Chapter 2

Multivariate time series analysis: Some essential notions

An overview of a modeling and learning framework for multivariate time series was presented in Chapter 1. In this chapter, some notions on **multivariate time series** analysis in time and frequency domains are succinctly introduced; tools and conventions used herein are essential to appreciate the contributions later in the thesis. Although they are widely available in textbooks, they have been adapted appropriately to suit this thesis.

In Section 2.1, a multivariate time series model and the concept of **weak stationarity** are formally defined; only those time series that are weakly stationary are considered throughout this thesis. Weak stationarity requires defining the **autocovariance function** of a time series. Autocovariance characteristics of a few relevant types of stationary multivariate time series are presented.

In Section 2.2, frequency or spectral domain concepts belonging to Fourier analysis are introduced. The motion of a **simple pendulum** is used as an example to motivate the presentation of **Fourier series**; this is subsequently extended to the **Fourier transform**. Clearly, this example to introduce Fourier analysis is a detour to a **continuous time process**; but it will enhance understanding of spectral domain tools, notations, and definitions.

In Section 2.3, the discrete time process is introduced as a limiting case of continuous time processes; this leads to **discrete time Fourier transform**. Discrete time Fourier transform gives a periodic and continuous spectrum and it underpins important developments in subsequent sections. There, **discrete Fourier transform** of a **discrete time process** is also discussed.

In Section 2.4, after having defined time and spectral domain characteristics of a deterministic process, spectral analysis of stationary time series is presented. The two most important ideas that need to be taken from this chapter are presented next: First, the relation between autocovariance function and **spectral density function** is simply that of a Fourier transform. Second, the probability distribution of discrete Fourier transform components of a linear process is **complex-Gaussian** within small **subbands** of frequencies. The first idea is a direct application of Fourier analysis derived in earlier sections. For the second idea, the asymptotic theory of spectral estimates is involved. In Section 2.5, therefore, the asymptotic probability distribution function of discrete Fourier transform components is provided without proof.

2.1 Temporal analysis of stationary processes

In this section, an introductory review of the time-domain or temporal analysis of time series is performed. It starts by adapting some definitions from the literature of random processes [29, 68, 86, 47]; the presented definitions might be termed differently by various authors elsewhere in the literature.

An infinite sequence of random variables forms a random process. Then, a **vector random process** is defined as an infinite sequence of vector random variables that is a set of random variables maintaining the same order of the vector components in every realization.

Definition 2.1. *A (multivariate)* ***time series*** *of a (vector) random process is a connected subsequence formed by its constituent (vector) random variables.*

A time series is called so because the index of the sequence is often attributed to time instants. If y_t is the random variable at time instant $t \in \mathbb{Z}$ of a random process of interest, then $\{y_t\}$, $t = 1, \ldots, \tau$ shall be called a τ-length realization of a time series whose t-th **sample** is y_t. An r-dimensional vector random process generates an infinite sequence $\{\mathsf{y}_t\}$ of r-dimensional vector random variables $\mathsf{y}_t = [\mathsf{y}_{1_t} \cdots \mathsf{y}_{r_t}]'$, where y_{i_t}, $i = 1, \ldots, r$ are the component random variables at time instant $t \in \mathbb{Z}$. Figure 2.1 selects a realization of a τ-length r-variate time series whose t-th sample is the vector $y_t \in \mathbb{R}^r$.

It may now be implied that when referring to the term 'process' in Chapter 1 in a broad sense, it meant the vector random process underlying a multivariate time series. On similar lines, the term 'model' there referred to the joint probability distribution function of the samples of the time series; this is because it is a set of random variables that is dealt with. Then, the model of a process corresponding to a τ-length r-variate time series requires evaluating the $\tau \times r$-dimensional joint probability distribution function $\mathsf{P}(\mathsf{y}_1 \leq c_1,\ \ldots,\ \mathsf{y}_\tau \leq c_\tau)$ for any constant vector $c_t \in \mathbb{R}^r$, $t = 1, \ldots, \tau$, where P denotes probability and the comparison of vectors are component-wise. Of course, direct evaluation of such a probability distribution is very unwieldy. Therefore, restricting the scope of the studies and bringing forth assumptions to simplify the process is inevitable for modeling a process generating a multivariate time series.

Let a few useful terms associated with random variables be first defined [95].

Definition 2.2. *The* ***probability density function*** p^{u} *of a random variable* u *is defined as* $\mathsf{p}^{\mathsf{u}}(a) = \frac{\mathsf{d}}{\mathsf{d}a}\mathsf{P}(\mathsf{u} \leq a)\ \forall a \in \mathbb{R}$, *wherever the derivative exists.*

In the above definition, $\mathsf{p}^{\mathsf{u}}(a)$ is any positive finite real number wherever the derivative does not exist. Then the joint probability density function $\mathsf{p}^{\mathsf{u}_1,\ldots,\mathsf{u}_r}$ of r random variables $\mathsf{u}_1, \ldots, \mathsf{u}_r$ may be given by $\mathsf{p}^{\mathsf{u}_1,\ldots,\mathsf{u}_r}\ (a_1, \ldots, a_r) = \partial^r \mathsf{P}(\mathsf{u}_1 \leq a_1,\ \ldots,\ \mathsf{u}_r \leq a_r)$ $/\partial a_1\ \cdots\ \partial a_r\ \forall a = [a_1\ \cdots\ a_r]' \in \mathbb{R}^r$ and $\mathsf{p}^{\mathsf{u}_1,\ldots,\mathsf{u}_r}(a_1, \ldots, a_r)$ is any positive finite real number wherever the derivative does not exist.

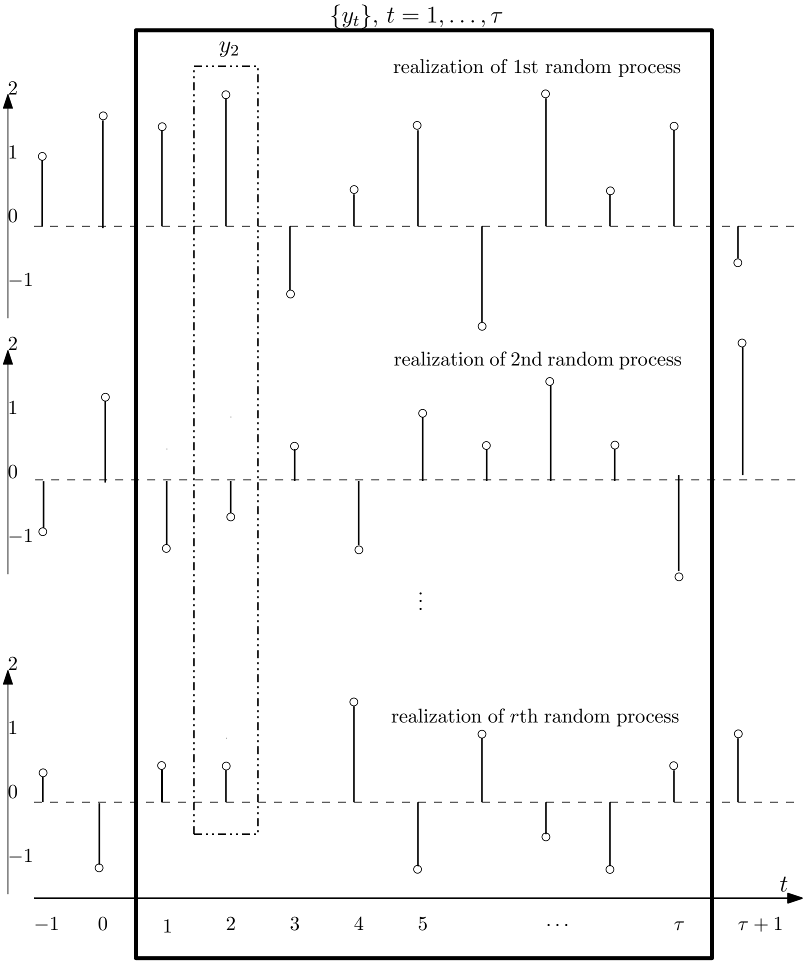

Figure 2.1: The second sample y_2 of a realization of a τ-length r-variate time series $\{\mathsf{y}_t\}$ is highlighted.

