# Chapter 5

# Random Processes

Random processes, also known as stochastic processes, are used to model uncertain quantities that evolve in time: the trajectory of a particle, the price of oil, the temperature in New York, the national debt of the United States, etc. In these notes we introduce a mathematical framework that makes it possible to reason probabilistically about such quantities.

## 5.1 Definition

We denote random processes using a tilde over an upper case letter $\widetilde{X}$. This is **not** standard notation, but we want to emphasize the difference with random variables and random vectors. Formally, a random process $\widetilde{X}$ is a function that maps elements in a sample space $\Omega$ to real-valued functions.

**Definition 5.1.1** (Random process)**.** *Given a probability space $(\Omega, \mathcal{F}, \mathrm{P})$, a random process $\widetilde{X}$ is a function that maps each element $\omega$ in the sample space $\Omega$ to a function $\widetilde{X}(\omega, \cdot) : \mathcal{T} \to \mathbb{R}$, where $\mathcal{T}$ is a discrete or continuous set.*

There are two possible interpretations for $\widetilde{X}(\omega, t)$:

- If we fix $\omega$, then $\widetilde{X}(\omega, t)$ is a *deterministic function* of $t$ known as a **realization** of the random process.
- If we fix $t$ then $\widetilde{X}(\omega, t)$ is a *random variable*, which we usually just denote by $\widetilde{X}(t)$.

We can consequently interpret $\widetilde{X}$ as an infinite collection of random variables indexed by $t$. The set of possible values that the random variable $\widetilde{X}(t)$ can take for fixed $t$ is called the **state space** of the random process. Random processes can be classified according to the indexing variable or to their state space.

- If the indexing variable $t$ is defined on $\mathbb{R}$, or on a semi-infinite interval $(t_0, \infty)$ for some $t_0 \in \mathbb{R}$, then $\widetilde{X}$ is a **continuous-time** random process.
- If the indexing variable $t$ is defined on a discrete set, usually the integers or the natural numbers, then $\widetilde{X}$ is a **discrete-time** random process. In such cases we often use a different letter from $t$, such as $i$, as an indexing variable.

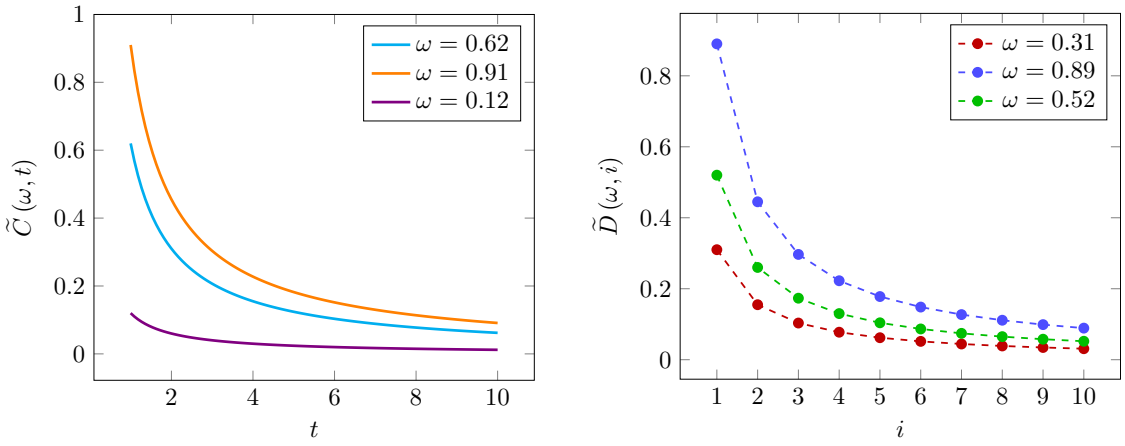


**Figure 5.1:** Realizations of the continuous-time (left) and discrete-time (right) random process defined in Example 5.1.2.

- If $\widetilde{X}(t)$ is a discrete random variable for all $t$, then $\widetilde{X}$ is a **discrete-state** random process. If the discrete random variable takes a finite number of values that is the same for all $t$, then $\widetilde{X}$ is a **finite-state** random process.
- If $\widetilde{X}(t)$ is a continuous random variable for all $t$, then $\widetilde{X}$ is a **continuous-state** random process.

Note that there are continuous-state discrete-time random processes and discrete-state continuous-time random processes. Any combination is possible.

The underlying probability space $(\Omega, \mathcal{F}, \mathrm{P})$ mentioned in the definition completely determines the stochastic behavior of the random process. In principle we can specify random processes by defining (1) a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ and (2) a mapping that assigns a function to each element of $\Omega$, as illustrated in the following example. This way of specifying random processes is only tractable for very simple cases.

**Example 5.1.2** (Puddle)**.** Bob asks Mary to model a puddle probabilistically. When the puddle is formed, it contains an amount of water that is distributed uniformly between 0 and 1 gallon. As time passes, the water evaporates. After a time interval $t$ the water that is left is $t$ times less than the initial quantity.

Mary models the water in the puddle as a continuous-state continuous-time random process $\widetilde{C}$. The underlying sample space is $(0, 1)$, the $\sigma$ algebra is the corresponding Borel $\sigma$ algebra (all possible countable unions of intervals in $(0, 1)$) and the probability measure is the uniform probability measure on $(0, 1)$. For a particular element in the sample space $\omega \in (0, 1)$

$$\widetilde{C}(\omega, t) := \frac{\omega}{t}, \quad t \in [1, \infty), \tag{5.1}$$

where the unit of $t$ is days in this example. Figure 6.1 shows different realizations of the random process. Each realization is a deterministic function on $[1, \infty)$.

Bob points out that he only cares what the state of the puddle is each day, as opposed to at any time $t$. Mary decides to simplify the model by using a continuous-state discrete-time random

process $\widetilde{D}$. The underlying probability space is exactly the same as before, but the time index is now discrete. For a particular element in the sample space $\omega \in (0, 1)$

$$\widetilde{D}(\omega, i) := \frac{\omega}{i}, \quad i = 1, 2, \ldots \tag{5.2}$$

Figure 6.1 shows different realizations of the continuous random process. Note that each realization is just a deterministic discrete sequence.

