8. Stationary Marked Point Processes

Many areas of engineering and statistics involve the study of a sequence of random events, described by points occurring over time (or space), together with a *mark* for each such point that contains some further information about it (type, class, etc.). Examples include image analysis, stochastic geometry, telecommunications, credit or insurance risk, discrete-event simulation, empirical processes, and general queueing theory. In telecommunications, for example, the events might be the arrival times of requests for bandwidth usage, and the marks the bandwidth capacity requested. In a mobile phone context, the points could represent the locations (at some given time) of all mobile phones, and the marks 1 or 0 as to whether the phone is in use or not. Such a stochastic sequence is called a *random marked point process*, an MPP for short. In a stationary stochastic setting (e.g., if we have moved our origin far away in time or space, so that moving further would not change the distribution of what we see) there are two versions of an MPP of interest depending on how we choose our origin: point-stationary and time-stationary (space-stationary). The first randomly chooses an event point as the origin, whereas the second randomly chooses a time (or space) point as the origin. Fundamental mathematical relationships exists between these two versions allowing for nice applications and computations. In what follows, we present this basic theory with emphasis on one-dimensional processes over time, but also include some recent results for d-dimensional Euclidean space, $\mathbb{R}^d$.

This chapter will primarily deal with marked point processes with points on the real line (time). Spatial point processes with points in $\mathbb{R}^d$ will be touched upon in the final section; some of the deepest results in multiple dimensions have only come about recently.

Topics covered include point- and time-stationarity, inversion formulas, the Palm distribution, Campbell's formula, MPPs jointly with a stochastic process, the rate conservation law, conditional intensities, and ergodicity.

8.1 **Basic Notation and Terminology** 138
 8.1.1 The Sample Space as a Sequence Space 138
 8.1.2 Two-sided MPPs 138
 8.1.3 Counting Processes 138
 8.1.4 Forward and Backward Recurrence Times . 138
 8.1.5 MPPs as Random Measures: Campbell's Theorem 139
 8.1.6 Stationary Versions 139
 8.1.7 The Relationship Between Ψ, Ψ^0 and Ψ^* 141
 8.1.8 Examples 142

8.2 **Inversion Formulas** 144
 8.2.1 Examples 144
 8.2.2 The Canonical Framework 145

8.3 **Campbell's Theorem for Stationary MPPs** 145
 8.3.1 Little's Law 145
 8.3.2 The Palm–Khintchine Formula 145

8.4 **The Palm Distribution: Conditioning in a Point at the Origin** 146

8.5 **The Theorems of Khintchine, Korolyuk, and Dobrushin** 146

8.6 **An MPP Jointly with a Stochastic Process** 147
 8.6.1 Rate Conservation Law 147

8.7 **The Conditional Intensity Approach** 148
 8.7.1 Time Changing to a Poisson Process 149
 8.7.2 Papangelou's Formula 149

8.8 **The Non-Ergodic Case** 150

8.9 **MPPs in $\mathbb{R}^d$** 150
 8.9.1 Spatial Stationarity in $\mathbb{R}^d$ 151
 8.9.2 Point Stationarity in $\mathbb{R}^d$ 151
 8.9.3 Inversion and Voronoi Sets 151

References 152

8.1 Basic Notation and Terminology

Here the basic framework is presented for MPPs on the real line, with the points distributed over time.

8.1.1 The Sample Space as a Sequence Space

A widely used class of MPPs has events corresponding to points in time,

$$0 \le t_0 < t_1 < t_2 < \cdots , \quad \lim_{n\to\infty} t_n = \infty . \tag{8.1}$$

An MPP is then defined as a stochastic sequence; a sequence of random variable (RVs),

$$\Psi = \{(t_n, k_n) : n \ge 0\} ,$$

where the marks k_n take values in a general space $\mathbb{K}$, the mark space, which is assumed to be a complete separable metric space, where the sample-paths of Ψ satisfy (8.1). (It helps to imagine that the arrivals correspond to *customers* arriving to some fixed location over time, each one bringing with them an object called their mark: the n-th customer arrives at time t_n and brings mark k_n.) Alternatively, with $T_n \stackrel{\text{def}}{=} t_{n+1} - t_n,\ n \ge 0$ denoting the n-th *interevent* (interarrival) time, Ψ can equivalently be defined by its *interevent time representation* $\{t_0, \{(T_n, k_n) : n \ge 0\}\}$.

Letting $\mathbb{R}_+$ and $\mathbb{Z}_+$ denote the non-negative real numbers and non-negative integers respectively, $\mathbb{S} = (\mathbb{R}_+ \times \mathbb{K})^{\mathbb{Z}_+}$ denotes sequence space, endowed with the product topology and corresponding Borel σ-field. $\boldsymbol{s} = \{(y_n, k_n) : n \in \mathbb{Z}_+\} \in \mathbb{S}$ denotes a sequence.

$\mathbb{M} \stackrel{\text{def}}{=} \{\boldsymbol{s} \in \mathbb{S} : \boldsymbol{s}$ satisfies (8.1)$\}$, and is the space of marked point processes with mark space $\mathbb{K}$, that is, the *MPP space*. Elements of $\mathbb{M}$ are denoted by $\psi = \{(t_n, k_n)\} \in \mathbb{M}$; they are the sample paths of an MPP $\Psi : \Omega \to \mathbb{M}$, formally a mapping from a probability space Ω into $\mathbb{M}$ with some underlying probability P. [It is standard to suppress the dependency of the random elements on $\omega \in \Omega$; e.g., $t_n(\omega), k_n(\omega),\ \Psi(\omega)$.] When $\Omega = \mathbb{M}$, this is called the *canonical* representation of Ψ. The sequence of points themselves, without marks, $\{t_n\}$, is called a *point process*.

The probability distribution of Ψ is denoted by $\boldsymbol{P} \stackrel{\text{def}}{=} P(\Psi \in \cdot)$; it is a distribution on the Borel sets $\mathcal{E} \subset \mathbb{M}$; $\boldsymbol{P}(\mathcal{E}) = P(\Psi \in \mathcal{E})$.

Two MPPs Ψ_1 and Ψ_2 are said to have the same distribution if $P(\Psi_1 \in \mathcal{E}) = P(\Psi_2 \in \mathcal{E})$ for all Borel sets $\mathcal{E} \subset \mathbb{M}$; equivalently *all finite-dimensional distributions of the two sequences are identical*, e.g., they agree for all Borel sets of the form

$$\begin{aligned}\mathcal{E} = \{\psi \in \mathbb{M} : t_{n_0} \le s_0, k_{n_0} \in K_0, \dots ,\\ t_{n_l} \le s_l, k_{n_l} \in K_l\} ,\end{aligned}$$

where $0 \le n_0 < \cdots < n_l, l \ge 0, s_i \ge 0, K_i \subset \mathbb{K}, 0 \le i \le l$.

The assumption (8.1) of strict monotonicity, $t_n < t_{n+1},\ n \ge 0$, can be relaxed to $t_n \le t_{n+1},\ n \ge 0$, to accommodate *batch* arrivals, such as busloads or other groups that arrive together, but if the inequalities are strict, then the MPP is called a *simple* MPP.

8.1.2 Two-sided MPPs

With $\mathbb{Z}$ denoting all integers, a *two-sided* MPP, $\Psi = \{(t_n, k_n) : n \in \mathbb{Z}\}$, has points defined on all of the real line $\mathbb{R}$ thus allowing for arrivals since the infinite past;

$$\cdots t_{-2} < t_{-1} < t_0 \le 0 < t_1 < t_2 < \cdots . \tag{8.2}$$

(In this case, by convention, $t_0 \le 0$.)

8.1.3 Counting Processes

For an MPP $\psi \in \mathbb{M}$, let $N(t) = \sum_j I\{t_j \in (0, t]\}$ denote the number of points that occur in the time interval $(0, t],\ t > 0$. ($I\{B\}$ denotes the indicator function for the event B.) $\{N(t) : t \ge 0\}$ is called the *counting process*. By convention $N(0) \stackrel{\text{def}}{=} 0$. For $0 \le s \le t$, $N(s, t] \stackrel{\text{def}}{=} N(t) - N(s)$, the number of points in $(s, t]$.

In a two-sided framework, counting processes can be extended by defining $N(-t) = \sum_j I\{t_j \in (-t, 0]\}$, the number of points in $(-t, 0],\ t \ge 0$. In this case

$$t_j = \begin{cases} \inf\{t > 0 : N(t) \ge j\},\ j \ge 1 ; \\ -\inf\{t > 0 : N(-t) \ge j+1\},\ j \le 0 , \end{cases}$$

and, for $t > 0$, $N(t) = \max\{j \ge 1 : t_j \le t\}$; $t_{N(t)}$ is thus the last point before or at time t, and $t_{N(t)+1}$ is the first point strictly after time t; $t_{N(t)} \le t < t_{N(t)+1}$. $T_{N(t)} = t_{N(t)+1} - t_{N(t)}$ is the interarrival time that covers t. Note that $\{t_j \le t\} = \{N(t) \ge j\},\ \ j \ge 1$: an obvious but useful identity. For example, in a stochastic setting it yields $P(N(t) = 0) = P(t_1 > t)$. [In the one-sided case, $P(N(t) = 0) = P(t_0 > t)$.]

For a fixed mark set $K \subset \mathbb{K}$, let $N_K(t) = \sum_j I\{t_j \in (0, t],\ k_j \in K\}$, the counting process of points restricted to the mark set K. The MPP corresponding to $\{N_K(t)\}$ is sometimes referred to as a *thinning* of ψ by the mark set K.

Counting processes uniquely determine the MPP, and can be extended to measures, as will be presented in Sect. 8.1.5.

8.1.4 Forward and Backward Recurrence Times

The *forward recurrence time* is defined by

$$A(t) \stackrel{\text{def}}{=} t_{N(t)+1} - t = \begin{cases} t_0 - t, & \text{if } 0 \le t < t_0 \,; \\ t_{n+1} - t, & \text{if } t_n \le t < t_{n+1},\ n \in \mathbb{Z}_+ \,. \end{cases}$$

It denotes the time until the next event strictly after time t and is also called the *excess* at time t. At an arrival time t_n, $A(t_n-) = 0$ and $A(t_n) = A(t_n+) = T_n$, then it decreases down to zero linearly with rate one, making its next jump at time t_{n+1} and so on.