Definition 2.3. *The* ***multivariate probability density function*** p^{u} *of an* r*-dimensional vector random variable* $\mathsf{u} = [\mathsf{u}_1 \ \cdots \ \mathsf{u}_r]'$ *is defined as the joint probability density function of its* r *component random variables, i.e.,* $\mathsf{p}^{\mathsf{u}}(a) = \mathsf{p}^{\mathsf{u}_1,\ldots,\mathsf{u}_r}(a_1, \ldots, a_r)\ \forall a = [a_1 \cdots a_r]' \in \mathbb{R}^r$.

Definition 2.4. *For an* r*-dimensional vector random variable* u *whose probability density function* $\mathsf{p}^{\mathsf{u}}(b)$ *exists* $\forall b \in \mathbb{R}^r$*, the* ***expectation*** *of a function* $g(\mathsf{u})$ *with respect to* p^{u} *is defined as* $\mathsf{E}^{\mathsf{u}}[g(\mathsf{u})] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(b)\ \mathsf{p}^{\mathsf{u}}(b)\ \mathsf{d}b$.

Definition 2.5. *The* ***mean*** $\mu^{\mathsf{u}} \in \mathbb{R}^r$ *of an* r*-dimensional vector random variable* u *is defined as its expectation with respect to its* r*-variate probability density function, i.e.,* $\mu^{\mathsf{u}} = \mathsf{E}^{\mathsf{u}}[\mathsf{u}]$.

Definition 2.6. *The **variance (covariance matrix)** Γ^{u} of the (r- dimensional vector) random variable* u *is defined as the expectation, with respect to its (multivariate) probability density function* p^{u}*, of the (outer) product of the (vector) random variable with itself about its mean* μ^{u}*, i.e.,* $\Gamma^{\mathsf{u}} = \mathsf{E}^{\mathsf{u}}\ [(\mathsf{u} - \mu^{\mathsf{u}})\ (\mathsf{u} - \mu^{\mathsf{u}})']$.

Definition 2.7. *The **cross-covariance** $\Gamma^{\mathsf{u},\mathsf{v}}$ between the (vector) random variables* u *and* v *is defined as the expectation, with respect to their joint (multivariate) probability density function* $\mathsf{p}^{\mathsf{u},\mathsf{v}}$*, of the (outer) product of the random variables about their respective means* μ^{u} *and* μ^{v}*, i.e.,* $\Gamma^{\mathsf{u},\mathsf{v}} = \mathsf{E}^{\mathsf{u},\mathsf{v}}\ [(\mathsf{u} - \mu^{\mathsf{u}})\ (\mathsf{v} - \mu^{\mathsf{v}})']$.

Definition 2.8. *The cross-covariance between any two constituent (vector) random variables of the same (vector) random process is called the **autocovariance function** between the (vector) random variables.*

Therefore, the autocovariance function (**acvf**) between the r- dimensional vector random variables y_t and y_s $\forall t, s \in \mathbb{Z}$ is

$$\Gamma^{\mathsf{y}_t,\mathsf{y}_s} = \mathsf{E}^{\mathsf{y}_t,\mathsf{y}_s}[(\mathsf{y}_t - \mu^{\mathsf{y}_t})(\mathsf{y}_s - \mu^{\mathsf{y}_s})']. \tag{2.1}$$

Note 2.1. *When it specifically concerns a univariate time series* $\{\mathsf{y}_t\}$ *and not a multivariate time series, its acvf will be denoted by* $\gamma^{\mathsf{y}_t,\mathsf{y}_s}$*. Then, for a multivariate time series* $\{\mathsf{y}_t\} = \{[\mathsf{y}_{1_t}\, \mathsf{y}_{2_t} \cdots \mathsf{y}_{r_t}]'\}$*, the* (i, j)*-th element of its acvf* $\Gamma^{\mathsf{y}_t,\mathsf{y}_s}$ *may be written as* $\gamma^{\mathsf{y}_{i_t},\mathsf{y}_{j_s}}$*, which according to Definition 2.7, may be interpreted as the cross-covariance between* y_{i_t} *and* y_{j_s} $\forall i, j \in 1, \ldots, r$ *and* $\forall t, s \in \mathbb{Z}$.

Definition 2.9. *A (vector) time series* $\{\mathsf{y}_t\}$ $\forall t \in \mathbb{Z}$ *is **weakly stationary** if the mean* μ^{y_t} *is a constant (vector)* μ^{y} *and its acvf* $\Gamma^{\mathsf{y}_t,\mathsf{y}_s}$ *between the (vector) random variables* y_t *and* y_s $\forall s \in \mathbb{Z}$ *depends on* s *and* t *only through* $s - t$.

The variable $h = s - t$ of the *acvf* $\Gamma_h^{\mathsf{y}_t,\mathsf{y}_s}$ will be referred to as the **lag**. It follows from Definition 2.9 that the *acvf* between y_{t+h} and y_t of a weakly stationary time series $\{\mathsf{y}_t\}$ is

$$\Gamma_h^{\mathsf{y}} \triangleq \Gamma^{\mathsf{y}_{t+h},\mathsf{y}_t} = \mathsf{E}^{\mathsf{y}_{t+h},\mathsf{y}_t}[(\mathsf{y}_{t+h} - \mu^{\mathsf{y}})(\mathsf{y}_t - \mu^{\mathsf{y}})'] \quad \forall h \in \mathbb{Z}. \tag{2.2}$$

It is easy to verify that $\gamma_h^{\mathsf{y}_i,\mathsf{y}_j} = \gamma_{-h}^{\mathsf{y}_j,\mathsf{y}_i}$, which gives rise to the following property of the *acvf*:

Property 2.1. *A weakly stationary acvf is transpose symmetric about* $h = 0$*, i.e.,*

$$(\Gamma_h^{\mathsf{y}})' = \Gamma_{-h}^{\mathsf{y}}. \tag{2.3}$$

In this thesis, the focus is on time series that are weakly stationary and the main references on that topic are [102, 99, 111, 19, 20]. Now, take a look at a few examples of weakly stationary multivariate time series.