△

Recall that the value of the random process at a specific time is a random variable. We can therefore characterize the behavior of the process at that time by computing the distribution of the corresponding random variable. Similarly, we can consider the joint distribution of the process sampled at $n$ fixed times. This is given by the $n$th-order distribution of the random process.

**Definition 5.1.3** ($n$th-order distribution)**.** *The $n$th-order distribution of a random process $\widetilde{X}$ is the joint distribution of the random variables $\widetilde{X}(t_1)$, $\widetilde{X}(t_2)$, ..., $\widetilde{X}(t_n)$ for any $n$ samples $\{t_1, t_2, \ldots, t_n\}$ of the time index $t$.*

**Example 5.1.4** (Puddle (continued))**.** The first-order cdf of $\widetilde{C}(t)$ in Example 5.1.2 is

$$F_{\widetilde{C}(t)}(x) := \mathrm{P}\left(\widetilde{C}(t) \leq x\right) \tag{5.3}$$

$$= \mathrm{P}(\omega \leq t\, x) \tag{5.4}$$

$$= \begin{cases} \int_{u=0}^{t\,x} \mathrm{d}u = t\, x & \text{if } 0 \leq x \leq \frac{1}{t}\ , \\ 1 & \text{if } x > \frac{1}{t}\ , \\ 0 & \text{if } x < 0. \end{cases} \tag{5.5}$$

We obtain the first-order pdf by differentiating.

$$f_{\widetilde{C}(t)}(x) = \begin{cases} t & \text{if } 0 \leq x \leq \frac{1}{t}\ , \\ 0 & \text{otherwise.} \end{cases} \tag{5.6}$$

△

If the $n$th order distribution of a random process is shift-invariant, then the process is said to be strictly or strongly stationary.

**Definition 5.1.5** (Strictly/strongly stationary process)**.** *A process is stationary in a strict or strong sense if for any $n \geq 0$ if we select $n$ samples $t_1$, $t_2$, ..., $t_n$ and any displacement $\tau$ the random variables $\widetilde{X}(t_1)$, $\widetilde{X}(t_2)$, ..., $\widetilde{X}(t_n)$ have the same joint distribution as $\widetilde{X}(t_1 + \tau)$, $\widetilde{X}(t_2 + \tau)$, ..., $\widetilde{X}(t_n + \tau)$.*

The random processes in Example 5.1.2 are clearly not strictly stationary because their first-order pdf and pmf are not the same at every point. An important example of strictly stationary processes are independent identically-distributed sequences, presented in Section 5.3.

As in the case of random variables and random vectors, defining the underlying probability space in order to specify a random process is usually not very practical, except for very simple

cases like the one in Example 5.1.2. The reason is that it is challenging to come up with a probability space that gives rise to a given $n$-th order distribution of interest. Fortunately, we can also specify a random process by directly specifying its $n$-th order distribution *for all values of* $n = 1, 2, \ldots$ This completely characterizes the random process. Most of the random processes described in this chapter, e.g. independent identically-distributed sequences, Markov chains, Poisson processes and Gaussian processes, are specified in this way.

Finally, random processes can also be specified by expressing them as functions of other random processes. A function $\widetilde{Y} := g(\widetilde{X})$ of a random process $\widetilde{X}$ is also a random process, as it maps any element $\omega$ in the sample space $\Omega$ to a function $\widetilde{Y}(\omega, \cdot) := g(\widetilde{X}(\omega, \cdot))$. In Section 5.6 we define random walks in this way.

## 5.2 Mean and autocovariance functions

As in the case of random variables and random vectors, the expectation operator allows to derive quantities that summarize the behavior of the random process. The mean of the random vector is the mean of $\widetilde{X}(t)$ at any fixed time $t$.

**Definition 5.2.1** (Mean)**.** *The mean of a random process is the function*

$$\mu_{\widetilde{X}}(t) := \mathrm{E}\left(\widetilde{X}(t)\right). \tag{5.7}$$

Note that the mean is a *deterministic* function of $t$. The autocovariance of a random process is another deterministic function that is equal to the covariance of $\widetilde{X}(t_1)$ and $\widetilde{X}(t_2)$ for any two points $t_1$ and $t_2$. If we set $t_1 := t_2$, then the autocovariance equals the variance at $t_1$.

**Definition 5.2.2** (Autocovariance)**.** *The autocovariance of a random process is the function*

$$R_{\widetilde{X}}(t_1, t_2) := \mathrm{Cov}\left(\widetilde{X}(t_1), \widetilde{X}(t_2)\right). \tag{5.8}$$

*In particular,*

$$R_{\widetilde{X}}(t, t) := \mathrm{Var}\left(\widetilde{X}(t)\right). \tag{5.9}$$

Intuitively, the autocovariance quantifies the correlation between the process at two different time points. If this correlation only depends on the separation between the two points, then the process is said to be wide-sense stationary.

**Definition 5.2.3** (Wide-sense/weakly stationary process)**.** *A process is stationary in a wide or weak sense if its mean is constant*

$$\mu_{\widetilde{X}}(t) := \mu \tag{5.10}$$

*and its autocovariance function is* shift invariant*, i.e.*

$$R_{\widetilde{X}}(t_1, t_2) := R_{\widetilde{X}}(t_1 + \tau, t_2 + \tau) \tag{5.11}$$

*for any $t_1$ and $t_2$ and any shift $\tau$. For weakly stationary processes, the autocovariance is usually expressed as a function of the difference between the two time points,*

$$R_{\widetilde{X}}(s) := R_{\widetilde{X}}(t, t + s) \quad \text{for any } t. \tag{5.12}$$