Similarly we can define the *backward recurrence time*

$$B(t) \stackrel{\text{def}}{=} t - t_{N(t)} = \begin{cases} t, & \text{if } 0 \le t < t_0 \,; \\ t - t_n, & \text{if } t_n \le t < t_{n+1},\ n \in \mathbb{Z}_+ \,, \end{cases}$$

which denotes the time since the last event prior to or at time t. In particular, $B(t) \le t$ and $B(0) = 0$. $B(t)$ is also called the *age* at time t. At an arrival time t_{n+1}, $B(t_{n+1}-) = T_n$ and $B(t_{n+1}+) = 0$ and then increases to T_{n+1} linearly with rate one. The sample paths of A and B are mirror images of each other.

In a two-sided framework, $A(t) = t_{n+1} - t$ and $B(t) = t - t_n$, if $t_n \le t < t_{n+1}$, $n \in \mathbb{Z}$; $B(t)$ is no longer bounded by t, $B(0) = |t_0|$ and $A(0) = t_1$ [recall (8.2)].

$S(t) = B(t) + A(t) = t_{N(t)+1} - t_{N(t)} = T_{N(t)}$ is called the *spread* or *total lifetime* at time t; $S(t) = T_n$ if $t_n \le t < t_{n+1}$, and is therefore piecewise constant. In a two-sided framework, $S(0) = |t_0| + t_1$.

In the context of consecutively replaced light bulbs at times t_n with lifetimes $\{T_n\}$, $A(t)$ denotes the remaining lifetime of the bulb in progress at time t, while $B(t)$ denotes its age. $S(t)$ denotes the total lifetime of the bulb in progress.

8.1.5 MPPs as Random Measures: Campbell's Theorem

An MPP ψ can equivalently be viewed as a σ-finite $\mathbb{Z}_+$ valued measure

$$\psi = \sum_j \delta_{(t_j, k_j)} \,,$$

on (the Borel sets of) $\mathbb{R} \times \mathbb{K}$, where $\delta_{(t_j,k_j)}$ is the Dirac measure at (t_j, k_j). For $A \subset \mathbb{R}$ and $K \subset \mathbb{K}$, $\psi(A \times K) =$ the number of points that occur in the time set A with marks taking values that fall in K;

$$\psi(A \times K) = \sum_j I(t_j \in A,\ k_j \in K) \,.$$

$\psi(A \times \mathbb{K}) < \infty$ for all bounded sets A. If $g = g(t, k)$ is a real-valued measurable function on $\mathbb{R} \times \mathbb{K}$, then the integral $\psi(g)$ is given by

$$\psi(g) = \int g \,\mathrm{d}\psi = \int g(t,k)\psi(\mathrm{d}t, \mathrm{d}k) = \sum_j f(t_j, k_j) \,.$$

An MPP Ψ can thus be viewed as a random measure and ν denotes its *intensity measure* on $\mathbb{R} \times \mathbb{K}$, defined by $\nu(A \times K) = E[\Psi(A \times K)]$, the expected value; $\nu(\mathrm{d}t, \mathrm{d}k) = E[\Psi(\mathrm{d}t, \mathrm{d}k)]$. Starting first with simple functions of the form $g(t,k) = I\{t \in A, k \in K\}$ and then using standard approximation arguments leads to

Theorem 8.1 (Campbell's theorem)
For any non-negative measurable function $g = g(t,k)$,

$$E\big[\Psi(g)\big] = \int g \,\mathrm{d}\nu \,.$$

8.1.6 Stationary Versions

An MPP can be stationary in one of two ways, either with respect to point shifts or time shifts (but not both); the basics are presented here.

Define for each $s \ge 0$, the MPP ψ_s by

$$\begin{aligned} \psi_s &= \{[t_n(s), k_n(s)] : n \in \mathbb{Z}_+\} \\ &\stackrel{\text{def}}{=} \{(t_{N(s)+n+1} - s, k_{N(s)+n+1}); n \in \mathbb{Z}_+\} \,, \end{aligned} \tag{8.3}$$

the MPP obtained from ψ by shifting to s as the origin and relabeling the points accordingly. For $s \ge 0$ fixed, there is a unique $m \ge 0$ such that $t_m \le s < t_{m+1}$, in which case $t_0(s) = t_{m+1} - s$; $t_1(s) = t_{m+2} - s$; and $t_n(s) = t_{m+n+1} - s$ for $n \in \mathbb{Z}_+$. Similarly, the marks become $k_0(s) = k_{m+1}$; and $k_n(s) = k_{m+n+1}$ for $n \in \mathbb{Z}_+$.

When choosing $s = t_j$, a particular point, then ψ_s is denoted by $\psi_{(j)}$. In this case ψ is shifted to the point t_j so $\psi_{(j)}$ always has its initial point at the origin: $t_0(t_j) = 0$, $j \ge 0$.

The mappings from $\mathbb{M} \to \mathbb{M}$ taking ψ to ψ_s and ψ to $\psi_{(j)}$ are called *shift mappings*.

Applying these shifts to the sample paths of an MPP Ψ yields the shifted MPPs Ψ_s and $\Psi_{(j)}$. It is noteworthy that, while Ψ_s is a deterministic shift of Ψ, $\Psi_{(j)}$ is a random shift because $t_j = t_j(\omega)$ depends upon the sample path.

In a two-sided framework, the shifts also include (and relabel) all points to the left of s, and s can be negative too.

Point Stationarity

Definition 8.1
Ψ is called a point-stationary MPP if $\Psi_{(j)}$ has the same distribution as Ψ for all $j \in \mathbb{Z}_+$. Equivalently its representation $\{t_0, \{(T_n, k_n) : n \in \mathbb{Z}_+\}\}$ has the properties that $P(t_0 = 0) = 1$ and $\{(T_n, k_n) : n \in \mathbb{Z}_+\}$ forms a stationary sequence of RVs.

If $\{(T_n, k_n) : n \in \mathbb{Z}_+\}$ is also ergodic, then Ψ is said to be a point-stationary and ergodic MPP.

For simplicity, we will always assume that a point-stationary MPP is ergodic.

In practical terms, ergodicity means that, for any measurable $f : \mathbb{M} \longrightarrow \mathbb{R}_+$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} f(\Psi_{(j)}) = E(f(\Psi)) ,\quad \text{with probability 1 (wp1)} . \tag{8.4}$$

(This is *Birkoff's ergodic theorem in its ergodic form*.) For example, if $f(\psi) = T_0$, then $f(\psi_{(j)}) = T_j$ and (8.4) yields the strong law of large numbers for the stationary ergodic sequence $\{T_n\}$; $\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} T_j = E(T_0)$, wp1. (The non-ergodic case is discussed in Sect. 8.8.)

Inherent in the definition of point-stationarity is the fact that there is a one-to-one correspondence between point-stationary point processes and stationary sequences of non-negative RVs; given any such stationary sequence $\{T_n\}$, $t_n \stackrel{def}{=} T_0 + \cdots + T_{n-1}$ (and $t_0 \stackrel{def}{=} 0$) defines a point-stationary point process.

When Ψ is point-stationary, we let T denote a generic interarrival time, define the *arrival rate* $\lambda = [E(T)]^{-1}$, and let $F(x) = P(T \le x)$, $x \ge 0$ denote the stationary interarrival time distribution with $\bar{F}(x) = 1 - F(x)$ being its tail. As in the classic *elementary renewal theorem*, it holds that $N(t)/t \to \lambda$ as $t \to \infty$, wp1.

From Kolmogorov's extension theorem in probability theory, a stationary sequence can be extended to be two-sided, $\{(T_n, k_n) : -\infty < n < \infty\}$, yielding a point-stationary MPP on all of $\boldsymbol{R}$:

$$\cdots t_{-2} < t_{-1} < t_0 = 0 < t_1 < t_2 < \cdots ,$$

where $t_{-n} \stackrel{\text{def}}{=} -(T_{-1} + \cdots + T_{-n})$, $n \ge 1$.

Point-stationary MPPs arise naturally as limits (in distribution) of $\Psi_{(j)}$ as $j \to \infty$. In applications the limit can be taken in a *Cesàro* sense. Independently take a discrete RV J with a uniform distribution on $\{1, \ldots, n\}$, and define an MPP Ψ^0 by defining its distribution as

$$\begin{aligned} \boldsymbol{P}^0(\cdot) = P(\Psi^0 \in \cdot) &\stackrel{\text{def}}{=} \lim_{n\to\infty} P(\Psi_{(J)} \in \cdot) \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} P(\Psi_{(j)} \in \cdot) . \end{aligned} \tag{8.5}$$

If the limit holds for all Borel sets of $\mathbb{M}$, then it can be shown that it holds uniformly over all Borel sets; known as *Cesàro total variation convergence*. Assuming the existence of such a limiting distribution $\boldsymbol{P}^0$, it is unique and is called the *point-stationary distribution* of Ψ (or of $\boldsymbol{P}$) and Ψ is said to be *point asymptotically stationary*. Any MPP $\Psi^0 = \{(t_n^0, k_n^0)\}$ distributed as $\boldsymbol{P}^0$ is called a *point-stationary version* of Ψ. Intuitively this is obtained from Ψ by randomly selecting a point t_j so far in the infinite future that shifting further to the next point t_{j+1} does not change the distribution; it is stationary with respect to such point shifts.

It is important to remember that a point-stationary MPP has (wp1) a point at the origin.

Time Stationarity

Definition 8.2
Ψ is called *time-stationary* if Ψ_s has the same distribution as Ψ, for all $s \ge 0$. In this case $P(t_0 > 0) = 1$ and $\{N_K(t) : t \ge 0\}$ has stationary increments for each mark set K.

When Ψ is time-stationary, the interevent time sequence $\{(T_n, k_n)\}$ will not be stationary in general; in particular, the distribution of T_j will generally be different for different choices of j. However, the stochastic process $\{A(t)\}$ is a stationary process.