Property 2.2. *A weakly stationary r-variate time series $\{\mathsf{z}_t\}$ is **idiosyncratic** if any two components z_{i_t} and $\mathsf{z}_{j_{t+h}}$ $\forall h \in \mathbb{Z}$, $i \neq j$ of its corresponding vector random variable $\mathsf{z}_t = [\mathsf{z}_{1_t} \cdots \mathsf{z}_{r_t}]$ $\forall t \in \mathbb{Z}$ have zero cross-covariance, i.e., $\gamma^{\mathsf{z}_{i_t},\mathsf{z}_{j_{t+h}}} \triangleq \gamma_h^{\mathsf{z}_i,\mathsf{z}_j} \in \mathbb{R}$ is zero whenever $i \neq j$ $\forall i, j \in 1, \ldots, r$ and $h \in \mathbb{Z}$.*

Note 2.2. *The diagonal elements of an acvf Γ^{u} of the vector random variable $\mathsf{u}_t = [\mathsf{u}_{1_t}\, \mathsf{u}_{2_t} \cdots \mathsf{u}_{r_t}]'$ will be written simply as $\gamma_h^{\mathsf{u}_i} \triangleq \gamma_h^{\mathsf{u}_i,\mathsf{u}_i}$ $\forall i \in 1, \ldots, r$.*

The *acvf* of an r-variate idiosyncratic time series $\{\mathsf{z}_t\}$ due to vector random variable

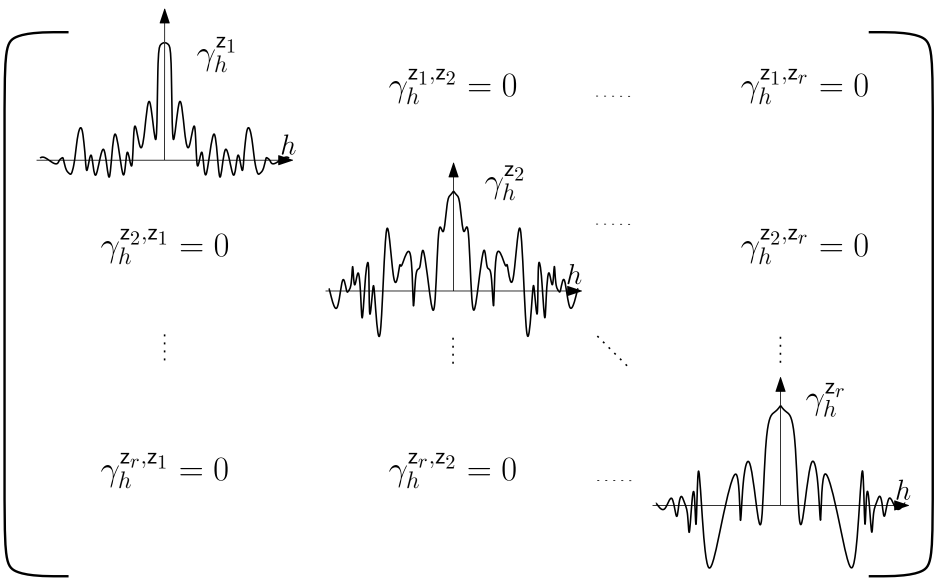

Figure 2.2: The structure of the $r \times r$ *acvf* matrix Γ_h^{z} of an r-variate idiosyncratic time series $\{\mathsf{z}_t\}$ shows zeros off-diagonal due to no cross-correlation between its components. The plots are hypothetical and interpolated for, an otherwise discrete h, for clarity.

$\mathsf{z}_t = [\mathsf{z}_{1_t} \ldots \mathsf{z}_{r_t}]'$ has an $r \times r$ matrix structure shown in Figure 2.2. Following Note 2.2, the cross-covariance $\gamma_h^{\mathsf{z}_i,\mathsf{z}_j}$ of Property 2.2 may be written as the *acvf* $\gamma_h^{\mathsf{z}_i}$ of z_i whenever $i = j$, whereas $\gamma_h^{\mathsf{z}_i,\mathsf{z}_j} = 0$ is zero otherwise. This means that the off-diagonal elements of such an *acvf* are always zero. Let a special case of an idiosyncratic time series whose each diagonal element of the *acvf* is an impulse function be now defined.

Definition 2.10. *For a weakly stationary (vector) time series $\{\mathsf{x}_t\}$, if (the components of) x_t are independently and identically distributed $\forall t \in \mathbb{Z}$, then $\{\mathsf{x}_t\}$ is said to be (multivariate) **white noise**.*

Note that the *acvf* Γ_h^{x} of a q-variate white noise $\{\mathsf{x}_t\}$ is $\Gamma_h^{\mathsf{x}} = 0_q$ $\forall h \neq 0$ and $\det(\Gamma_0^{\mathsf{x}}) \neq 0$. Definition 2.10 implies that mean-subtracted white noise components may be defined completely by their component variances $\sigma_{i_h}^2 = \sigma_i^2$ $\forall i = 1, \ldots, q$ so that $\Gamma_h^{\mathsf{x}} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_q^2)$ $\forall h \in \mathbb{Z}$, which is said to be **isotropic** if $\sigma_i = \sigma$,

$i = 1, 2, \ldots, q$. A special case of the zero-mean isotropic white noise is the following: If the component random variables of a q-variate white noise $\{\mathsf{x}_t\}$ have zero means and unit variances so that $\Gamma_h^{\mathsf{x}} = \mathsf{diag}(1, \ldots, 1) \in \mathbb{R}^{q\times q}\, \forall h \in \mathbb{Z}$, then $\{\mathsf{x}_t\}$ is termed a **zero-mean unit-variance white noise**. The white noise is used as the ingredient in many weakly stationary time series process models because of the simplicity of its *acvf*. Defined below is one such model; refer §11.1 of [20] or §9.2 of [79] among many references for its details.

Definition 2.11. *For a q-variate zero-mean white noise $\{\mathsf{x}_t\}$ $\forall t \in \mathbb{Z}$ and matrices $C_j \in \mathbb{R}^{q\times q}$ $\forall j \in \mathbb{Z}$ with absolutely summable elements, a* ***linear process*** *is defined as the q-variate time series*

$$\mathsf{u}_t = \sum_{j\in\mathbb{Z}} C_j \mathsf{x}_{t-j}, \tag{2.4}$$

which is weakly stationary with zero mean and acvf

$$\Gamma_h^{\mathsf{u}} = \sum_{j\in\mathbb{Z}} C_{j+h} \Gamma_0^{\mathsf{x}} C_j'. \tag{2.5}$$

2.2 Spectral analysis of continuous processes

The purpose of this section is to introduce certain Fourier analysis concepts required for this thesis.

Note 2.3. *This section deviates momentarily to discuss continuous time processes; the time index is $t \in \mathbb{R}$; everywhere else in this thesis $t \in \mathbb{Z}$.*

Consider the motion of a simple pendulum as an example of a periodic continuous process. It is assumed for simplicity that the mass of the string attached to the bob of the pendulum is negligible. The oscillation is restricted to a plane so that the constant string length and an angle, viz., the instantaneous angle that the string forms with respect to its equilibrium position, are sufficient to describe its motion. It is also assumed that the amplitude α, which is the maximum displacement of the bob from its position of equilibrium, is very small relative to the length of the pendulum. Refer to Figure 2.3; let τ be the time period of oscillation so that τ^{-1} is the frequency of oscillation. The standard association of 2π radians to be equivalent to one complete oscillation may be made. Let ϕ radians be the part of 2π radians of an oscillation the pendulum has completed at time $t = 0$; its sign depends on the choice of direction of reference of the bob's trajectory. If it is assumed that the pendulum is undamped by any kinds of friction and disturbances, then the pendulum's displacement with respect to the equilibrium position of the string at time $t \in \mathbb{R}$ is $y_t = \alpha \cos(2\pi \frac{t}{\tau} + \phi)$. Basic trigonometric identities enable writing y_t in various combinations of sinusoids, e.g.,

$$y_t = a \cos(2\pi t/\tau) + b \sin(2\pi t/\tau), \tag{2.6}$$

where $a = \alpha \cos(\phi)$ and $b = -\alpha \sin(\phi)$.

