-known for $M/G/\infty$ systems, where Palm probabilities are not needed.

Proposition 28 *For the queueing system described above,*

$$E[X(t)] = \int_{(0,t]} P_s(W(s) > t - s)\mu(ds), \quad t \geq 0.$$

Proof This follows by applying the Campbell-Mecke formula to

$$X(t) = \sum_n \mathbf{1}(T_n + W_n > t, T_n \leq t) = \int_{(0,t]} \mathbf{1}(W(s) > t - s)N(ds). \quad (33)$$

■

This was proved in Riaño [20], where the main interest was in controlling the input process to achieve a desired output process. A sample-path version of the result is in [2].

In many queues, the entire distribution of $X(t)$ has the following nice relation with its sojourn times.

Theorem 29 (Distributional Little Law) *Suppose the queueing discipline is overtake-free (the output order is the same as the input order). Then*

$$P(X(t) \geq n) = \int_0^t P_s(W(s) > t - s, N(s, t] = n - 1)\mu(ds). \quad (34)$$

If in addition, $W(s)$ is independent of N under P_s on the interval (s, ∞) , for μ -a.e.s, then

$$E[z^{X(t)}] = 1 + (z - 1) \int_0^t P_s(W(s) > t - s)E_s[z^{N(s, t)}]\mu(ds). \quad (35)$$

Proof Assertion (34) follows by applying the Campbell-Mecke formula to

$$\mathbf{1}(X(t) \geq n) = \int_{(0,t]} \mathbf{1}(W(s) > t - s, N(s, t] = n - 1)N(ds).$$

Next, note that under the independence assumption, (34) becomes

$$P(X(t) \geq n) = \int_0^t P_s(W(s) > t - s)P_s(N(s, t] = n - 1)\mu(ds).$$

The generating function of $X(t)$ can now be easily derived with simple algebra. ■

Formula (34) was proved in [2] under the additional assumption that future arrivals do not influence the service times of all customers currently in the overtake-free system. In the Palm context, this is the same as the independence assumption for (34).

We end by showing how to use Palm probabilities given that N has multiple points at certain locations. For a two-point case, under our assumptions on \mathcal{F} , there exists a (μ_2 -a.e. unique) probability kernel $P_{\mathbf{t}}(A)$ such that

$$E[N(B_1)N(B_2)\mathbf{1}(\omega \in A)] = \int_{B_1} \int_{B_2} P_{(t_1, t_2)}(A) \mu_2(d(t_1, t_2))$$

where $\mu_2(B_1 \times B_2) = E[N(B_1)N(B_2)]$. One can interpret $P_{(t_1, t_2)}(A)$ as the probability of A , given that N has points at t_1 and t_2 . Furthermore, it's clear that this idea can be carried further to construct Palm probabilities $P_{\mathbf{t}}$ that condition on the locations $\mathbf{t} = (t_1, \dots, t_n)$ of $n \geq 1$ points based on

$$\mu_n(B_1 \times B_2 \times \dots \times B_n) = E\left[\prod_{k=1}^n N(B_k)\right], \quad B_k \in \mathcal{B}.$$

Theorem 30 (Little Law for Moments) For any $n \geq 1$,

$$E[X(t)^n] = \int_{(0, t]^n} P_{\mathbf{s}}(W(s_1) > t - s_1, W(s_2) > t - s_2, \dots, W(s_n) > t - s_n) \mu_n(d\mathbf{s}).$$

Proof For simplicity, we will prove the assertion for $n = 2$, which is

$$E[X(t)^2] = \int_{(0, t]^2} P_{\mathbf{s}}(W(s_1) > t - s_1, W(s_2) > t - s_2) \mu_2(d\mathbf{s}). \quad (36)$$

The more general formula follow similarly. From the representation (36),

$$X(t)^2 = \int_{(0, t]^2} \mathbf{1}(W(s_1) > t - s_1) \mathbf{1}(W(s_2) > t - s_2) N(ds_1) N(ds_2).$$

Then (36) follows by repeated applications of the Campbell-Mecke formula and Lemma 11.2 in [12]. ■

References

- [1] Baccelli, F. and P. Brémaud (2003). *Elements of Queueing Theory: Palm-Martingale Calculus and Stochastic Recurrences*. Springer, New York.
- [2] Bertsimas, D. and G. Mourtzinou (1997). Transient laws of non-stationary queueing systems and their applications. *Queueing Systems* **25** 115–155.
- [3] Brandt, A. and G. Last (1995). *Marked Point Processes on the Real Line: The Dynamic Approach*. Springer-Verlag, New York.