Ergodicity is defined as requiring that the measure-preserving flow of shifts $\theta_s : \mathbb{M}$ to $\mathbb{M}$, $s \ge 0$, $\theta_s \psi = \psi_s$ be ergodic under the distribution of Ψ. (In the point-stationary case, ergodicity is equivalent to requiring that the measure-preserving shift map $\theta_{(1)} = \theta_{t_1}$ be ergodic.)

For simplicity, we will always assume that a time-stationary MPP is ergodic. In practical terms, ergodicity means that, for any measurable $f : \mathbb{M} \longrightarrow \mathbb{R}_+$ (satisfying $\int_0^t f(\Psi_s)\,\mathrm{d}s < \infty,\ t \geq 0,\ \text{wp1}$),

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t f(\Psi_s)\,\mathrm{d}s = E\big[f(\psi)\big],\ \text{wp1}\,. \tag{8.6}$$

When Ψ is time-stationary, the arrival rate is defined by $\lambda \stackrel{\text{def}}{=} E[N(1)]$ and it holds that $E[N(t)] = \lambda t,\ t \geq 0$. It also holds that $N(t)/t \to \lambda$ as $t \to \infty$, wp1.

Time-stationary MPPs can be extended to be two-sided

$$\cdots t_{-2} < t_{-1} < t_0 < 0 < t_1 < t_2 < \cdots\ , \tag{8.7}$$

where $P(t_0 < 0,\ t_1 > 0) = 1$. In this case $\{B(t)\}$ and $\{S(t)\}$ are stationary processes in which case $B(0) = |t_0|$ and $A(0) = t_1$ are identically distributed.

Time-stationary MPPs arise naturally as limits (in distribution) of Ψ_t as time $t \to \infty$. In applications the limit can be taken in a Cesàro sense: independently take a continuous RV, U, uniformly distributed over $(0, t)$, and define an MPP Ψ^* by defining its distribution as

$$\begin{aligned} \boldsymbol{P}^*(\cdot) = P(\Psi^* \in \cdot) &\stackrel{\text{def}}{=} \lim_{t\to\infty} P(\Psi_U \in \cdot) \\ &= \lim_{t\to\infty} \frac{1}{t} \int_0^t P(\Psi_s \in \cdot)\,\mathrm{d}s\,. \end{aligned} \tag{8.8}$$

If the limit holds for all Borel sets of $\boldsymbol{M}$, then it can be shown that it holds uniformly over all Borel sets; Cesàro total variation convergence. Assuming the existence of such a limiting distribution $\boldsymbol{P}^*$, it is unique and is called the *time-stationary distribution* of Ψ (or of $\boldsymbol{P}$) and Ψ is said to be *time asymptotically stationary*. Any MPP $\Psi^* = \{(t_n^*, k_n^*)\}$ distributed as $\boldsymbol{P}^*$ is called a *time-stationary version* of Ψ. Intuitively it is obtained from Ψ by randomly selecting a time t as the origin that is so far in the infinite future that shifting s time units further does not change the distribution; it is stationary with respect to such time shifts.

It is important to remember that a time-stationary MPP has (wp1) no point at the origin.

8.1.7 The Relationship Between Ψ, Ψ^0 and Ψ^*

Suppose that Ψ has a point-stationary version Ψ^0. What then is the time-stationary distribution of Ψ^0? Intuitively it should be the same as the time-stationary distribution of Ψ, and this turns out to be so:

Proposition 8.1

Ψ is point asymptotically stationary (defined as in (8.5)) with point-stationary (and ergodic) $\boldsymbol{P}^0$ under which $0 < \lambda < \infty$, if and only if Ψ is time asymptotically stationary (defined as in (8.8)) with time-stationary (and ergodic) $\boldsymbol{P}^*$ under which $0 < \lambda < \infty$. In this case $\boldsymbol{P}^*$ is the time-stationary distribution of $\boldsymbol{P}^0$, and $\boldsymbol{P}^0$ is the point-stationary distribution of $\boldsymbol{P}^*$. (All three of $\Psi,\ \Psi^0,\ \Psi^*$ share the same point- and time-stationary distributions.)

Because of the above proposition, Ψ is called *asymptotically stationary* if one (hence both) of $\boldsymbol{P}^0$, $\boldsymbol{P}^*$ exist with $0 < \lambda < \infty$.

Proposition 8.2

Suppose that Ψ is asymptotically stationary (and ergodic). Then the two definitions of the arrival rate λ coincide; $\lambda = E[N^*(1)] = [E(T^0)]^{-1}$. Moreover, the ergodic limits in (8.4) and (8.6) hold for all three MPPs, Ψ, Ψ^0, Ψ^* with the right-hand sides replaced by $E[f(\Psi^0)]$ and $E[f(\Psi^*)]$ respectively.

It turns out that, in fact, all three MPPs, Ψ, Ψ^0, Ψ^* *shift couple*, and that is the key to understanding the above two propositions ($\stackrel{\mathrm{d}}{\sim}$ denotes "is distributed as"):

Proposition 8.3

If Ψ is asymptotically stationary, then there exist versions of $\Psi \stackrel{\mathrm{d}}{\sim} \boldsymbol{P}$, $\Psi^0 \stackrel{\mathrm{d}}{\sim} \boldsymbol{P}^0$, $\Psi^* \stackrel{\mathrm{d}}{\sim} \boldsymbol{P}^*$ all on a common probability space together with three random times, S_1, S_2, S_2 such that $\Psi_{S_1} = \Psi^0_{S_2} = \Psi^*_{S_3}$. In other words, they share the same sample paths modulo some time shifts.

Given an asymptotically stationary MPP Ψ, the superscripts ^{0} and * are used to denote point- and time-stationary versions of all associated processes of Ψ. $\Psi^0 = \{(t_n^0, k_n^0)\}$, and $\Psi^* = \{(t_n^*, k_n^*)\}$ denote the two versions, and, for example, $\{(T_n^0, k_n^0)\}$ denotes the stationary sequence of interevent times and marks for Ψ^0, and T^0 denotes such a generic interevent time with F being its distribution; $F(x) = P(T^0 \leq x),\ x \geq 0$. $\{A^*(t)\}$ denotes the forward recurrence time process for Ψ^*, etc.

To illustrate the consequences of Proposition 8.8.2, suppose that $f(\psi) = t_0$. Then $f(\psi_s) = t_0(s) = A(s)$, forward recurrence time, and it holds that

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t A^*(s)\,\mathrm{d}s = E(t_0^*),\ \text{wp1}\,,$$

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t A(s)\,\mathrm{d}s = E(t_0^*),\ \text{wp1}\,,$$

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t A^0(s)\,\mathrm{d}s = E(t_0^*), \ \text{wp1} .$$

8.1.8 Examples

Some simple examples are presented. In some of these examples, marks are left out for simplicity and to illustrate the ideas of stationarity better.

1. *Poisson process:* A (time-homogenous) Poisson process with rate λ has independent and identically distributed (iid) interarrival times $T_n, \ n \geq 0$ with an exponential distribution, $P(T \leq x) = 1 - \mathrm{e}^{-\lambda x}, \ x \geq 0$. Its famous defining feature is that $\{N(t)\}$ has both stationary and independent increments, and that these increments have a Poisson distribution; $N(t)$ is Poisson-distributed with mean $E[N(t)] = \lambda t, \ t \geq 0$; $P[N(t) = n] = \left[\mathrm{e}^{-\lambda t}(\lambda t)^n\right]/n!, n \in \mathbb{Z}_+$.
   If we place t_0 at the origin, $t_0 = 0$, then the Poisson process is point-stationary, whereas if we (independently) choose t_0 distributed as exponential at rate λ, then the Poisson process becomes time-stationary. Thus, for a Poisson process, removing the point at the origin from Ψ^0 yields Ψ^*, while placing a point at the origin for Ψ^* yields Ψ^0. Observe that, by the memoryless property of the exponential distribution, $A(t)$ is distributed as exponential with rate λ for all $t \geq 0$.
   A two-sided time-stationary version is obtained as follows: Choose both $|t_0^*| = B^*(0)$ and $t_1^* = A^*(0)$ as iid with an exponential λ distribution. All interarrival times $T_n^*, -\infty < n < \infty$ are iid exponential at rate λ except for $T_0^* = t_1^* - t_0^* = B^*(0) + A^*(0) = S^*(0)$, the spread, which has an Erlang distribution (mean $2/\lambda$). That the distribution of T_0^* is different (larger) than T results from the *inspection paradox*: Randomly choosing the origin in time, we are more likely to land in a larger than usual interarrival time because larger intervals cover a larger proportion of the time line. $S^*(t)$ is distributed as Erlang (mean $2/\lambda$) for all $t \in \mathbb{R}$, by stationarity.
   The Poisson process is the unique simple point process with a counting process that possesses both stationary and independent increments.
2. *Renewal process*: Interarrival times $\{T_n : n \geq 0\}$, are iid with a general distribution $F(x) = P(T \leq x)$ and mean $\lambda^{-1} = E(T)$. If $t_0 = 0$ then the renewal process is point-stationary, and is called a non-delayed version of the renewal process. If instead, independently, $t_0 = A(0) > 0$ and has the *stationary excess* distribution, F_e, defined by
$$F_\mathrm{e}(x) = \lambda \int_0^x \bar{F}(y)\,\mathrm{d}y, \ x \geq 0 , \tag{8.9}$$
   then the renewal process is time-stationary and $A^*(t)$ is distributed as F_e for all $t \geq 0$. (In the Poisson process case $F_\mathrm{e} = F$.) In general, when $t_0 > 0$ the renewal process is said to be *delayed*. For any renewal process (delayed or not) $\Psi_{(j)}$ always yields a point-stationary version Ψ^0 (for any $j \geq 0$), while Ψ_s always yields a delayed version with delay $t_0(s) = A(s)$. Only when this delay is distributed as F_e is the version time-stationary. As $s \to \infty$, the distribution of $A(s)$ converges (in a Cesàro total variation sense) to F_e; this explains why the distribution of Ψ_s converges (in a Cesàro total variation sense) to the time-stationary version we just described.
   A two-sided time-stationary version Ψ^* is obtained when $T_n^*, \ n \neq 0$ are iid distributed as F, and independently $[B^*(0), A^*(0)] = (|t_0^*|, t_1^*)$ has the joint distribution $P(|t_0^*| > x, t_1^* > y) = \bar{F}_\mathrm{e}(x+y), \ x \geq 0, \ y \geq 0$. Here, as for the Poisson process, $T_0^* = S^*(0)$ has, due to the inspection paradox, a distribution that is stochastically larger than F, $P(T_0^* > x) \geq P(T > x), \ x \geq 0$; this is called the *spread* distribution of F and has tail
$$P(T_0^* > x) = \lambda x \bar{F}(x) + \bar{F}_\mathrm{e}(x) ; \tag{8.10}$$
   while $E(T_0^*) = E(T^2)/E(T)$. If F has a density $f(x)$, then the spread has a density $\lambda x f(x)$, which expresses the length biasing contained in the spread. F_e always has a density, $f_\mathrm{e}(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_\mathrm{e}(x) = \lambda \bar{F}(x)$, whether or not F does.
3. *Compound renewal process*: Given the counting process $\{N(t)\}$ for a renewal process, and independently an iid sequence of RVs $\{X_n\}$ (called the jumps), with jump distribution $G(x) = P(X \leq x), \ x \in \mathbb{R}$, the process
$$X(t) = \sum_{j=1}^{N(t)} X_j, \ t \geq 0$$
   is called a compound renewal process with jump distribution G. A widely used special case is when the renewal process is a Poisson process, called a *compound Poisson process*.
   This can elegantly be modeled as the MPP $\Psi = \{(t_n, k_n)\}$, where $\{t_n\}$ are the points and $k_n = X_n$. Because it is assumed that $\{X_n\}$ is