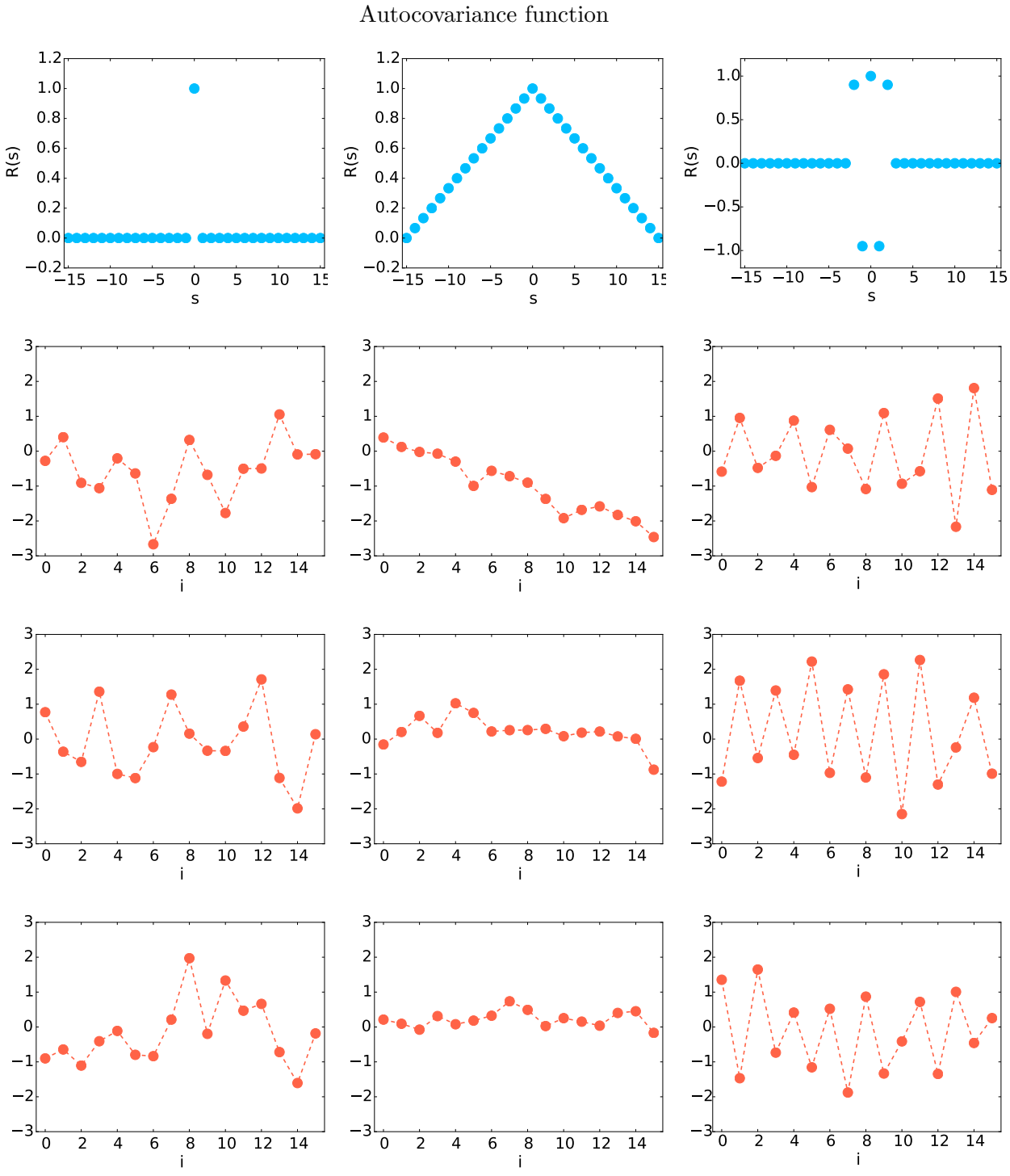


**Figure 5.2:** Realizations (bottom three rows) of Gaussian processes with zero mean and the autocovariance functions shown on the top row.

Note that any strictly stationary process is necessarily weakly stationary because its first and second-order distributions are shift invariant.

Figure 5.2 shows several stationary random processes with different autocovariance functions. If the autocovariance function is zero everywhere except at the origin, then the values of the random processes at different points are uncorrelated. This results in erratic fluctuations. When the autocovariance at neighboring times is high, the trajectory random process becomes smoother. The autocorrelation can also induce more structured behavior, as in the right column of the figure. In that example $\widetilde{X}(i)$ is negatively correlated with its two neighbors $\widetilde{X}(i-1)$ and $\widetilde{X}(i+1)$, but positively correlated with $\widetilde{X}(i-2)$ and $\widetilde{X}(i+2)$. This results in rapid periodic fluctuations.

## 5.3 Independent identically-distributed sequences

An independent identically-distributed (iid) sequence $\widetilde{X}$ is a discrete-time random process where $\widetilde{X}(i)$ has the same distribution for any fixed $i$ and $\widetilde{X}(i_1)$, $\widetilde{X}(i_2)$, $\ldots$, $\widetilde{X}(i_n)$ are mutually independent for any $n$ fixed indices and any $n \geq 2$. If $\widetilde{X}(i)$ is a discrete random variable (or equivalently the state space of the random process is discrete), then we denote the pmf associated to the distribution of each entry by $p_{\widetilde{X}}$. This pdf completely characterizes the random process, since for any $n$ indices $i_1, i_2, \ldots, i_n$ and any $n$:

$$p_{\widetilde{X}(i_1),\widetilde{X}(i_2),\ldots,\widetilde{X}(i_n)}(x_{i_1}, x_{i_2}, \ldots, x_{i_n}) = \prod_{i=1}^{n} p_{\widetilde{X}}(x_i). \tag{5.13}$$

Note that the distribution that does not vary if we shift every index by the same amount, so the process is strictly stationary.

Similarly, if $\widetilde{X}(i)$ is a continuous random variable, then we denote the pdf associated to the distribution by $f_{\widetilde{X}}$. For any $n$ indices $i_1, i_2, \ldots, i_n$ and any $n$ we have

$$f_{\widetilde{X}(i_1),\widetilde{X}(i_2),\ldots,\widetilde{X}(i_n)}(x_{i_1}, x_{i_2}, \ldots, x_{i_n}) = \prod_{i=1}^{n} f_{\widetilde{X}}(x_i). \tag{5.14}$$

Figure 5.3 shows several realizations from iid sequences which follow a uniform and a geometric distribution.