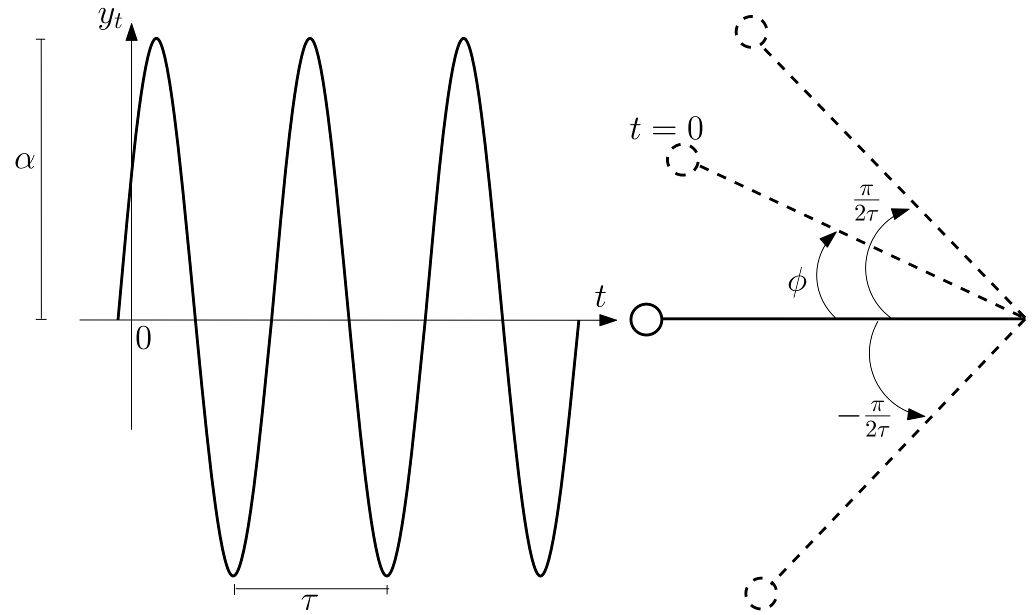

Figure 2.3: The motion of the simple pendulum registers a continuous function based on the displacement of its bob.

Just seen is the decomposition of a basic equation of oscillation into two sinusoids of frequency τ^{-1}. Consider that y_t was expressed as a weighted sum of two basis functions; this is because the sinusoids here are orthogonal functions, i.e., $\int_{-\frac{1}{2}\tau}^{\frac{1}{2}\tau} \cos(2\pi t/\tau) \sin(2\pi t/\tau)\, \mathrm{d}t = 0$. The necessity of orthogonal functions in many problems is analogous to the necessity of orthogonal coordinate axes in expressing the position of a point in a Cartesian plane.

The decomposition of a time-domain process to its frequency components is known as **Fourier (spectral) analysis** and the definitions presented in this and Section 2.3 on this topic can be found in references such as [99, 44, 93, 89]. Fourier analysis is based on one of the most important contributions to the sciences originally formalized by Joseph Fourier in 1807 that any 'well-behaved' deterministic continuous periodic function y_t could be expressed as a sum of orthogonal functions if and only if the orthogonal functions are sinusoids, where a 'well-behaved' function satisfies the following condition:

Definition 2.12. *A function $y_t\, \forall t \in \mathbb{R}$ is said to be **absolutely summable** if*

$$\int_{-\infty}^{\infty} |y_t|\, \mathrm{d}t < \infty. \tag{2.7}$$

The unique decomposition of such a deterministic periodic function y_t into a possibly infinite number of sinusoids is called its Fourier series representation: $y_t = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos(2\pi nt/\tau) + b_n \sin(2\pi nt/\tau))$, where $a_m = \frac{2}{\tau} \int_{-\frac{1}{2}\tau}^{\frac{1}{2}\tau} y_t \cos(2m\pi t/\tau)\, \mathrm{d}t$, $m = 0, 1, 2, \ldots$ and $b_l = \frac{2}{\tau} \int_{-\frac{1}{2}\tau}^{\frac{1}{2}\tau} y_t \sin(2l\pi t/\tau)\, \mathrm{d}t$, $l = 1, 2, 3, \ldots$. It is often convenient to use Euler identity $e^{\mathrm{i}\theta} = \cos(\theta) + \mathrm{i}\sin(\theta)$ to reach the following definition which holds for complex-valued functions also.

Definition 2.13. *The **Fourier series** representation of a deterministic continuous periodic function $y_t \in \mathbb{C}\ \forall t \in \mathbb{R}$ satisfying (2.7) is*

$$y_t = \sum_{m=-\infty}^{\infty} c_m e^{\mathrm{i}2\pi mt/\tau} \tag{2.8}$$

where $c_m \in \mathbb{C}\ \forall m \in \mathbb{Z}$ is

$$c_m = \frac{1}{\tau}\int_{-\frac{1}{2}\tau}^{\frac{1}{2}\tau} y_t e^{-\mathrm{i}2\pi mt/\tau}\mathrm{d}t. \tag{2.9}$$

Note that if $c_m = \overline{c}_m\ \forall m \in \mathbb{Z}$, then the function $y_t \in \mathbb{R}$, else $y_t \in \mathbb{C}$.
Involve frequency spacing $\Delta u = 2\pi/\tau$ and write the Fourier series coefficients as $c_m = \frac{\Delta u}{2\pi}\int_{-\frac{1}{2}\tau}^{\frac{1}{2}\tau} y_t\ e^{-\mathrm{i}mt\Delta u}\mathrm{d}t$. Substituting these coefficients back into the series summation gives $y_t = \sum_{m=-\infty}^{\infty} \frac{\Delta u}{2\pi}\int_{-\frac{1}{2}\tau}^{\frac{1}{2}\tau} y_t e^{-\mathrm{i}nt\Delta u}\ \mathrm{d}t\, e^{\mathrm{i}mt\Delta u}$. Suppose

$$\mathbf{y}(n\Delta u) = \int_{-\frac{1}{2}\tau}^{\frac{1}{2}\tau} y_t e^{-\mathrm{i}nt\Delta u}\mathrm{d}t,$$

then $y_t = \sum_{m=-\infty}^{\infty} \frac{\Delta u}{2\pi}\ \mathbf{y}(n\Delta u)\ e^{\mathrm{i}mt\Delta u}$. As $\tau \to \infty$ or $\Delta u \to 0$,

$$y_t = \frac{1}{2\pi}\int_{u=-\infty}^{\infty} \mathbf{y}(u)e^{\mathrm{i}tu}\mathrm{d}u.$$

This is regarded as the inverse relation of a very important transform in mathematics that is defined below. The term $\mathbf{y}(u)$ in the above development is the result of limiting $\Delta u \to 0$ in $\mathbf{y}(n\Delta u)$ and acquires the following definition:

Definition 2.14. ***Fourier transform*** *of a function $y_t\ \forall t \in \mathbb{R}$ satisfying (2.7) is*

$$\mathbf{y}(u) = \int_{-\infty}^{\infty} y_t e^{-\mathrm{i}tu}\mathrm{d}t. \tag{2.10}$$

2.3 Spectral analysis of discrete processes

In the Fourier transform relation of (2.10), a continuous function defined for $t \in \mathbb{R}$ is dealt with. Consider the continuous time function $y_t\ \forall t \in \mathbb{R}$ such that $y_t = 0$ whenever $t \neq m\Delta\tau\ \forall m \in \mathbb{Z}$ for some constant $\Delta\tau > 0$. This is equivalent to sampling the continuous function $y_t\ \forall t \in \mathbb{R}$ at discrete instants separated by $\Delta\tau$ and zero at all other instants. Since only discrete instants are relevant here from the Fourier transform perspective, y_t would be called a discrete time function. Therefore, the discrete time Fourier transform using (2.10) becomes

$$\mathbf{y}(u) = \sum_{m=-\infty}^{\infty} y_{m\Delta\tau} e^{-\mathrm{i}um\Delta\tau}.$$

Denote $y_m \triangleq y_{m\Delta\tau}$ and refer to it as the m-th sample. Then a sufficient condition for the existence of such a relation is $|\mathbf{y}(u)| < \infty$, i.e., $|\sum_{m=-\infty}^{\infty} y_m e^{-\mathrm{i}um\Delta\tau}| \leq \sum_{m=-\infty}^{\infty} |y_m| \, |e^{-\mathrm{i}um\Delta\tau}| < \infty$. This results in the absolute summability condition

$$\sum_{m=-\infty}^{\infty} |y_m| < \infty. \tag{2.11}$$

Using angular frequency ω as $u\Delta\tau = 2\pi\omega$ in the above development results in the following definition of Fourier transform for discrete time functions.

Definition 2.15. *The **discrete time Fourier transform** of a complex-valued discrete function $y_m \; \forall m \in \mathbb{Z}$ satisfying (2.11) is*

$$\mathbf{y}(\omega) = \sum_{m=-\infty}^{\infty} y_m e^{-\mathrm{i}2\pi\omega m}. \tag{2.12}$$

Although real-valued discrete functions were being discussed, the discrete time Fourier transform is valid for complex-valued functions also. Furthermore, since $e^{\mathrm{i}2\pi k} = 1 \, \forall k \in \mathbb{Z}$, the following property holds:

Property 2.3. *The discrete time Fourier transform has **unit periodicity**, i.e.,*

$$\mathbf{y}(\omega) = \mathbf{y}(k+\omega) \, \forall k \in \mathbb{Z}. \tag{2.13}$$

Another easily verifiable property, which holds true for any absolutely summable discrete or continuous function, is due to the following theorem; refer §22.1 of [41] or Chapter 3 of [72]:

Theorem 2.1. *According to the **Plancherel-Parseval theorem** for the discrete time Fourier transform $\mathbf{y}(\omega) \, \forall \omega \in [-\frac{1}{2}, \frac{1}{2}]$ of the function $y_m \, \forall m \in \mathbb{Z}$,*

$$\sum_{m\in\mathbb{Z}} |y_m|^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} |\mathbf{y}(\omega)|^2 \, \mathrm{d}\omega. \tag{2.14}$$

Just as the discrete time Fourier transform was defined being valid for complex-valued discrete functions, the Fourier series discussed earlier in (2.8) and (2.9) is applicable to any complex valued periodic function defined over any continuous domain. Hence, replacing $(y, -\frac{t}{\tau}) \to (\mathbf{y}, \omega)$ in (2.8) makes it equivalent to (2.12). In other words, the discrete time Fourier transform of a sequence of equally spaced samples of a real function is also a Fourier series whose coefficients form the sequence. Therefore, allowing the same replacement in (2.9) gives the differential $\frac{1}{\tau}\mathrm{d}t \to -\mathrm{d}\omega$ and the integral limits $t = \pm\frac{1}{2}\tau \to \omega = \mp\frac{1}{2}$ resulting in the following inverse of the relation in (2.12):

Definition 2.16. *The* ***inverse discrete time Fourier transform*** *of a complex-valued continuous function* $\mathbf{y}(\omega)\,\forall\omega \in \mathbb{R}$ *is defined as*

$$y_m = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{\mathrm{i}2\pi m\omega}\mathbf{y}(\omega)\mathrm{d}\omega. \tag{2.15}$$

The Fourier series gave discrete frequency components of a continuous time process and the discrete time Fourier transform gave continuous frequency components of a discrete time process. On the other hand, the following transform gives discrete frequency components of a finite realization of a discrete time process:

Definition 2.17. *The* ***discrete Fourier transform*** *of a series* $\{y_t\}$*,* $t = 1, \ldots, \tau$ *is defined as*

$$\mathbf{y}(\omega_j) = \frac{1}{\sqrt{\tau}}\sum_{t=1}^{\tau} y_t e^{-\mathrm{i}2\pi\omega_j t} \tag{2.16}$$

at discrete frequencies $\omega_j = \frac{j}{\tau}$*,* $j = 0, \ldots, \tau - 1$ *and the* ***inverse discrete Fourier transform*** *at the discrete instants is defined as*

$$y_t = \frac{1}{\sqrt{\tau}}\sum_{t=1}^{\tau} \mathbf{y}(\omega_j)e^{\mathrm{i}2\pi\omega_j t}. \tag{2.17}$$

The equivalent of Theorem 2.1 for the discrete Fourier transform is as follows [17]:

Property 2.4. *According to the* ***Plancherel-Parseval theorem*** *for the discrete Fourier transform* $\mathbf{y}(\omega_j)$ *of the sequence* $y_t\,\forall j, t = 1, \ldots, \tau$*,*

$$\sum_{t=1}^{\tau} |y_t|^2 = \sum_{j=1}^{\tau} |\mathbf{y}(\omega_j)|^2. \tag{2.18}$$

In this thesis, for a given finite length realization of a multivariate time series, certain asymptotic properties of the discrete Fourier transform will be used to define, derive, and optimize the dynamic transformation of the latent time series into commonalities. These asymptotic properties will be discussed in Section 2.5. The Plancherel-Parseval theorem will enable measuring and containing the commonalities that are retained during the transition between the time-domain and the frequency-domain. In Section 2.4, how the frequency-domain analysis finds utility in a stationary process will be discussed. Specifically, in Theorem 2.2, it will be learned how the Fourier transform relates two important statistical properties of a time series.