independent of $\{t_n\}$, obtaining point and time-stationary versions merely amounts to joining in the iid marks to Example 2's renewal constructions: $k_n^0 = X_n = k_n^*$.

4. *Renewal process with marks depending on interarrival times*: Consider a two-sided renewal process and define the marks as $k_n = T_{n-1}$, the length of the preceding interarrival time. The interesting case is to construct a time-stationary version. This can be done by using the two-sided time-stationary version of the point process, $\{t_n^*\}$, from Example 2. Note that, for $n \neq 1$, the k_n^* are iid distributed as F, defined by $k_n^* = T_{n-1}^*$; only k_1^* is different (biased via the inspection paradox). $k_1^* = T_0^*$ and has the spread distribution.
5. *Cyclic deterministic*: Starting with interarrival time sequence $\{T_n\} = \{1, 2, 3, 1, 2, 3, 1, 2, 3\ldots\}$, Ψ^0 is given by defining $t_0^0 = 0$ and
$$\{T_n^0 : n \geq 0\} = \begin{cases} \{1, 2, 3, 1, 2, 3, \ldots\}, \text{wp} = 1/3\,; \\ \{2, 3, 1, 2, 3, 1, \ldots\}, \text{wp} = 1/3\,; \\ \{3, 1, 2, 3, 1, 2, \ldots\}, \text{wp} = 1/3\,. \end{cases} \tag{8.11}$$
(By randomly selecting a j and choosing t_j as the origin, we are equally likely to select a T_j with length 1, 2, or 3; $P(T^0 = i) = 1/3,\ \ i = 1, 2, 3$.) The two-sided extension is given by defining $t_0^0 = 0$ and
$$\{\ldots, T_{-1}^0, T_0^0, T_1^0, \ldots\} = \begin{cases} \{\ldots, 3, 1, 2, \ldots\}, \text{wp} = 1/3; \\ \{\ldots, 1, 2, 3, \ldots\}, \text{wp} = 1/3\,; \\ \{\ldots, 2, 3, 1, \ldots\}, \text{wp} = 1/3\,. \end{cases}$$
A construction of Ψ^* is given as follows. Let U denote a random variable having a continuous uniform distribution over (0, 1). Then
$$\{t_0^*,\ \{T_n^* : n \geq 0\}\} = \begin{cases} U,\ \{2, 3, 1, 2, 3, 1\ldots\}, \text{wp} = 1/6\,; \\ 2U,\ \{3, 1, 2, 3, 1, 2\ldots\}, \text{wp} = 1/3\,; \\ 3U,\ \{1, 2, 3, 1, 2, 3\ldots\}, \text{wp} = 1/2\,. \end{cases} \tag{8.12}$$
By randomly selecting a time s as the origin, we would land inside an interarrival time of length 1, 2, or 3 with probability $1/6$, $1/3$ and $1/2$ respectively (they are proportions of time). Given that we land inside one of length i, $t_0(s)$ would be distributed as $iU,\ \ i = 1, 2, 3$ (e.g., uniform on $(0, i)$). Unlike $\{T_n^0 : n \geq 0\}$, $\{T_n^* : n \geq 0\}$ is not a stationary sequence because of the unequal probabilities in the mixture.
This illustrates the general fact that t_0^* has the stationary excess distribution $F_e(x)$ of the point-stationary distribution $F(x) = P(T^0 \leq x)$ [recall (8.9)]. In a two-sided extension, the distribution of $T_0^* = |t_0^*| + t_1^* = S^*(0)$ is the spread distribution of F; in this case $P(T_0^* = i) = i/6,\ \ i = 1, 2, 3$, and the joint distribution of $(|t_0^*|, t_1^*)$ is of the mixture form $(1-U, U)$, $(2-2U, 2U)$, $(3-3U, 3U)$ with probabilities $1/6,\ \ 1/3,\ \ 1/2$ respectively.
This example also illustrates the general fact that the time reversal of an MPP Ψ has a different distribution from Ψ; the sequence $\{T_n^0 : n \geq 0\}$ has a different distribution from that of the sequence $\{T_n^0 : n \leq 0\}$.
6. *Single-server queue*: t_n denotes the arrival time of the n-th customer, denoted by C_n, to a system (such as a bank with one clerk) that has one server behind which customers wait in queue (line) in a first-in-first-out manner (FIFO). Upon entering service, C_n spends an amount of time S_n with the server and then departs. D_n denotes the length of time that C_n waits in line before entering service and is called the *delay* of C_n in queue. Thus C_n enters service at time $t_n + D_n$ and departs at time $t_n + D_n + S_n$; $W_n = D_n + S_n$ is called the *sojourn time*. The total number of customers in the system at time t, is denoted by $L(t)$ and can be constructed from $\{W_n\}$;
$$L(t) = \sum_{j=1}^{N(t)} I(W_j > t - t_j), \tag{8.13}$$
because C_j is in the system at time t if $t_j \leq t$ and $W_j > t - t_j$.
Letting $\Psi = [(t_n, S_n)]$ yields an MPP, with marks $k_n = S_n$, called the *input* to the queueing model; from it the queueing processes of interest can be constructed. It is known that D_n satisfies the recursion $D_{n+1} = (D_n + S_n - T_n)_+,\quad n \geq 0$, where $x_+ \overset{\text{def}}{=} \max(x, 0)$ denotes the positive part of x, and yet another MPP of interest is $\Psi = \{[t_n, (S_n, D_n)]\}$, where now $k_n = (S_n, D_n)$. Letting $\boldsymbol{D}_{(n)} = (D_{n+m} : m \geq 0)$, another important MPP with an infinite-dimensional mark space is $\Psi = \{[t_n, (S_n, \boldsymbol{D}_{(n)})]\}$, where $k_n = (S_n, \boldsymbol{D}_{(n)})$. The *workload* $V(t)$ is defined by $V(t) = D_n + S_n - (t - t_n)$, $t \in [t_n, t_{n+1})$, $n \geq 0$, and $D_n = V(t_n-)$; it rep-

resents the sum of all remaining service times in the system at time t. It can also model the water level of a reservoir into which the amounts S_n are inserted at the times t_n while water is continuously drained out at rate 1.

A point-stationary version $\Psi^0 = \{[t_n^0, (S_n^0, D_n^0)]\}$ yields a stationary version of the delay sequence $\{D_n^0\}$ with stationary delay distribution $P(D \le x) = P(D_0^0 \le x)$, which is an important measure of congestion from the point of view of customers, as is its mean, $d \overset{\text{def}}{=} E(D)$, the average delay.

A time-stationary version $\Psi^* = \{[t_n^*, (S_n^*, D_n^*)]\}$ yields a time-stationary version of workload $\{V^*(t)\}$ and corresponding stationary distribution $P(V \le x) = P(V^*(0) \le x)$, which is an important measure of congestion from the point of view of the system, as is its mean, $E(V)$, is the average workload.

If the input MPP is asymptotically stationary (ergodic) with $0 < \lambda E(S^0) < 1$, then it is known that $\Psi = \{[t_n, (S_n, D_n)]\}$ is asymptotically stationary, e.g., the stationary versions and distributions for such things as delay and workload exist.

8.2 Inversion Formulas

Inversion formulas allow one to derive $\boldsymbol{P}^0$ from $\boldsymbol{P}^*$, and visa versa.

Theorem 8.2 (Inversion formulas)
Suppose that Ψ is asymptotically stationary (and ergodic) and $0 < \lambda < \infty$. Then

$$P(\Psi^* \in \cdot) = \lambda E\left[\int_0^{T_0^0} I(\Psi_s^0 \in \cdot)\,\mathrm{d}s\right], \tag{8.14}$$

$$P(\Psi^0 \in \cdot) = \lambda^{-1} E\left[\sum_{j=0}^{N^*(1)} I(\Psi_{(j)}^* \in \cdot)\right], \tag{8.15}$$

which, in functional form, become

$$E(f(\Psi^*)) = \lambda E\left[\int_0^{T_0^0} f(\Psi_s^0)\,\mathrm{d}s\right], \tag{8.16}$$

$$E(f(\Psi^0)) = \lambda^{-1} E\left[\sum_{j=0}^{N^*(1)} f(\Psi_{(j)}^*)\right]. \tag{8.17}$$

Recalling (8.6) and Proposition 8.8.2, it is apparent that (8.14) and (8.16) are generalizations (to a stationary ergodic setting) of the renewal reward theorem from renewal theory:

> *The time average equals the expected value over a cycle divided by the expected cycle length.*

Here a cycle length is (by point stationarity) represented by any interarrival time, so the first one, $T_0^0 = t_1^0$, is chosen for simplicity. Equations (8.15) and (8.17) are the inverse [recalling (8.4)]:

> *The point average equals the expected value over a unit of time divided by the expected number of points during a unit of time.*

Here a unit of time is (by time stationarity) represented by any such unit, so the first one, (0, 1], is chosen for simplicity.