The mean of an iid random sequence is constant and equal to the mean of its associated distribution, which we denote by $\mu$,

$$\mu_{\widetilde{X}}(i) := \mathrm{E}\left(\widetilde{X}(i)\right) \tag{5.15}$$

$$= \mu. \tag{5.16}$$

Let us denote the variance of the distribution associated to the iid sequence by $\sigma^2$. The autocovariance function is given by

$$R_{\widetilde{X}}(i,j) := \mathrm{E}\left(\widetilde{X}(i)\,\widetilde{X}(j)\right) - \mathrm{E}\left(\widetilde{X}(i)\right)\mathrm{E}\left(\widetilde{X}(j)\right) \tag{5.17}$$

$$= \begin{cases} \sigma^2, \\ 0. \end{cases} \tag{5.18}$$

This is not surprising, $\widetilde{X}(i)$ and $\widetilde{X}(j)$ are independent for all $i \neq j$, so they are also uncorrelated.

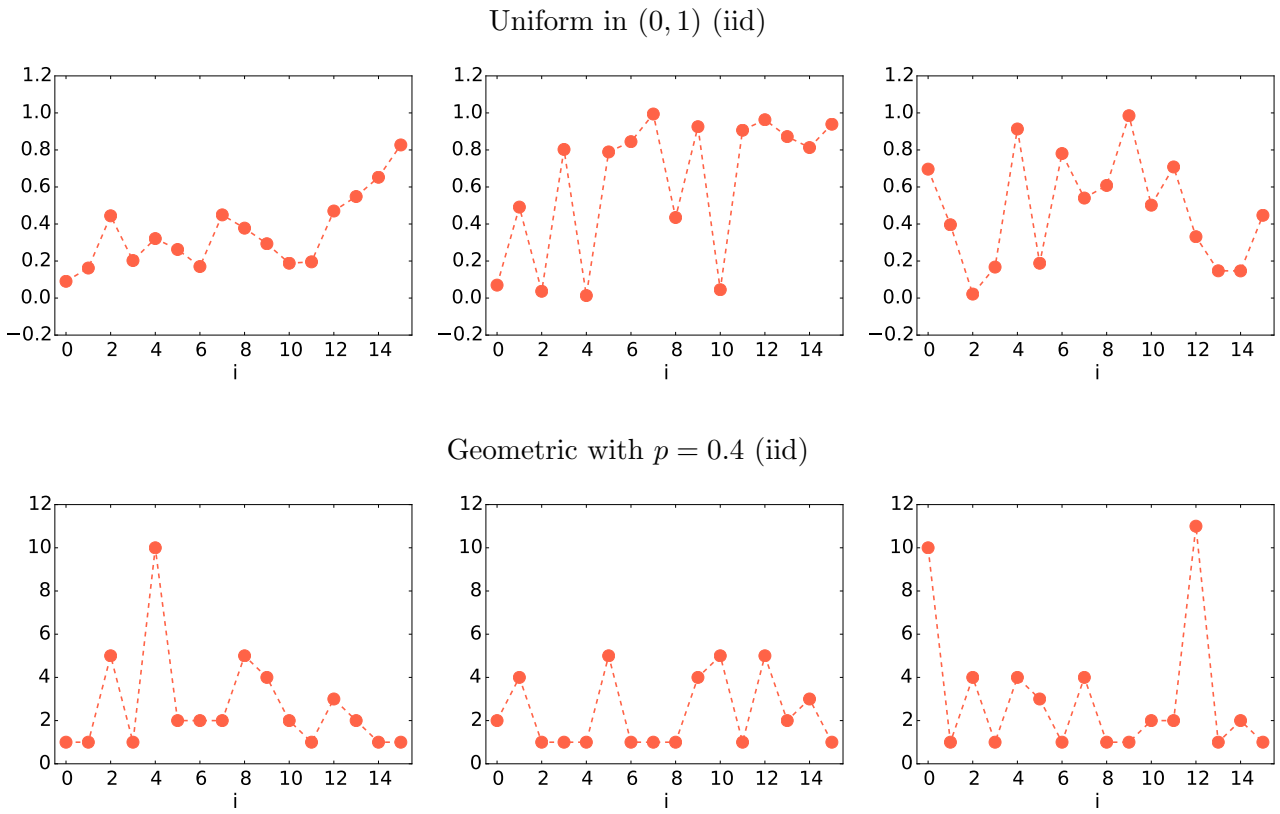


**Figure 5.3:** Realizations of an iid uniform sequence in $(0, 1)$ (first row) and an iid geometric sequence with parameter $p = 0.4$ (second row).

## 5.4 Gaussian process

A random process $\widetilde{X}$ is Gaussian if any set of samples is a Gaussian random vector. A Gaussian process $\widetilde{X}$ is fully characterized by its mean function $\mu_{\widetilde{X}}$ and its autocovariance function $R_{\widetilde{X}}$. For all $t_1$, $t_2$, ..., $t_n$ and any $n \geq 1$, the random vector

$$\vec{X} := \begin{bmatrix} \widetilde{X}(t_1) \\ \widetilde{X}(t_2) \\ \cdots \\ \widetilde{X}(t_n) \end{bmatrix} \tag{5.19}$$

is a Gaussian random vector with mean

$$\vec{\mu}_{\vec{X}} := \begin{bmatrix} \mu_{\widetilde{X}}(t_1) \\ \mu_{\widetilde{X}}(t_2) \\ \cdots \\ \mu_{\widetilde{X}}(t_n) \end{bmatrix} \tag{5.20}$$

and covariance matrix

$$\Sigma_{\vec{X}} := \begin{bmatrix} R_{\widetilde{X}}(t_1,t_1) & R_{\widetilde{X}}(t_1,t_2) & \cdots & R_{\widetilde{X}}(t_1,t_n) \\ R_{\widetilde{X}}(t_1,t_2) & R_{\widetilde{X}}(t_2,t_2) & \cdots & R_{\widetilde{X}}(t_2,t_n) \\ \vdots & \vdots & \ddots & \vdots \\ R_{\widetilde{X}}(t_2,t_n) & R_{\widetilde{X}}(t_2,t_n) & \cdots & R_{\widetilde{X}}(t_n,t_n) \end{bmatrix} \tag{5.21}$$

Figure 5.2 shows realizations of several discrete Gaussian processes with different autocovariance functions. Sampling from a Gaussian random process boils down to sampling a Gaussian random vector with the appropriate mean and covariance matrix.