- [4] Brémaud, P. (1981). *Point Processes and Queues*. Springer, New York.
- [5] Brémaud, P. (1989). Characteristics of queueing systems observed at events and the connection between stochastic intensity and Palm probability. *Queueing Systems* **5** 99–112.
- [6] Brémaud, P., Kannurpatti, R. and Mazumdar, R. (1992). Event and Time Averages: A Review. *Adv. Appl. Prob.* **24** 377–411.
- [7] Chang, K., Serfozo, R. F. and W. Szczotka (1998). Tree-like queueing networks: asymptotic stationarity and heavy traffic. *Ann. Appl. Probab.*, **8**, 541–568.
- [8] Chung, K.L. (2001). *A Course in Probability Theory*. Academic Press, San Diego.
- [9] Daley, D. and D. Vere-Jones (1988). *An Introduction to the Theory of Point Processes*. Springer, New York.
- [10] Davis, M.H.A. (1993). *Markov Models and Optimization*. Chapman and Hall, London.
- [11] Kallenberg, O. (2002). *Foundations of Modern Probability*. Springer, New York.
- [12] Kallenberg, O. (1983). *Random Measures*. Akademie-Verlag, Berlin.
- [13] Kleinrock, L. (1975). *Queueing Systems, Volume I: Theory*. John Wiley and Sons, New York.
- [14] Mecke, J. (1967). Stationäre Zufällige auf Lokalkompakten Abelschen Gruppen. *Z. Wahrscheinlich keitsth.* **8**, 39-56.
- [15] Melamed, B. and W. Whitt (1990a). On arrivals that see time averages. *Operations Research*. **38**, No.1, 156-172.
- [16] Melamed, B. and W. Whitt (1990b). On arrivals that see time averages: a martingale approach. *J. of Appl. Prob.* **27** 376–384.
- [17] Melamed, B. and D. Yao (1995). The ASTA property in queueing. *Advances in Queueing: Theory, Methods, and Open Problems* (J. H. Dshalalow, Ed.) 195–224, CRC Press.

- [18] Nieuwenhuis, G. (2005). Dominated families of shifted Palm distributions. Technical Report, Tilburg University, The Netherlands.
- [19] Papangelou, F. (1972). Integrability of expected increments of point processes and a related change of time. *Trans. Amer. Math. Soc.* **165** 483–506
- [20] Riaño, G. (2002). *Transient Behavior of Stochastic Networks: Application to Production Planning with Load-Dependent Lead Times*. Ph. D. thesis, Georgia Institute of Technology.
- [21] Rolski, T. (1989). Relationships Between Characteristics in Periodic Poisson Queues. *Queueing Systems* **4**, 17–26.
- [22] Ryll-Nardzewski, C. (1961). Remarks on process of call. *Proc. 4th Berkeley Symposium, vol 2*. University of California Press, Berkeley, 455–465.
- [23] Serfozo, R. (1999). *Introduction to Stochastic Networks*. Springer-Verlag, New York.
- [24] Szczotka, W. (1986). Stationary representation of queues I and II. *Adv. in Appl. Prob.*, **18**, 815–859.
- [25] Szczotka, W. (1992). Weak convergence under mappings. *Probab. Math, Statist.* **13**, 39–58.
- [26] van Doorn, E.A., and Regterschot, G.J.K. (1988). Conditional *PASTA*. *Operations Research Letters* **7**, October No. 5, 229–232.
- [27] Watanabe, S. (1964). On discontinuous additive functionals and Lévy measures of a Markov process. *Japanese J. Math.*, **34**, 53–70.
- [28] Wolff, R. (1982). Poisson arrivals see time averages. *Operations Research* **30** March-April No.2, 223–231.