8.2.1 Examples

The following examples illustrate how some well-known results that hold for renewal processes, involving the stationary excess distribution (8.9) and the inspection paradox and spread distribution (8.10) also hold in general. Throughout, assume that Ψ is asymptotically stationary (and ergodic).

1. *Stationary forward recurrence time*: $P(t_0^* \le x) = P[A^*(t) \le x] = F_e(x)$ where $F(x) = P(T^0 \le x)$. This is derived by applying (8.17) with $f(\psi) = I(t_0 > x)$: $f(\psi_s^0) = I[t_0^0(s) > x]$ and $t_0^0(s) = A^0(s) = t_1^0 - s$, $s \in [0, t_1^0)$; $\int_0^{T_0^0} f(\Psi_s^0)\,\mathrm{d}s = \int_0^{T_0^0} I\{s < T_0^0 - x\}\,\mathrm{d}s = (T_0^0 - x)_+$. $\lambda E[(T_0^0 - x)_+] = \lambda \int_x^\infty \bar{F}(y)\,\mathrm{d}y = \bar{F}_e(x)$.
2. *Stationary backwards recurrence time*: $P[B(0)^* \le x] = F_e(x)$. Here, a two-sided framework must be assumed so that $B(0) = |t_0|$. Applying (8.17) with $f(\psi) = I[B(0) > x]$: $f(\Psi_s^0) = I[B^0(s) > x]$ where $B^0(s) = s$, $s \in [0, t_1^0)$; $\int_0^{T_0^0} f(\Psi_s^0)\,\mathrm{d}s = \int_0^{T_0^0} I(s > x)\,\mathrm{d}s = (T_0^0 - x)_+$. $\lambda E[(T_0^0 - x)_+] = \bar{F}_e(x)$.
3. *Stationary spread*: $P(T_0^* > x) = \lambda x \bar{F}(x) + \bar{F}_e(x)$. Here again, a two-sided framework must be assumed so that $S(0) = |t_0| + t_1$. Applying (8.17)

with $f(\psi) = I(T_0 > x)$: $f(\psi_s) = I[S(s) > x]$ and $S^0(s) = T_0^0, s \in [0, t_1^0)$; $\int_0^{T_0^0} f(\Psi_s^0)\,ds = \int_0^{T_0^0} I(T_0^0 > x)\,ds = T_0^0 I(T_0^0 > x)$. $\lambda E(T_0^0 I(T_0^0 > x)) = \lambda x \times \bar{F}(x) + \bar{F}_e(x)$ by carrying out the integration $E[T_0^0 I(T_0^0 > x)] = \int_0^\infty P(T_0^0 > y, T_0^0 > x)\,dy$.

8.2.2 The Canonical Framework

In the canonical framework $\boldsymbol{E}$ denotes expectation under $\boldsymbol{P}$, $\boldsymbol{E^0}$ denotes expectation under $\boldsymbol{P}^0$ and $\boldsymbol{E}^*$ denotes expectation under $\boldsymbol{P}^*$ and $\Psi : \mathbb{M} \to \mathbb{M}$ is the identity map; $\Psi(\psi) = \psi$. This makes for some elegance and simplicity in notation. For example, the inversion formulas in functional form become

$$\boldsymbol{E}^*[f(\Psi)] = \lambda \boldsymbol{E^0}\left[\int_0^{T_0} f(\Psi_s)\,ds\right],$$

$$\boldsymbol{E^0}[f(\Psi)] = \lambda^{-1} \boldsymbol{E}^*\left[\sum_{j=0}^{N(1)} f(\Psi_{(j)})\right]. \tag{8.18}$$

8.3 Campbell's Theorem for Stationary MPPs

Suppose that $\Psi = \Psi^*$ is time-stationary (and ergodic), with point-stationary version Ψ^0. From the inversion formula (8.15), $P(k^0 \in K) = \lambda^{-1} E\{\Psi^*[(0, 1] \times K]\}$, yielding $E\{\Psi^*[(0, 1] \times K]\} = \lambda P(k^0 \in K)$. This implies that the intensity measure from Campbell's theorem becomes $\nu(A \times K) = E[\Psi^*(A \times K)] = \lambda l(A) P(k^0 \in K)$, where $l(A)$ denotes Lebesgue measure {e.g., $E[\Psi^*(dt \times dk)] = \lambda\, dt P(k^0 \in dk)$}. This can be rewritten as $\nu(A \times K) = \lambda l(A) E[I(k_0^0 \in K)]$, in terms of the mark at the origin k_0^0 of Ψ^0. This yields

Theorem 8.3 [Campbell's theorem under stationarity (and ergodicity)]
For any non-negative measurable function $g = g(t, k)$,

$$E[\Psi^*(g)] = \lambda E\left[\int_{\mathbb{R}} g\left(t, k_0^0\right) dt\right].$$

8.3.1 Little's Law

A classic application of Campbell's theorem in queueing theory is when $\Psi^* = [(t_n^*, W_n^*)]$ (two-sided) represents a time-stationary queueing model, where t_n^* is the arrival time of the n-th customer, and W_n^* their sojourn time. Using $g(t, w) = 0,\ t > 0$ and $g(t, w) = I(w > |t|)$, $t \le 0$ yields $\Psi^*(g) = \sum_{j \le 0} I(W_j^* > |t_j^*|) = L^*(0)$, denoting the time-stationary number of customers in the system at time $t = 0$ [recall (8.13)]. Campbell's theorem then yields $E[L^*(0)] = \lambda E(W^0)$, known as *Little's Law* or $L = \lambda w$.

8.3.2 The Palm–Khintchine Formula

Another application of interest for Campbell's theorem is the *Palm–Khintchine formula*: for all $n \ge 0$ and $t > 0$,

$$P[N^*(t) > n] = \lambda \int_0^t P[N^0(s) = n]\,ds. \tag{8.19}$$

Proof: Since this result does not involve any marks, the marks can be replaced by new ones: define $k_j = \psi_{(j)}$. With these new marks Ψ^* remains stationary (and ergodic). For fixed $t > 0$ and $n \ge 0$, define $g(s, \psi) = I[0 \le s \le t,\ N(t - s) = n]$. Then

$$\begin{aligned}\Psi^*(g) &= \sum_{j=1}^{N^*(t)} I\left(N^*(t_j, t] = n\right)\\ &= I[N^*(t) > n],\end{aligned}$$

where the last equality is obtained by observing that $N(t) > n$ if and only if there exists a j (unique) such that $t_j < t$ and there are exactly n more arrivals during $(t_j, t]$. Campbell's theorem then yields

$$\begin{aligned}P[N^*(t) > n] &= \lambda E \int_0^t I[N^0(t - s) = n]\,ds\\ &= \lambda \int_0^t P[N^0(t - s) = n]\,ds,\\ &= \lambda \int_0^t P[N^0(s) = n]\,ds. \qquad \blacksquare\end{aligned}$$

8.4 The Palm Distribution: Conditioning in a Point at the Origin

Given any time-stationary MPP Ψ, its *Palm* distribution (named after C. Palm) is defined by

$$Q(\cdot) = \lambda^{-1} E\left[\sum_{j=0}^{N(1)} I(\Psi_{(j)} \in \cdot)\right] ,$$

and the mapping taking $P(\Psi \in \cdot)$ to $Q(\cdot)$ is called the *Palm transformation*. From (8.15), it follows that, if Ψ is also ergodic, then Q is the same as the point-stationary distribution P^0 [as defined in (8.5)]. If ergodicity does not hold, however, then Q and P^0 are different (in general), but the Palm distribution still yields a point-stationary distribution and any version distributed as Q is called a *Palm version* of Ψ.

Similarly, if we start with any point-stationary MPP Ψ, we can define a time-stationary distribution by

$$H(\cdot) = \lambda E\left[\int_0^{T_0} I(\Psi_s \in \cdot)\,\mathrm{d}s\right] ,$$

which under ergodicity agrees with P^*, but otherwise does not (in general). This mapping is called the *Palm inverse transformation* because applying it to Q yields back the original time-stationary distribution $P(\Psi \in \cdot)$. Together the two formulas are called the *Palm inversion formulas*. It should be emphasized that only in the non-ergodic case does the distinction between Q and P^0 (or H and P^*) become an issue because only when ergodicity holds can Q be interpreted as a point average [as defined in (8.5)], so one might ask if there is some other intuitive way to interpret Q. The answer is yes: if Ψ is time-stationary, then its Palm distribution Q can be interpreted as *the conditional distribution of* Ψ *given a point at the origin*:

Theorem 8.4

If Ψ is time-stationary, then the Palm distribution Q can be obtained as the limiting distribution

$$Q(\cdot) = \lim_{t\to 0} P(\Psi \in \cdot \mid t_0 \le t) ,$$

in the sense of weak convergence. Total variation convergence is obtained if Ψ is first shifted to t_0:

$$Q(\cdot) = \lim_{t\to 0} P(\Psi_{(0)} \in \cdot \mid t_0 \le t) ,$$

in total variation.

As an immediate consequence, we conclude that (under ergodicity)

$$P(\Psi^0 \in \cdot) = \lim_{t\to 0} P(\Psi^* \in \cdot \mid t_0^* \le t)$$

(weak convergence) ,

$$P(\Psi^0 \in \cdot) = \lim_{t\to 0} P(\Psi^*_{(0)} \in \cdot \mid t_0 \le t)$$

(total variation convergence) .

Under ergodicity P^0 can be viewed as the conditional distribution of P^* given a point at the origin.

A proof of such results can be carried out using inversion formulas and Khintchine–Korolyuk's Theorem 8.8.1 given in the next section which asserts that $P[N^*(t) > 0] \approx \lambda t$ as $t \to 0$.

Putting the one-sided renewal process aside, it is not true in general that $\Psi^*_{(0)}$ has a point-stationary distribution: shifting a time-stationary MPP to its initial point does not in general make it point-stationary; conditioning on $\{t_0^* \le t\}$ and taking the limit as $t \to 0$ is needed. [Recall the cyclic deterministic example in (8.12), for example.]