**Algorithm 5.4.1** (Generating a Gaussian random process)**.** *To sample from an Gaussian random process with mean function $\mu_{\widetilde{X}}$ and autocovariance function $\Sigma_{\widetilde{X}}$ at $n$ points $t_1, \ldots, t_n$ we:*

1. *Compute the mean vector $\vec{\mu}_{\vec{X}}$ given by (5.20) and the covariance matrix $\Sigma_{\vec{X}}$ given by (5.21).*
2. *Generate $n$ independent samples from a standard Gaussian.*
3. *Color the samples according to $\Sigma_{\vec{X}}$ and center them around $\vec{\mu}_{\vec{X}}$, as described in Algorithm 4.3.15.*

## 5.5 Poisson process

In Example 2.2.8 we motivate the definition of Poisson random variable by deriving the distribution of the number of events that occur in a fixed time interval under the following conditions:

1. Each event occurs independently from every other event.
2. Events occur uniformly.
3. Events occur at a rate of $\lambda$ events per time interval.

We now assume that these conditions hold in the semi-infinite interval $[0, \infty)$ and define a random process $\widetilde{N}$ that counts the events. To be clear $\widetilde{N}(t)$ is the number of events that happen between 0 and $t$.

By the same reasoning as in Example 2.2.8, the distribution of the random variable $\widetilde{N}(t_2) - \widetilde{N}(t_1)$, which represents the number of events that occur between $t_1$ and $t_2$, is a Poisson random variable with parameter $\lambda(t_2 - t_1)$. This holds for any $t_1$ and $t_2$. In addition the random variables $\widetilde{N}(t_2) - \widetilde{N}(t_1)$ and $\widetilde{N}(t_4) - \widetilde{N}(t_3)$ are independent as along as the intervals $[t_1, t_2]$ and $(t_3, t_4)$ do not overlap by Condition 1. A Poisson process is a discrete-state continuous random process that satisfies these two properties.

Poisson processes are often used to model events such as earthquakes, telephone calls, decay of radioactive particles, neural spikes, etc. Figure 2.6 shows an example of a real scenario where the number of calls received at a call center is well approximated as a Poisson process (as long as we only consider a few hours). Note that here we are using the word *event* to mean *something that happens*, such as the arrival of an email, instead of a set within a sample space, which is the meaning that it usually has elsewhere in these notes.