2.4 Spectral analysis of stationary processes

Suppose the pendulum motion expressed in (2.6) is subject to random amplitude α and phase ϕ disturbances so that (a, b) become uncorrelated zero-mean unit-variance random variables (a, b). Moreover, the discrete time domain is considered so that the equation of motion in (2.6) takes the form $\mathsf{y}_t = \mathsf{a}\cos(2\pi\omega t) + \mathsf{b}\sin(2\pi\omega t)\ \forall t \in \mathbb{Z}$; it will be called the '*perturbed pendulum.*' Its mean $\mu^{\mathsf{y}} = \mathsf{E}^{\mathsf{a},\mathsf{b}}[\mathsf{y}_t] = 0$. The *acvf* is $\gamma^{\mathsf{y}}_h = \mathsf{E}^{\mathsf{a},\mathsf{b}}[\mathsf{y}_{t+h}\mathsf{y}_t]$, which due to non-correlated a and b takes the form $\gamma^{\mathsf{y}}_h = \mathsf{E}^{\mathsf{a},\mathsf{b}}[\mathsf{a}^2\cos(2\pi\omega(t+h))\cos(2\pi\omega t)] + \mathsf{E}^{\mathsf{a},\mathsf{b}}[\mathsf{b}^2\sin(2\pi\omega(t+h))\sin(2\pi\omega t)]\ \forall h \in \mathbb{Z}$. Since it was assumed that $\mathsf{E}^{\mathsf{a}}[\mathsf{a}^2] = \mathsf{E}^{\mathsf{b}}[\mathsf{b}^2] = 1$, what one gets[1] is $\gamma^{\mathsf{y}}_h = \cos(2\pi\omega h)$; spectral analysis of such an *acvf* could be found in [111, 19]. Since μ^{y} and γ^{y}_h are independent of t, it is found that $\{\mathsf{y}_t\}$ is weakly stationary. And, since weakly stationary $\{\mathsf{y}_t\}$ does not satisfy (2.11), its Fourier transform simply does not exist.

Using Euler's identity, the *acvf* of the perturbed pendulum is written as a summation $\gamma^{\mathsf{y}}_h = \sum_{i=1}^{k} \alpha_i g(\omega_i)$, where $g(\omega) = e^{\mathrm{i}2\pi\omega h}$, $k = 2$, $\omega_1 = -\omega$, $\omega_2 = \omega$, and $\alpha_1 = \alpha_2 = \frac{1}{2}$. But such a summation with a general $g(\omega)$ has an integral representation $\sum_{i=1}^{k} \alpha_i g(\omega_i) = \int g(\omega)\ \mathrm{d}\mathfrak{S}^{\mathsf{y}}(\omega)$, where $\mathfrak{S}^{\mathsf{y}}(\omega) \triangleq \sum_{i=1}^{k} \alpha_i\, 1(\omega_i \leq \omega)$ is a monotonically increasing function bounded between $\mathfrak{S}^{\mathsf{y}}(-\infty) = 0$ and $\mathfrak{S}^{\mathsf{y}}(\infty) = 1$, and $1(\omega_i \leq \omega)$ is the step function which jumps from zero to unity at $\omega = \omega_i$.

However, due to periodicity of $g(\omega) = e^{\mathrm{i}2\pi\omega h}$ in the above example of a perturbed pendulum, the *acvf* is essentially represented in the integral form $\gamma^{\mathsf{y}}_h = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{\mathrm{i}2\pi\omega h}$ $\mathrm{d}\mathfrak{S}^{\mathsf{y}}(\omega)$, where $\mathfrak{S}^{\mathsf{y}}(\omega)$ is a monotonically increasing function bounded between $[-\frac{1}{2}, \frac{1}{2}]$ while $\mathfrak{S}^{\mathsf{y}}(-\frac{1}{2}) = 0$ and $\mathfrak{S}^{\mathsf{y}}(\frac{1}{2}) = \gamma^{\mathsf{y}}_0$. The reader is referred to [99, 20] for the details of this representation and other properties that $\mathfrak{S}^{\mathsf{y}}(\omega)$ adheres to. The notion carried forward is that whenever the derivate $\mathsf{s}^{\mathsf{y}}(\omega) = \frac{\mathrm{d}}{\mathrm{d}\omega}\mathfrak{S}^{\mathsf{y}}(\omega)$ exists, it is possible to write the *acvf* as

$$\gamma^{\mathsf{y}}_h = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{\mathrm{i}2\pi\omega h}\ \mathsf{s}^{\mathsf{y}}(\omega)\,\mathrm{d}\omega. \tag{2.19}$$

But if there are discontinuities in $\mathfrak{S}^{\mathsf{y}}(\omega)$, e.g., the perturbed pendulum, it will not be possible to write the *acvf* according to (2.19) because $\mathsf{s}^{\mathsf{y}}(\omega)$ does not exist.

Now refer back to (2.15) to see its analogy with (2.19) which requires that a condition

$$\sum_{h=-\infty}^{\infty} |\gamma^{\mathsf{y}}_h| < \infty, \tag{2.20}$$

equivalent to (2.11) be satisfied by γ^{y}_h. This enables the following theorem and definition; refer §4.3 of [20]:

[1] Using the trigonometric identity $\cos(\theta_1 - \theta_2) = \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2$

Theorem 2.2. *For acvf γ_h^{y} of a weakly stationary time series $\{\mathsf{y}_t\}\ \forall t \in \mathbb{Z}$ satisfying (2.20), according to Herglotz's theorem, the **spectral density function** $\mathsf{s}^{\mathsf{y}}(\omega)$ at the frequency $\omega \in \mathbb{R}$ exists and is defined as the discrete time Fourier transform of its acvf, i.e.,*

$$\mathsf{s}^{\mathsf{y}}(\omega) \triangleq \sum_{h=-\infty}^{\infty} \gamma_h^{\mathsf{y}} e^{-\mathrm{i}2\pi\omega h}. \tag{2.21}$$

In this context, it may be noted that the perturbed pendulum does not satisfy the absolute sum condition because $\sum_{h=-\infty}^{\infty} |\gamma_h^{\mathsf{y}}| = \sum_{h=-\infty}^{\infty} |\cos(2\pi f h)| = \infty$ and it does not have a spectral density function.
In light of (2.21) and similar to Property 2.3, the following property of the spectral density function is arrived at:

Property 2.5. *The spectral density function has **unit periodicity**, i.e., $\mathsf{s}^{\mathsf{y}}(\omega) = \mathsf{s}^{\mathsf{y}}(k+\omega)\,\forall k \in \mathbb{Z}$.*

For an r-variate time series $\{\mathsf{y}_t\}\ \forall t \in \mathbb{Z}$, where $\mathsf{y}_t = [\mathsf{y}_{1_t} \cdots \mathsf{y}_{r_t}]'$, let $\gamma_h^{\mathsf{y}_i,\mathsf{y}_j}$ be the (i,j)-th element of its autocovariance matrix Γ_h^{y}. Referring to §1.1.2 of [102], the condition equivalent to (2.20) for vector random variables becomes

$$\sum_{h=-\infty}^{\infty} |\gamma_h^{\mathsf{y}_i,\mathsf{y}_j}| < \infty, \tag{2.22}$$

which is valid $\forall i,j \in 1,\ldots,r$ in defining the matrix of spectral density function $\mathsf{S}^{\mathsf{y}}(\omega) \in \mathbb{C}^{r\times r}$ whose (i,j)-th element is $\mathsf{s}^{\mathsf{y}_i,\mathsf{y}_j}$. Then, due to the development of Property 2.1 and the relation (2.21), the following is easily got:

Property 2.6. *The spectral density function $\mathsf{S}^{\mathsf{y}}(\omega)$ is **Hermitian symmetric** about $\omega = 0$, i.e.,*

$$(\mathsf{S}^{\mathsf{y}}(\omega))' = \mathsf{S}^{\mathsf{y}}(-\omega) = \overline{\mathsf{S}}^{\mathsf{y}}(\omega). \tag{2.23}$$

Referring to Theorem 2.7.1 of [18], Theorem 4.4.1 of [39], and [92, 34], another important property of the r-variate time series follows:

Property 2.7. *If $\{\mathsf{y}_t\}\ \forall t \in \mathbb{Z}$ is a linear process, then $|\mathsf{S}^{\mathsf{y}}(\omega)| \neq 0\ \forall \omega \in [0,1]$.*

The above discussion is very relevant to the intention in this thesis to assess the commonalities of a multivariate time series via its spectral density function. For the purpose of learning multivariate time series based on the commonalities, the hope is to take the following approaches: Firstly, two multivariate time series are compared by evaluating how similar the components of their spectral density functions are. Secondly, the future evolution of a multivariate time series is predicted by estimating the *acvf*, via its spectral density function, that inherits maximum commonalities.