8.5 The Theorems of Khintchine, Korolyuk, and Dobrushin

For a Poisson process with rate λ, $P[N(t) = n] = \frac{\mathrm{e}^{-\lambda t}(\lambda t)^n}{n!}$, $n \in \mathbb{Z}_+$; thus $P[N(t) > 0] = 1 - \mathrm{e}^{-\lambda t}$ yielding (by L'Hospital's rule for example)

$$\lim_{t\to 0} \frac{P[N(t) > 0]}{t} = \lambda . \tag{8.20}$$

Similarly, $P[N(t) > 1] = 1 - \mathrm{e}^{-\lambda t}(1 + \lambda t)$ yielding

$$\lim_{t\to 0} \frac{P[N(t) > 1]}{t} = 0 . \tag{8.21}$$

Both (8.20) and (8.21) remain valid for any simple time-stationary point process, and the results are attributed to A. Y. Khintchine, V. S. Korolyuk, and R. L. Dobrushin. Any point process satisfying (8.21) is said to be *orderly*.

Theorem 8.5 (Khintchine–Korolyuk)
If Ψ is time stationary (and simple), then (8.20) holds.

Theorem 8.6 (Dobrushin)
If Ψ is time stationary (and simple), then (8.21) holds.

Proofs can easily be established using inversion formulas. For example, assume ergodicity and let $\Psi^* = \Psi$ with Ψ^0 being a point-stationary version with $F(x) = P(T^0 \le x)$ and $F_e(x) = \lambda \int_0^x [1 - F(y)]\,dy$. Then $P[N^*(t) > 0] = P(t_0^* \le t) = F_e(t)$, from the inversion formula (8.14). L'Hospital's rule then reduces the limit in (8.20) to $\lim_{t\to 0} \lambda[1 - F(t)] = \lambda$ $[F(0) = 0$ by simplicity]. Equation (8.21) can be proved from the Palm–Khintchine formula (8.19) for $n = 1$:

$$
\begin{aligned}
P[N^*(t) > 1] &= \lambda \int_0^t P[N^0(s) = 1]\,ds \\
&= \lambda \int_0^t P\big(t_1^0 \le s, t_2^0 > s\big)\,ds \\
&= \lambda \int_0^t P\big(t_1^0 \le s, t_2^0 > s\big)\,ds \\
&\le \lambda \int_0^t P\big(t_1^0 \le s\big)\,ds \le \lambda t F(t)\,;
\end{aligned}
$$

the result then follows since $F(0) = 0$ by simplicity.

8.6 An MPP Jointly with a Stochastic Process

In many applications an MPP Ψ is part of or interacts with some stochastic process $\boldsymbol{X} = [X(t) : t \ge 0]$, forming a joint process $(\boldsymbol{X}, \Psi)$. For example, Ψ might be the arrival times and service times to a queueing model, and $X(t)$ the state of the queue at time t. To accommodate this it is standard to assume that the sample paths of $\boldsymbol{X}$ are functions $x : \boldsymbol{R}_+ \to \boldsymbol{S}$ in the space

$$
\begin{aligned}
\mathcal{D}_{\boldsymbol{S}}[0,\infty) \stackrel{\text{def}}{=} \{x :\ & x \text{ is continuous from the right} \\
& \text{and has left-hand limits}\}\,,
\end{aligned}
$$

endowed with the *Skorohod* topology. The state-space $\boldsymbol{S}$ can be a general complete separable metric space, but in many applications $\boldsymbol{S} = \boldsymbol{R}$, or a higher-dimensional Euclidean space. $\mathcal{D}_{\boldsymbol{S}}[0,\infty)$ is denoted by $\mathcal{D}$ for simplicity.

Continuous from the right means thats for each $t \ge 0$: $x(t+) \stackrel{\text{def}}{=} \lim_{h\downarrow 0} x(t+h) = x(t)$, while *has left-hand limits* means that for each $t > 0$: $x(t-) \stackrel{\text{def}}{=} \lim_{h\downarrow 0} x(t-h)$ exits (and is finite). Such functions are also called *cadlag* (continue à droit, limits à gauchee) from the French.

It can be shown that such a function has, at most, countably many discontinuities, and is bounded on any finite interval $[a, b]$: $\sup_{t\in[a,b]} |x(t)| < \infty$. If t is a discontinuity, then the *jump* of X at t is defined by $x(t+) - x(t-)$.

Jointly the sample paths are pairs $(x, \psi) \in \mathcal{D} \times \boldsymbol{M}$ and this canonical space is endowed with the product topology and corresponding Borel sets. $(\boldsymbol{X}, \Psi) : \Omega \to \mathcal{D} \times \boldsymbol{M}$ formally is a mapping into the canonical space under some probability P; its distribution is denoted by $\boldsymbol{P}(\cdot) = P[(\boldsymbol{X}, \Psi) \in \cdot]$. The shifts θ_s and $\theta_{(j)}$ extend to this framework in a natural way by defining $\boldsymbol{X}_s = \theta_s \boldsymbol{X} = [X(s+t) : t \ge 0]$; $\theta_s(\boldsymbol{X}, \Psi) = (\boldsymbol{X}_s, \Psi_s)$. The notions of point and time stationarity (and ergodicity) go right through as does the notion of asymptotic stationarity, and the inversion formulas also go through. For example, the functional form of the inversion formulas in the canonical framework are:

$$
\boldsymbol{E}^0[f(\boldsymbol{X}, \Psi)] = \lambda^{-1} \boldsymbol{E}^* \left[\sum_{j=0}^{N(t)} f(\boldsymbol{X}_{(j)}, \Psi_{(j)}) \right], \tag{8.22}
$$

$$
\boldsymbol{E}^*[f(\boldsymbol{X}, \Psi)] = \lambda \boldsymbol{E}^0 \left[\int_0^{T_0} f(\boldsymbol{X}_s, \Psi_s)\,ds \right]. \tag{8.23}
$$

A point-stationary version is denoted by $(\boldsymbol{X}^0, \Psi^0)$, and has the property that $\boldsymbol{X}^0$ can be broken up into a stationary sequence of cycles $\mathcal{C}_n = [X^0(t_n^0 + t) : 0 \le t < T_n^0]$, $n \in \boldsymbol{Z}_+$, with cycle lengths being the interevent times $\{T_n\}$.

A time-stationary version is denoted by $(\boldsymbol{X}^*, \Psi^*)$, and $\boldsymbol{X}^*$ is a stationary stochastic process.

The two-sided framework goes through by letting $x : \boldsymbol{R} \to \boldsymbol{S}$ and using the extended space $\mathcal{D}(-\infty, +\infty)$.

8.6.1 Rate Conservation Law

Given a asymptotically stationary (and ergodic) pair $(\boldsymbol{X}, \Psi)$, with $\boldsymbol{X}$ real-valued, assume also that the sample paths of $\boldsymbol{X}$ are right differentiable, $x'(t) = \lim_{h \downarrow 0}[X(t+h) - x(t)]/x(t)$ exists for each t. Further assume that the points t_n of Ψ include all the discontinuity points (jumps) of $\boldsymbol{X}$ (if any); if for some t it holds that $X(t-) \neq X(t+)$, then $t = t_n$ for some n. Noting that (wp1) $\boldsymbol{E}^*[X'(0)] = \lim_{t\to\infty} \frac{1}{t}\int_0^t X'(s)\,\mathrm{d}s$ and $\boldsymbol{E}^0[X(0+) - X(0-)] = \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^n [X(t_j+) - X(t_j-)]$, average jump size, the following is known as *Miyazawa's rate conservation law* (RCL):

Theorem 8.7
If $\boldsymbol{E}^*|X'(0)| < \infty$ and $\boldsymbol{E}^0|X(0-) - X(0+)| < \infty$, then

$$\boldsymbol{E}^*(X'(0)) = \lambda \boldsymbol{E}^0[X(0-) - X(0+)]\ .$$

The time-average right derivative equals the arrival rate of jumps multiplied by the (negative of) the average jump size.

As an easy example, for $x \geq 0$ let $X(t) = [A(t) - x]_+$, where $A(t)$ is the forward recurrence time for Ψ. Then $A'(t) = -1$ and $X'(t) = -I[A(t) > x]$. Jumps are of the form $X(t_n+) - X(t_n-) = (T_n - x)_+$. The RCL then yields $P[A^*(0) > x] = \lambda E(T_0^0 - x)_+ = 1 - F_e(x)$. The RCL has many applications in queueing theory. For example consider Example 6 from Sect. 8.1.8 and let $X(t) = V^2(t)$. Then $V'(t) = -I[V(t) > 0]$ so $X'(t) = -2V(t)$ and $X(t_n+) - X(t_n-) = 2S_n D_n + S_n^2$; the RCL thus yields *Brumelle's formula*, $E(V) = \lambda E(SD) + \lambda E(S^2)/2$. (Here $SD = S_0^0 D_0^0$.) A sample-path version of the RCL can be found in [8.1].

8.7 The Conditional Intensity Approach

Motivated by the fact that $\{N(t) - \lambda t : t \geq 0\}$ forms a mean-zero martingale for a time-homogenous Poisson process with rate λ, the *conditional intensity* $\lambda(t)$ of a point process (when it exists) satisfies the property that $\{N(t) - \int_0^t \lambda(s)\,\mathrm{d}s\}$ forms a mean-zero martingale. The framework requires a history $\mathcal{F}_t$ supporting $N(t)$ and a heuristic definition is then $\lambda(t)\,\mathrm{d}t = E(N(\mathrm{d}t) \mid \mathcal{F}_t)$ which asserts that for each t the conditional expected number of new arrivals in the next $\mathrm{d}t$ time units, conditional on the history up to time t, is equal to $\lambda(t)\,\mathrm{d}t$. For a time-homogenous Poisson process at rate λ, $\lambda(t) = \lambda$; $E[N(\mathrm{d}t) \mid \mathcal{F}_t] = \lambda\,\mathrm{d}t$ due to stationary and independent increments; but for general point processes, $\lambda(t)$ (if it exists) depends on the past evolution (before time t).