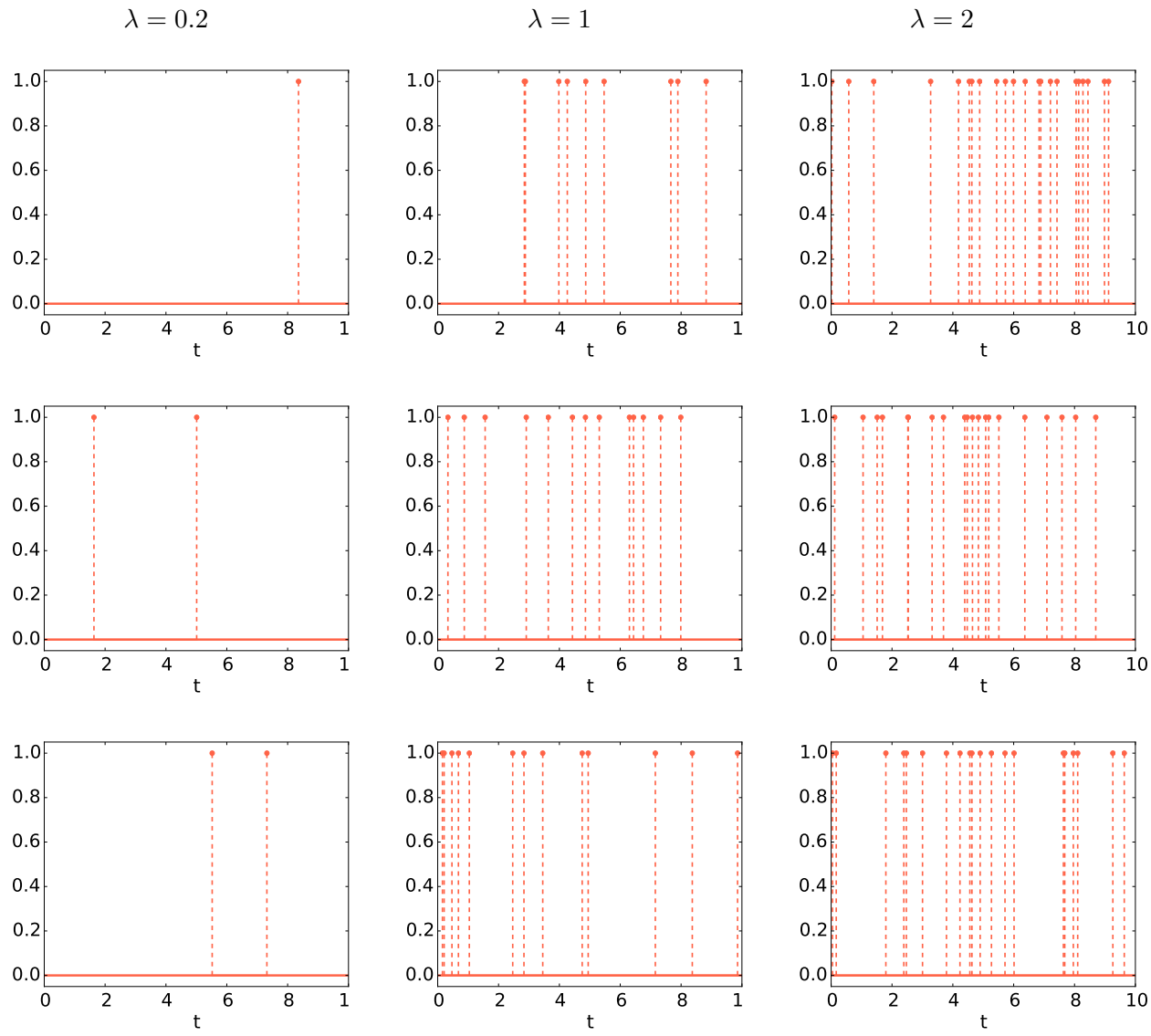


**Figure 5.4:** Events corresponding to the realizations of a Poisson process $\widetilde{N}$ for different values of the parameter $\lambda$. $\widetilde{N}(t)$ equals the number of events up to time $t$.

**Definition 5.5.1** (Poisson process)**.** *A Poisson process with parameter $\lambda$ is a discrete-state continuous random process $\widetilde{N}$ such that*

1. $\widetilde{N}(0) = 0$.
2. *For any $t_1 < t_2 < t_3 < t_4$ $\widetilde{N}(t_2) - \widetilde{N}(t_1)$ is a Poisson random variable with parameter $\lambda(t_2 - t_1)$.*
3. *For any $t_1 < t_2 < t_3 < t_4$ the random variables $\widetilde{N}(t_2) - \widetilde{N}(t_1)$ and $\widetilde{N}(t_4) - \widetilde{N}(t_3)$ are independent.*

We now check that the random process is well defined, by proving that we can derive the joint pmf of $\widetilde{N}$ at any $n$ points $t_1 < t_2 < \ldots < t_n$ for any $n \geq 0$. To alleviate notation let $p\left(\tilde{\lambda}, x\right)$ be the value of the pmf of a Poisson random variable with parameter $\tilde{\lambda}$ at $x$, i.e.

$$p\left(\tilde{\lambda}, x\right) := \frac{\tilde{\lambda}^x \, e^{-\tilde{\lambda}}}{x!}. \tag{5.22}$$

We have

$$\begin{aligned}
&p_{\widetilde{N}(t_1),\ldots,\widetilde{N}(t_n)}\left(x_1, \ldots, x_n\right) && (5.23)\\
&= \mathrm{P}\left(\widetilde{N}\left(t_1\right) = x_1, \ldots, \widetilde{N}\left(t_n\right) = x_n\right) && (5.24)\\
&= \mathrm{P}\left(\widetilde{N}\left(t_1\right) = x_1, \widetilde{N}\left(t_2\right) - \widetilde{N}\left(t_1\right) = x_2 - x_1, \ldots, \widetilde{N}\left(t_n\right) - \widetilde{N}\left(t_{n-1}\right) = x_n - x_{n-1}\right) && (5.25)\\
&= \mathrm{P}\left(\widetilde{N}\left(t_1\right) = x_1\right) \mathrm{P}\left(\widetilde{N}\left(t_2\right) - \widetilde{N}\left(t_1\right) = x_2 - x_1\right) \ldots \mathrm{P}\left(\widetilde{N}\left(t_n\right) - \widetilde{N}\left(t_{n-1}\right) = x_n - x_{n-1}\right)\\
&= p\left(\lambda t_1, x_1\right) p\left(\lambda\left(t_2 - t_1\right), x_2 - x_1\right) \ldots p\left(\lambda\left(t_n - t_{n-1}\right), x_n - x_{n-1}\right). && (5.26)
\end{aligned}$$

In words, we have expressed the event that $\widetilde{N}\left(t_i\right) = x_i$ for $1 \leq i \leq n$ in terms of the random variables $\widetilde{N}\left(t_1\right)$ and $\widetilde{N}\left(t_i\right) - \widetilde{N}\left(t_{i-1}\right)$, $2 \leq i \leq n$, which are independent Poisson random variables with parameters $\lambda t_1$ and $\lambda\left(t_i - t_{i-1}\right)$ respectively.

Figure 5.4 shows several sequences of events corresponding to the realizations of a Poisson process $\widetilde{N}$ for different values of the parameter $\lambda$ ($\widetilde{N}\left(t\right)$ equals the number of events up to time $t$). Interestingly, the interarrival time of the events, i.e. the time between contiguous events, always has the same distribution: it is an exponential random variable.

**Lemma 5.5.2** (Interarrival times of a Poisson process are exponential)**.** *Let $T$ denote the time between two contiguous events in a Poisson process with parameter $\lambda$. $T$ is an exponential random variable with parameter $\lambda$.*

The proof is in Section 5.7.1 of the appendix. Figure 2.11 shows that the interarrival times of telephone calls at a call center are indeed well modeled as exponential.

Lemma 5.5.2 suggests that to simulate a Poisson process all we need to do is sample from an exponential distribution.

**Algorithm 5.5.3** (Generating a Poisson random process)**.** *To sample from a Poisson random process with parameter $\lambda$ we:*

1. *Generate independent samples from an exponential random variable with parameter $\lambda$ $t_1$, $t_2$, $t_3$, ...*
2. *Set the events of the Poisson process to occur at $t_1$, $t_1 + t_2$, $t_1 + t_2 + t_3$, ...*

Figure 5.4 was generated in this way. To confirm that the algorithm allows to sample from a Poisson process, we would have to prove that the resulting process satisfies the conditions in Definition 5.5.1. This is indeed the case, but we omit the proof.

The following lemma, which derives the mean and autocovariance functions of a Poisson process is proved in Section 5.7.2.

**Lemma 5.5.4** (Mean and autocovariance of a Poisson process)**.** *The mean and autocovariance of a Poisson process equal*

$$\mathrm{E}\left(\widetilde{X}\left(t\right)\right) = \lambda t, \tag{5.27}$$

$$R_{\widetilde{X}}\left(t_1, t_2\right) = \lambda \min\left\{t_1, t_2\right\}. \tag{5.28}$$

The mean of the Poisson process is not constant and its autocovariance is not shift-invariant, so the process is neither strictly nor wide-sense stationary.