2.5 Asymptotic properties of linear processes

In practical problems, it is infeasible to have a dataset consisting of an infinite collection of samples to compute true statistical properties such as mean, *acvf*, variance, etc. Characteristics of a time series have to be estimated from a given finite length subset of its realization. With a limited number of samples, sample estimates may be questioned for their reliability. The field of study of asymptotic statistics strives to design properties, procedures, tests, and estimators in the limit that the sample size becomes large [122, 58]. A broad review of the asymptotic techniques will not be resorted to; however, presented below are the essentials for the thesis's purposes.

Consider the scenario in which, due to computational or limited access to data, time series characteristics have to be gleaned from one realization forming a finite length data stream. For a weakly stationary time series, these characteristics include its mean and *acvf* for which, first, the following asymptotic properties referring to §11.2 of [20] are presented:

Theorem 2.3. *For a finite τ-length realization $\{y_t\}$ $t = 1, \ldots, \tau$ of a weakly stationary time series $\{\mathsf{y}_t\}$ $t \in \mathbb{Z}$ whose acvf Γ^{y}_h satisfies (2.22), the **sample mean***

$$\hat{y} = \frac{1}{\tau} \sum_{t=1}^{\tau} y_t \tag{2.24}$$

converges in a mean square sense to the population mean μ^{y}.

Theorem 2.4. *For a finite τ-length realization $\{y_t\}$ $t = 1, \ldots, \tau$ of a weakly stationary time series $\{\mathsf{y}_t\}$ $t \in \mathbb{Z}$ with sample mean $\hat{y}$, the $r \times r$ **sample acvf***

$$\widehat{\Gamma}^{\mathsf{y}}_h = \begin{cases} \frac{1}{\tau} \sum_{t=1}^{\tau-h} (y_{t+h} - \hat{y})(y_{t+h} - \hat{y})' & 0 \leq h \leq \tau - 1, \\ \frac{1}{\tau} \sum_{t=-h+1}^{\tau} (y_{t+h} - \hat{y})(y_{t+h} - \hat{y})' & -\tau + 1 \leq h < 0 \end{cases} \tag{2.25}$$

converges in probability to the population acvf Γ^{y}_h.

With the sample *acvf* $\widehat{\Gamma}^{\mathsf{y}}_h$ for finite lags, the best hope is for estimates of the spectral density function $\mathsf{S}^{\mathsf{y}}(\omega_k)$ at finite discrete frequencies $\omega_k = \frac{k}{\tau}$ $\forall |k| = 0, \ldots, \tau - 1$ via inverse discrete Fourier transform. For an otherwise continuous spectral density function $\mathsf{S}^{\mathsf{y}}(\omega)$, $0 \leq \omega < 1$, those estimates at discrete frequencies is an approximation of $\mathsf{S}^{\mathsf{y}}(\omega_k)$ dependent on how good the sample estimation $\widehat{\Gamma}^{\mathsf{y}}_h$ is. Therefore, in what follows, described is the asymptotic property of $\mathsf{S}^{\mathsf{y}}(\omega)$ near any target frequency $\omega_j = j/\hat{\jmath}$ $\forall j = 0, \ldots, \hat{\jmath} - 1$, or $0 \leq \omega_j < 1$ and $\hat{\jmath} \ll \tau$.

It starts by splitting a period of $\omega \in [0, 1)$ of the spectral density function into $\hat{\jmath}$ non-overlapping frequency bands. Suppose there is a total of $\tau = n\hat{\jmath}$ discrete frequencies that are considered for the splitting so that each band will have n discrete frequencies. By the 0-th frequency band represented by the target frequency $\omega_0 = 0$, implied are n discrete frequencies $\omega_{0,l} > 0$, $l = 1, \ldots, n$ closest to 0. By the j-th frequency band $\omega_{j,l}$ $\forall l = 1, \ldots, n$; $j = 1, \ldots, \hat{\jmath} - 1$, implied are n frequencies closest to the target frequency $\omega_j = j/\hat{\jmath}$ and between $\omega_j - b$ and $\omega_j + b$, where $2b = n/\hat{\jmath}$ is called the

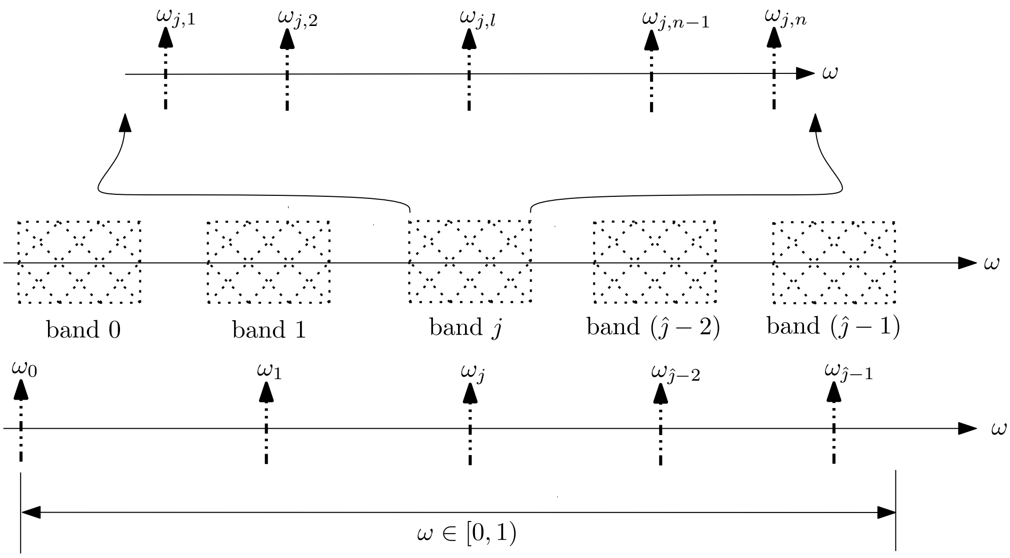

Figure 2.4: The scheme of splitting the frequency range $\omega = [0,1)$ into $\hat{\jmath}$ non-overlapping subbands containing n discrete frequency components each.

bandwidth. Suppose $2b < \omega_1 - b < \omega_{\hat{\jmath}-1} + b < 1$ and $n \ll \hat{\jmath}$ is choosen so that the bandwidth is very low.