A *non-stationary* Poisson process is a simple and very useful example, where the arrival rate λ changes over time, but $N(t)$ still has a Poisson distribution. A common example of this is when $\lambda(t)$ is a deterministic alternating function [e.g., $\lambda(t) = 2$ during the first 12 hours of each day, and $\lambda(t) = 1$ during the second 12 hours]. Intuitively then, a point process with an intensity is a generalization of a non-stationary Poisson process allowing for more complicated correlations over time.

Given any MPP Ψ, if $E[N(t)] < \infty$, $t \geq 0$, then $\{N(t)\}$ is always a non-negative right-continuous submartingale (with respect to its internal history), so the *Doob–Meyer decomposition* yields a right-continuous (and predictable) increasing process $\Lambda(t)$ (called the *compensator*) for which $\{N(t) - \Lambda(t)\}$ forms a mean-zero martingale. If $\Lambda(t)$ is of the form $\Lambda(t) = \int_0^t \lambda(s)\,\mathrm{d}s$, $t \geq 0$, where $\lambda(t)$ satisfies the regularity conditions of being non-negative, measurable, adapted to $\mathcal{F}_t$ and locally integrable [$\int_A \lambda(s)\,\mathrm{d}s < \infty$ for all bounded sets A], then $\lambda(t)$ is called the conditional intensity of the point process, or the *intensity* for short. (A predictable version of the intensity can always be chosen; this is done so by convention.) By the martingale property, an intensity can equivalently be defined as a stochastic process $\{\lambda(t)\}$ that satisfies the aforementioned regularity conditions and satisfies for all $s \leq t$

$$E[N(s,t] \mid \mathcal{F}_s] = E\left[\int_s^t \lambda(u)\,\mathrm{d}u \mid \mathcal{F}_s\right]\ .$$

Not all point processes admit an intensity. For example, a deterministic renewal process does not admit an intensity. The only part of $\mathcal{F}_t$ that is relevant for predicting the future of a renewal process is the backwards recurrence time $B(t)$, and if the interarrival time distribution F has a density f, then the renewal process admits an intensity $\lambda(t) = f[B(t-)]/\bar{F}[B(t-)]$, the hazard rate function of F evaluated at $B(t-)$. The fact that a density is needed illustrates the general fact that the existence of an intensity requires some smoothness in the distribution of points over time.

Incorporating marks into an intensity amounts to making rigorous the heuristic $\lambda(t, \mathrm{d}k)\mathrm{d}t = E[\Psi(\mathrm{d}t \times \mathrm{d}k)|\mathcal{F}_t]$, for some *intensity kernal* $\lambda(t, \mathrm{d}k)$ which in integral form becomes

$$E\Big[\int H(t,k)\Psi(\mathrm{d}t \times \mathrm{d}k)\Big] = E\Big[\int H(t,k)\lambda(t, \mathrm{d}k)\Big],$$

for non-negative and predictable H.

Here $\lambda(t, \mathrm{d}k)$ is a measure on the mark space for each t. Equivalently such an intensity kernel must have the properties that, for each mark set K, the process $\{\lambda(t, K) : t \geq 0\}$ is adapted to $\{\mathcal{F}_t\}$ and serves as an intensity for the thinned point process (defined by its counting process) $N_K(t) = \Psi[(0,t] \times K]$.

An elementary example is given by the compound Poisson process at rate λ with (independent of its points) iid jumps $k_n = X_n$ with some distribution $\mu(\mathrm{d}x) = P(X \in \mathrm{d}x)$. Then $\lambda(t, \mathrm{d}x) = \lambda\mu(\mathrm{d}x)$.

8.7.1 Time Changing to a Poisson Process

In some applications, it is desirable to construct (or simulate) a point process with a given intensity or corresponding compensator $\Lambda(t) = \int_0^t \lambda(s)\mathrm{d}s$. This can generally be accomplished by defining $N(t) = M[\Lambda(t)]$, where $M(t)$ is the counting process for an appropriate time-homogenous Poisson process at rate $\lambda = 1$. Conversely, the Poisson process can be retrieved by inverting the time change; $M(t) = N[\Lambda^{-1}(t)]$.

Theorem 8.8

Consider the counting process $\{N(t)\}$ of a (simple) MPP with intensity $\{\lambda(t)\}$ that is strictly positive and bounded. [Also assume that $\Lambda(t) \to \infty$, as $t \to \infty$, wp1.] Then $M(t) \stackrel{\text{def}}{=} N[\Lambda^{-1}(t)]$ defines a time-homogenous Poisson process at rate $\lambda = 1$.

There are some extensions of this result that incorporates the marks, in which case the time-homogenous Poisson process is replaced by a compound Poisson process.

8.7.2 Papangelou's Formula

Suppose that Ψ is asymptotically stationary (and ergodic), and that Ψ^* (two-sided) admits a conditional intensity $\lambda(t)$ with respect to a history $\mathcal{F}_t$.

Proposition 8.4 (Papangelou's formula)

For all non-negative random variables $X \in \mathcal{F}_{0-} \stackrel{\text{def}}{=} \cup_{t<0} \mathcal{F}_t$

$$\boldsymbol{E}^*[\lambda(0)X] = \lambda \boldsymbol{E}^0(X).$$

In other words, conditional on $\mathcal{F}_{0-}$, $\boldsymbol{P}^0$ is absolutely continuous with respect to $\boldsymbol{P}^*$, $\boldsymbol{P}^0 \ll \boldsymbol{P}^*$, having Radon–Nikodým derivative

$$\frac{\mathrm{d}\boldsymbol{P}^0}{\mathrm{d}\boldsymbol{P}^*}|\mathcal{F}_{0-} = \frac{\lambda(0)}{\lambda}.$$

Note that, when $X = 1$, the basic fact that $\boldsymbol{E}^*(\lambda(t)) = \lambda$ is retrieved. In many applications, $\mathcal{F}_t$ supports a stochastic process $\boldsymbol{X}$ jointly with Ψ in which case, letting $X = X(0-)$, Papangelou's formula yields $\boldsymbol{E}^*[\lambda(0)X(0-)] = \lambda\boldsymbol{E}^0[X(0-)]$. For example in the single-server queue Example 6 from Sect. 8.1.8, letting $X(t) = I[V(t) \leq x]$ yields $\boldsymbol{E}^*\{\lambda(0)I[V(0-) \leq x]\} = \lambda\boldsymbol{P}^0[V(0-) \leq x] = \lambda P(D \leq x)$. If arrivals are Poisson $[\lambda(t) = \lambda]$, this reduces to $P(V \leq x) = P(D \leq x)$ illustrating *Poisson arrivals see time averages* (PASTA), which asserts that when arrival times are Poisson for a queueing process $\{X(t)\}$ and satisfy a lack of anticipation condition (LAC), then the distribution of $X(0-)$ is the same under $\boldsymbol{P}^0$ and $\boldsymbol{P}^*$. LAC is defined as: for each $t \geq 0$ the future increments $\{N(t+s) - N(t) : s \geq 0\}$ are independent of the joint past $\{\{X(s) : s \leq t\}, \{N(s) : s \leq t\}\}$. Recalling the definition of $\boldsymbol{P}^0$ and $\boldsymbol{P}^*$ as Cesàro averages, PASTA says that if arrivals are Poisson then

the proportion of arrivals who find the queueing process in a given state is equal to the proportion of time the system is in that state;

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} I[X(t_j-) \in A] = \lim_{t\to\infty} \frac{1}{t} \int_0^t I[X(s) \in A]\mathrm{d}s.$$

Although its origins and primary applications are in queueing theory, PASTA can be applied to any joint pair $(\boldsymbol{X}, \Psi)$ for which the points of Ψ form a Poisson process satisfying LAC.

8.8 The Non-Ergodic Case

The assumption of ergodicity can be relaxed to allow for a stationary but non-ergodic framework. The key is in conditioning first on the *invariant σ-field* of $\mathbb{M}$,

$$\Im \stackrel{\text{def}}{=} \{\mathcal{E} \subset \mathbb{M} : \theta_t^{-1}\mathcal{E} = \mathcal{E},\ t \geq 0\}\,,$$

where $\theta_t^{-1}\mathcal{E} = \{\psi \in \mathbb{M} : \psi_t \in \mathcal{E}\}$ and the $\mathcal{E}$ are restricted to be Borel sets. Ergodicity is the case when $\boldsymbol{P}(\mathcal{E}) = 0$ or 1 for all $\mathcal{E} \in \Im$. Conditional expectation is denoted by $\boldsymbol{E}_{\Im}(X) \stackrel{\text{def}}{=} \boldsymbol{E}(X|\ \Im)$, and so on. A typical example of an invariant set is $\mathcal{E}_2 = \{\psi \in \mathbb{M} :\ \lim_{t\to\infty} N(t)/t = 2\}$, the set of all ψ with arrival rate exactly 2; it follows immediately that $\theta_t^{-1}\mathcal{E}_2 = \mathcal{E}_2$ because shifting a point process does not effect its long-term arrival rate.

A typical example of a stationary but non-ergodic MPP is a mixture of two Poisson processes, one at rate 1 and one at rate 2: flip a fair coin once. If it lands heads, then choose Ψ as a Poisson process at rate 1; if it lands tails, then choose Ψ as a Poisson process at rate 2. In this case $\boldsymbol{P}(\mathcal{E}_2) = 0.5$, Ψ is non-ergodic.