**Example 5.5.5** (Earthquakes)**.** The number of earthquakes with intensity at least 3 on the Richter scale occurring in the San Francisco peninsula is modeled using a Poisson process with parameter 0.3 earthquakes/year. What is the probability that there are no earthquakes in the next ten years and then at least one earthquake over the following twenty years?

We define a Poisson process $\widetilde{X}$ with parameter 0.3 to model the problem. The number of earthquakes in the next 10 years, i.e. $\widetilde{X}\left(10\right)$, is a Poisson random variable with parameter $0.3 \cdot 10 = 3$. The earthquakes in the following 20 years, $\widetilde{X}\left(30\right) - \widetilde{X}\left(10\right)$, are Poisson with parameter $0.3 \cdot 20 = 6$. The two random variables are independent because the intervals do not overlap.

$$\mathrm{P}\left(\widetilde{X}\left(10\right) = 0, \widetilde{X}\left(30\right) \geq 1\right) = \mathrm{P}\left(\widetilde{X}\left(10\right) = 0, \widetilde{X}\left(30\right) - \widetilde{X}\left(10\right) \geq 1\right) \tag{5.29}$$

$$= \mathrm{P}\left(\widetilde{X}\left(10\right) = 0\right) \mathrm{P}\left(\widetilde{X}\left(30\right) - \widetilde{X}\left(10\right) \geq 1\right) \tag{5.30}$$

$$= \mathrm{P}\left(\widetilde{X}\left(10\right) = 0\right) \left(1 - \mathrm{P}\left(\widetilde{X}\left(30\right) - \widetilde{X}\left(10\right) = 0\right)\right) \tag{5.31}$$

$$= e^{-3}\left(1 - e^{-6}\right) = 4.97\,10^{-2}. \tag{5.32}$$

The probability is 4.97%.

△

## 5.6 Random walk

A random walk is a discrete-time random process that models a sequence of steps in random directions. To specify a random walk formally, we first define an iid sequence of steps $\widetilde{S}$ such that

$$\widetilde{S}\left(i\right) = \begin{cases} +1 & \text{with probability } \frac{1}{2}, \\ -1 & \text{with probability } \frac{1}{2}. \end{cases} \tag{5.33}$$

We define a random walk $\widetilde{X}$ as the discrete-state discrete-time random process

$$\widetilde{X}\left(i\right) := \begin{cases} 0 & \text{for } i = 0, \\ \sum_{j=1}^{i} \widetilde{S}\left(j\right) & \text{for } i = 1, 2, \ldots \end{cases} \tag{5.34}$$

We have specified $\widetilde{X}$ as a function of an iid sequence, so it is well defined. Figure 5.5 shows several realizations of the random walk.

$\widetilde{X}$ is symmetric (there is the same probability of taking a positive step and a negative step) and begins at the origin. It is easy to define variations where the walk is non-symmetric and begins

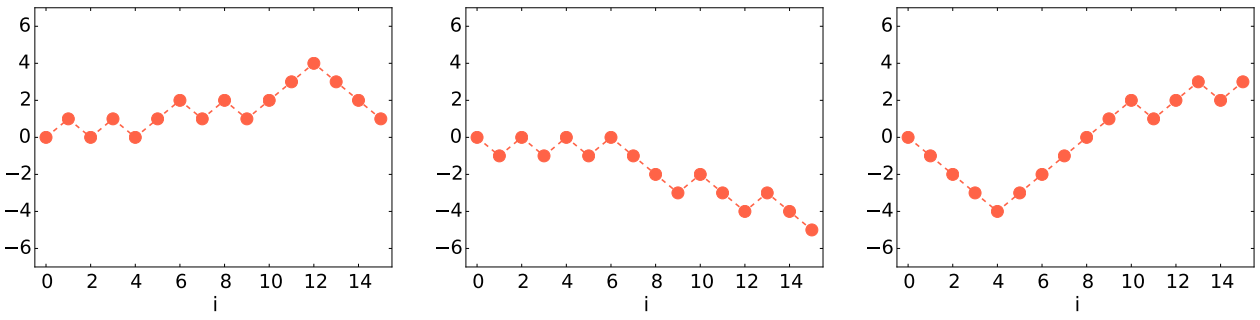


**Figure 5.5:** Realizations of the random walk defined in Section 5.5.

at another point. Generalizations to higher dimensional spaces– for instance to model random processes on a 2D surface– are also possible.

We derive the first-order pmf of the random walk in the following lemma, proved in Section 5.7.3 of the appendix.

**Lemma 5.6.1** (First-order pmf of a random walk)**.** *The first-order pmf of the random walk* $\widetilde{X}$ *is*

$$p_{\widetilde{X}(i)}(x) = \begin{cases} \binom{i}{\frac{i+x}{2}} \frac{1}{2^i} & \text{if } i+x \text{ is even and } -i \leq x \leq i \\ 0 & \text{otherwise.} \end{cases} \tag{5.35}$$

The first-order distribution of the random walk is clearly time-dependent, so the random process is not strictly stationary. By the following lemma, the mean of the random walk is constant (it equals zero). The autocovariance, however, is not shift invariant, so the process is not weakly stationary either.

**Lemma 5.6.2** (Mean and autocovariance of a random walk)**.** *The mean and autocovariance of the random walk* $\widetilde{X}$ *are*

$$\mu_{\widetilde{X}}(i) = 0, \tag{5.36}$$

$$R_{\widetilde{X}}(i,j) = \min\{i,j\}. \tag{5.37}$$

*Proof.*

$$\mu_{\widetilde{X}}(i) := \mathrm{E}\left(\widetilde{X}(i)\right) \tag{5.38}$$

$$= \mathrm{E}\left(\sum_{j=1}^{i} \widetilde{S}(j)\right) \tag{5.39}$$

$$= \sum_{j=1}^{i} \mathrm{E}\left(\widetilde{S}(j)\right) \qquad \text{by linearity of expectation} \tag{5.40}$$

$$= 0. \tag{5.41}$$

$$R_{\widetilde{X}}(i,j) := \mathrm{E}\left(\widetilde{X}(i)\,\widetilde{X}(j)\right) - \mathrm{E}\left(\widetilde{X}(i)\right)\mathrm{E}\left(\widetilde{X}(j)\right) \tag{5.42}$$

$$= \mathrm{E}\left(\sum_{k=1}^{i}\sum_{l=1}^{j}\widetilde{S}(k)\,\widetilde{S}(l)\right) \tag{5.43}$$

$$= \mathrm{E}\left(\sum_{k=1}^{\min\{i,j\}}\widetilde{S}(k)^2 + \sum_{k=1}^{i}\sum_{\substack{l=1\\ l\neq k}}^{j}\widetilde{S}(k)\,\widetilde{S}(l)\right) \tag{5.44}$$

$$= \sum_{k=1}^{\min\{i,j\}} 1 + \sum_{k=1}^{i}\sum_{\substack{l=1\\ l\neq k}}^{j}\mathrm{E}\left(\widetilde{S}(k)\right)\mathrm{E}\left(\widetilde{S}(l)\right) \tag{5.45}$$

$$= \min\{i,j\}, \tag{5.46}$$

where (5.45) follows from linearity of expectation and independence. □

The variance of $\widetilde{X}$ at $i$ equals $R_{\widetilde{X}}(i,i) = i$ which means that the standard deviation of the random walk scales as $\sqrt{i}$.

**Example 5.6.3** (Gambler)**.** A gambler is playing the following game. A fair coin is flipped sequentially. Every time the result is heads the gambler wins a dollar, every time it lands on tails she loses a dollar. We can model the amount of money earned (or lost) by the gambler as a random walk, as long as the flips are independent. This allows us to estimate that the expected gain equals zero or that the probability that the gambler is up 6 dollars or more after the first 10 flips is

$$\mathrm{P}(\text{gambler is up \$6 or more}) = p_{\widetilde{X}(10)}(6) + p_{\widetilde{X}(10)}(8) + p_{\widetilde{X}(10)}(10) \tag{5.47}$$

$$= \binom{10}{8}\frac{1}{2^{10}} + \binom{10}{9}\frac{1}{2^{10}} + \frac{1}{2^{10}} \tag{5.48}$$

$$= 5.47\,10^{-2}. \tag{5.49}$$

△

## 5.7 Proofs

### 5.7.1 Proof of Lemma 5.5.2

We begin by deriving the cdf of $T$,

$$F_T(t) := \mathrm{P}(T \leq t) \tag{5.50}$$

$$= 1 - \mathrm{P}(T > t) \tag{5.51}$$

$$= 1 - \mathrm{P}(\text{no events in an interval of length } t) \tag{5.52}$$

$$= 1 - e^{-\lambda t} \tag{5.53}$$

because the number of points in an interval of length $t$ follows a Poisson distribution with parameter $\lambda t$. Differentiating we conclude that

$$f_T(t) = \lambda e^{-\lambda t}. \tag{5.54}$$

### 5.7.2 Proof of Lemma 5.5.4

By definition the number of events between 0 and $t$ is distributed as a Poisson random variables with parameter $\lambda\, t$ and hence its mean is equal to $\lambda\, t$.

The autocovariance equals

$$R_{\widetilde{X}}\left(t_1, t_2\right) := \mathrm{E}\left(\widetilde{X}\left(t_1\right)\widetilde{X}\left(t_2\right)\right) - \mathrm{E}\left(\widetilde{X}\left(t_1\right)\right)\mathrm{E}\left(\widetilde{X}\left(t_2\right)\right) \tag{5.55}$$

$$= \mathrm{E}\left(\widetilde{X}\left(t_1\right)\widetilde{X}\left(t_2\right)\right) - \lambda^2 t_1 t_2. \tag{5.56}$$

By assumption $\widetilde{X}\left(t_1\right)$ and $\widetilde{X}\left(t_2\right) - \widetilde{X}\left(t_1\right)$ are independent so that

$$\mathrm{E}\left(\widetilde{X}\left(t_1\right)\widetilde{X}\left(t_2\right)\right) = \mathrm{E}\left(\widetilde{X}\left(t_1\right)\left(\widetilde{X}\left(t_2\right) - \widetilde{X}\left(t_1\right)\right) + \widetilde{X}\left(t_1\right)^2\right) \tag{5.57}$$

$$= \mathrm{E}\left(\widetilde{X}\left(t_1\right)\right)\mathrm{E}\left(\widetilde{X}\left(t_2\right) - \widetilde{X}\left(t_1\right)\right) + \mathrm{E}\left(\widetilde{X}\left(t_1\right)^2\right) \tag{5.58}$$

$$= \lambda^2 t_1\left(t_2 - t_1\right) + \lambda t_1 + \lambda^2 t_1^2 \tag{5.59}$$

$$= \lambda^2 t_1 t_2 + \lambda t_1. \tag{5.60}$$

### 5.7.3 Proof of Lemma 5.6.1

Let us define the number of positive steps $S_+$ that the random walk takes. Given the assumptions on $\widetilde{S}$, this is a binomial random variable with parameters $i$ and $1/2$. The number of negative steps is $S_- := i - S_+$. In order for $\widetilde{X}\left(i\right)$ to equal $x$ we need for the net number of steps to equal $x$, which implies

$$x = S_+ - S_- \tag{5.61}$$

$$= 2S_+ - i. \tag{5.62}$$

This means that $S_+$ must equal $\frac{i+x}{2}$. We conclude that

$$p_{\widetilde{X}(i)}\left(i\right) = \mathrm{P}\left(\sum_{j=0}^{i}\widetilde{S}\left(i\right) = x\right) \tag{5.63}$$

$$= \binom{i}{\frac{i+x}{2}}\frac{1}{2^i} \qquad \text{if } \frac{i+x}{2} \text{ is an integer between 0 and } i. \tag{5.64}$$