Inversion formulas still hold between Ψ^* and Ψ^0 if one conditions on $\Im$ first. For example, letting $\lambda_{\Im} \stackrel{\text{def}}{=} \boldsymbol{E}^*_{\Im}[N(1)]$, it holds that

$$P(t_0^* > x) = \boldsymbol{E}^0\{\lambda_{\mathrm{i}}\boldsymbol{E}^0_{\Im}(T_0 - x)_+\}\,. \tag{8.24}$$

In the above Poisson process case, $\lambda_{\Im} = 1$ if the coin lands heads, or 2 if it lands tails, and (8.24) reduces to

$$P(t_0^* > x) = \frac{(\mathrm{e}^{-x} + \mathrm{e}^{-2x})}{2}\,.$$

If the mixture was for two renewal processes with interarrival time distributions F_1 and F_2 respectively, then (8.24) reduces to

$$P(t_0^* > x) = \frac{[\overline{F}_{1,e}(x) + \overline{F}_{2,e}(x)]}{2}\,,$$

involving the two stationary excess distributions. The general inversion formula from $\boldsymbol{P}^0$ to $\boldsymbol{P}^*$ in functional form becomes

$$\boldsymbol{E}^*[f(\Psi)] = \boldsymbol{E}^0\left\{\lambda_{\mathrm{i}}\boldsymbol{E}^0_{\Im}\left[\int_0^{T_0} f(\Psi_s)\,\mathrm{d}s\right]\right\}\,.$$

8.9 MPPs in $\mathbb{R}^d$

When a point process has points in a higher-dimensional space such as $\mathbb{R}^d$, then the theory becomes more complicated. The main reason for this is that there is no longer a natural ordering for the points, e.g., there is no "next" point as is the case on $\mathbb{R}$. So "shifting to the j-th point" to obtain $\Psi_{(j)}$ is no longer well-defined. To make matters worse, point Cesàro limits as in (8.5) depend upon the ordering of the points. Whereas when $d = 1$ there is a one-to-one correspondence between stationary sequences of non-negative RVs (interarrival times) and point-stationary point processes, in higher dimensions such a simple correspondence is elusive. A good example to keep in mind is mobile phone usage, where the points (in $\mathbb{R}^2$ for simplicity) denote the locations of mobile phone users at some given time, and for each user the marks might represent whether a phone call is in progress or not. As in one dimension, it would be useful to consider analyzing this MPP from two perspectives: from the perspective of a "typical" user, and from the perspective of a "typical" spatial position in $\mathbb{R}^2$. For example, one might wish to estimate the average distance from a typical user to a base station, or the average distance from a typical position to a user with a call in progress. A mobile phone company trying to decide where to place some new base stations would benefit by such an analysis.

Some of the multidimensional complications can be handled, and initially it is best to use the measure approach from Sect. 8.1.5 to define an MPP. Starting with $\psi = \{(x_j, k_j)\}$, where $x_j \in \mathbb{R}^d$, it can equivalently be viewed as a σ-finite $\mathbb{Z}_+$-valued measure

$$\psi = \sum_j \delta_{(x_j, k_j)}\,,$$

on (the Borel sets of) $\mathbb{R}^d \times \mathbb{K}$.

The counting process is replaced by the counting measure $N(A) =$ the number of points that fall in the Borel set $A \subset \mathbb{R}^d$, and it is assumed that $N(A) < \infty$ for all bounded A. Simple means that the points x_j are distinct; $N(\{x\}) \leq 1$ for all $x \in \mathbb{R}^d$.

For any x, the shift mapping $\theta_x\psi = \psi_x$ is well defined via $\psi_x(A \times K) = \psi(A + x, K)$, where $A + x = \{y + x : y \in A\}$.

8.9.1 Spatial Stationarity in $\mathbb{R}^d$

Analogous to time stationarity in $\mathbb{R}$, the definition of *spatial stationarity* is that Ψ_x has the same distribution for all $x \in \mathbb{R}^d$, and as in (8.8) such MPPs can be viewed as arising as a Cesàro average over space, as follows.

Let B_r denote the d-dimensional ball of radius r centered at 0. Then (with l denoting Lebesgue measure in $\mathbb{R}^d$) a spatially stationary MPP is obtained via

$$P(\Psi^* \in \cdot) \stackrel{\text{def}}{=} \lim_{r\to\infty} \frac{1}{l(B_r)} \int_{B_r} P(\Psi_x \in \cdot)\,\mathrm{d}x\,.$$

In essence, we have randomly chosen our origin from over all of space.

Ergodicity means that the flow of shifts $\{\theta_x\}$ is ergodic. Stationarity implies that $E[N(A)] = \lambda l(A)$ for some λ, called the *mean density*; it can be computed by choosing (say) A as the unit hypercube $H = [0, 1]^d$; $\lambda = E[N(H)]$, the expected number of points in any set of volume 1.

An important example in applications is the Poisson process in $\mathbb{R}^d$. $N(A)$ has a Poisson distribution with mean $\lambda l(A)$ for all bounded Borel sets A, and $N(A_1)$ and $N(A_2)$ are independent if $A_1 \cap A_2 = \emptyset$.

8.9.2 Point Stationarity in $\mathbb{R}^d$

Coming up with a definition of point stationarity, however, is not clear, for what do we mean by "randomly selecting a point as the origin", and even if we could do just that what stationarity property would the resulting MPP have? (For example, even for a spatially stationary two-dimensional Poisson process, if a point is placed at the origin, it is not clear in what sense such a point process is stationary.) One would like to be able to preserve the distribution under a point shift, but which point can be chosen as the one to shift to as the new origin? Under ergodicity, one could define $P(\Psi^0 \in \cdot)$ as a sample-path average

$$P(\Psi^0 \in \cdot) \stackrel{\text{def}}{=} \lim_{r\to\infty} \frac{1}{N(B_r)} \sum_{x\in B_r} I(\Psi_x \in \cdot), \text{ wp1}\,.$$

It turns out that this can be improved to be more like (8.5) as follows. Let p_n denote the n-th point hit by B_r as $r \to \infty$ (if there are ties just order lexicographically). For each sample path of Ψ, $\{p_n\}$ is a permutation of $\{x_n\}$. Define

$$P(\Psi^0 \in \cdot) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} P(\Psi_{p_j} \in \cdot)\,.$$

Another approach involves starting with the spatially stationary MPP Ψ^* and defining $P(\Psi^0 \in \cdot)$ by inversion in the spirit of (8.15) and the Palm transformation, replacing a "unit of time" by any set A with volume 1, such as the unit hypercube $H = (0, 1]^d$:

$$P(\Psi^0 \in \cdot) = \lambda^{-1} E\left[\sum_{x\in H} I(\Psi_x^* \in \cdot)\right]. \tag{8.25}$$

Under ergodicity all these methods yield the same distribution. Ψ^0 has the property that there is a point at the origin, and its distribution is invariant under a two-step procedure involving an external randomization followed by a random point shift as follows (see Chapt. 9 of *Thorisson* [8.2]):

First, randomly place a ball B_r of any fixed radius $r > 0$ over the origin, e.g., take U distributed uniformly over the open ball B_r and consider the region $R = B_r + U$. There is at least one point in R, the point at the origin, but in any case let $n = N(R)$ denote the total number. Second, randomly choose one of the n points (e.g., according to the discrete uniform distribution) and shift to that point as the new origin. This shifted MPP *has the same distribution $P(\Psi^0 \in \cdot)$ as it started with.*

A recent active area of research is to determine whether or not one can achieve this invariance without any randomization. In other words is there an algorithm for choosing the "next point" to move to only using the sample paths of Ψ^0? In one dimension we know this is possible; always choose (for example) the point to the right (or left) of the current point. It turns out that in general this can be done (*Heveling* and *Last* [8.3]), but what is still not known is whether it can be done in such a way that all the points of the point process are exhaustively visited if the algorithm is repeated (as is the case in one dimension). For the Poisson process with $d = 2$ or 3 simple algorithms have indeed been found (*Ferrari* et al. [8.4]).

8.9.3 Inversion and Voronoi Sets

There is an analogue for the inverse part of the formula (8.25) in the spirit of (8.14), but now there is no "cycle" to average over so it is not clear what to do. It turns out that a random *Voronoi cell* is needed. For an MPP ψ with points $\{x_j\}$, for each point x_i define the

Voronoi cell about x_i by

$$V_{x_i}(\psi) = \{x \in \boldsymbol{R}^d : ||x - x_i|| < ||x - x_j|| ,\ \text{for all points } x_j \neq x_i\} ,$$

the set of elements in $\boldsymbol{R}^d$ that are closer to the point x_i than they are to any other point of ψ. For an MPP, this set is a random set containing x_i and of particular interest is when $\Psi = \Psi^0$ and $x_i = 0$, the point at the origin. We denote this Voronoi cell by V_0. It turns out that $E[l(V_0)] = \lambda^{-1}$, and

$$P(\Psi^* \in \cdot) = \lambda E\left[\int_{V_0} I\left(\Psi^0_x \in \cdot\right) \mathrm{d}x\right] . \tag{8.26}$$

The Voronoi cell V_0 plays the role that the interarrival time $T^0_0 = t^0_1$ does when $d = 1$. But, even when $d = 1$, V_0 is not the same as an interarrival time; instead it is given by the random interval $V_0 = (-T^0_{-1}/2, T^0_0/2) = (t^0_{-1}/2, t^0_1/2)$ which has length $l(V_0) = (t^0_1 + |t^0_{-1}|)/2$ and hence mean λ^{-1}. It is instructive to look closer at this for a Poisson process at rate λ, for then $l(V_0)$ has an Erlang distribution with mean λ^{-1}.

In the mobile phone context, if the points x_i are now the location of base stations (instead of phones) then V_{x_i} denotes the service zone for the base station, the region about x_i for which x_i is the closest base station. Any mobile user in that region would be best served (e.g., minimal distance) by being connected to the base at x_i. Thus all of space can be broken up into a collection of disjoint service zones corresponding to the Voronoi cells.

Finally, analogous to the $d = 1$ case, starting with a spatially stationary MPP it remains valid (in a limiting sense as in Theorem 8.8.4) that the distribution of Ψ^0 can be obtained as the conditional distribution of Ψ^* given a point at the origin. For example, placing a point at the origin for a spatially stationary Poisson process Ψ^* in $\mathbb{R}^d$ yields Ψ^0.

References

8.1 K. Sigman: *Stationary Marked Point Processes: An Intuitive Approach* (Chapman Hall, New York 1995)

8.2 H. Thorisson: *Coupling, Stationarity, and Regeneration* (Springer, Heidelberg Berlin New York 2000)

8.3 M. Heveling, G. Last: Characterization of Palm measures via bijective point-shifts, Annals of Probability **33**(5), 1698–1715 (2004)

8.4 P. A. Ferrari, C. Landim, H. Thorisson: Poisson trees, succession lines and coalescing random walks, Annals de L'Institut Henry Poincaré **40**, 141–152 (2004)