For proceeding further, the following definitions are needed; refer to [45]:

Definition 2.18. *An r-dimensional '**complex-valued vector random variable**' $\xi = [\xi_1 \cdots \xi_r]' = \Re(\xi) + \mathrm{i}\Im(\xi) \in \mathbb{C}^r$ is defined as the $2r$-variate vector random variable $\eta = [\Re(\xi_1)\,\Im(\xi_1) \cdots \Re(\xi_r)\,\Im(\xi_r)]' \in \mathbb{R}^{2r}$ formed by its real and imaginary components.*

As established in [45], the covariance matrix Γ^ξ of an r-variate complex valued vector random variable ξ is isomorphic, i.e., equivalent upto a row and a column permutation, to the covariance matrix Γ^η of its corresponding $2r$ component vector random variable η via

$$\Gamma^\xi \approxeq 2\Gamma^\eta;\ (\Gamma^\xi)^{-1} \approxeq \frac{1}{2}(\Gamma^\eta)^{-1};$$

whereas the means are isomorphic via $\mu^\xi \approxeq \mu^\eta$. Also, it was shown there that $\det(\Gamma^\xi) = 2^r(\det(\Gamma^\eta))^{\frac{1}{2}}$ and $\xi^*\Gamma^\xi\xi = \eta'\Gamma^\eta\eta$. Then, following the convention of a Gaussian distribution of an r-variate random variable u with mean a and covariance matrix B denoted by

$$\mathcal{N}(\mathrm{u}|a,B) = \frac{\exp\left(-\frac{1}{2}(\mathrm{u}-a)'B^{-1}(\mathrm{u}-a)\right)}{(2\pi)^r\ (\det(B))^{\frac{1}{2}}}, \tag{2.26}$$

the following definition could be arrived at:

Definition 2.19. *The r-variate '**complex Gaussian probability density**' of a complex valued random variable* u *with mean $a \in \mathbb{C}^r$ and covariance matrix $B \in \mathbb{C}^{r\times r}$ is defined as*

$$\mathcal{N}_{\mathbb{C}}(\mathsf{u} \mid a, B) = \frac{\exp\left(-(\mathsf{u}-a)^* B^{-1}(\mathsf{u}-a)\right)}{\pi^r \det(B)}. \tag{2.27}$$

Now an essential theorem for this thesis is presented; refer Theorem 4.4.1 of [18], §C.2 of [111], and [53].

Theorem 2.5. *The discrete Fourier transform components of a realization of an r-variate linear process at frequencies $\omega_{j,l}$ such that $\lim_{\hat{\jmath}\to\infty} |\omega_j - \omega_{j,l}| \to 0\ \forall l \in 1,\dots,n$; $\forall j = 1,\dots,\hat{\jmath} \gg n$ are iid samples of an r-dimensional 'complex-valued vector random variable'* y_j *at frequency $\omega_j \in [0,1]$ with a probability density*

$$\mathsf{p}^{\mathsf{y}(\omega_j)}(u) = \begin{cases} \mathcal{N}_{\mathbb{C}}(u \mid 0, \mathsf{S}^{\mathsf{y}}(\omega_j)), & \omega_j \in (0,1) \\ \mathcal{N}(u \mid 0, 2\,\mathsf{S}^{\mathsf{y}}(\omega_j)), & \omega_j \in \{0,1\}. \end{cases} \tag{2.28}$$

Theorem 2.5 furnishes a probabilistic model for discrete Fourier transform samples obtained from a finite realization of a time series of a linear process. The theorem simply recommends that discrete Fourier transform components within a 'small' bandwidth near a target frequency ω_j is Gaussian with the covariance matrix equal to the spectral density function $\mathsf{S}^{\mathsf{y}}(\omega_j)$; at zero frequency the covariance matrix is twice $\mathsf{S}^{\mathsf{y}}(0)$.

In order to use this theorem, the following procedure is adhered to: Given τ samples of a time series realization, first compute the τ-length discrete Fourier transform $\mathbf{y}(\omega_k)$, $k = 0,\dots,\tau-1$. Then, n discrete Fourier transform components $\mathbf{y}(\omega_{j,l})$, $l = 1,\dots,n$; $j = 0,1,\dots,\hat{\jmath}-1$, contained in the j-th subband may be assigned as

$$\omega_{j,l} : |\omega_j - \omega_{j,l}| \le n\tau^{-1}; \quad n \ll \hat{\jmath}. \tag{2.29}$$

For the j-th frequency band, the sample covariance matrix is computed as

$$\widehat{\mathsf{S}}^{\mathsf{y}}(\omega_j) = \frac{1}{n}\sum_{l=1}^{n}(\mathbf{y}(\omega_{j,l}) - \hat{\mathbf{y}}(\omega_j))(\mathbf{y}(\omega_{j,l}) - \hat{\mathbf{y}}(\omega_j))^*, \tag{2.30}$$

and

$$\hat{\mathbf{y}}(\omega_j) = \frac{1}{n}\sum_{l=1}^{n}\mathbf{y}(\omega_{j,l})$$

is the sample mean of the discrete Fourier transform $\mathbf{y}(\omega_k)$ and $\omega_j - b < \omega_k < \omega_j + b$. To ensure robustness of the estimate $\widehat{\mathsf{S}}^{\mathsf{y}}(\omega_j)$, typically, one would also want to maintain

$$n \ge r^2, \tag{2.31}$$

refer [106, 12].

It could be shown, as done in §4.2 of [18] or §12.4 of [25] that for a linear process

$$\lim_{n\to\infty} \mathsf{E}^{\mathsf{y}}[\widehat{\mathsf{S}}^{\mathsf{y}}(\omega \mid n)] = \mathsf{S}^{\mathsf{y}}(\omega) \quad \forall\omega \in [0,1). \tag{2.32}$$

Hence, while maintaining $\hat{\jmath} \gg n$, a sufficiently large n should provide an unbiased estimate of $\mathsf{S}^{\mathsf{y}}(\omega_j)$ through (2.30). This process is given in Algorithm 1, where care should be taken to ensure that there are sufficient subbands as required by Theorem 2.5.

Algorithm 1: Prepare discrete Fourier transform subbands

Input: $\mathcal{D} = \{y_t\}, t = 1, \ldots, \tau;\ y_t \in \mathbb{R}^r; n; \hat{\jmath};$
Output: $\{\widehat{\mathsf{S}}^{\mathsf{y}}(\omega_j)\};\ \{\mathbf{y}(\omega_{j,l})\};\ j = 1, \ldots, \hat{\jmath};\ l = 1, \ldots, n;$
compute $y_t \xleftrightarrow{\mathcal{F}} \mathbf{y}(\omega_k);\ \omega_k = \frac{k}{\tau},\ k = 0, \ldots, \tau - 1$ using (2.16);
assign $\mathbf{y}(\omega_{j,l});\ j = 1, \ldots, \hat{\jmath};\ l = 1, \ldots, n$ using (2.29);
estimate $\widehat{\mathsf{S}}^{\mathsf{y}}(\omega_j)$ using (2.30);

2.6 Summary

This chapter introduced certain frequently sought after notions pertaining to time and frequency domain analyses of time series. These notions include *acvf*, spectral density function, discrete Fourier transform, white noise, etc. The relation between the spectral density function and the *acvf* was recapped on. Also introduced were some of the notations adhered to for the remaining chapters. As discussed, the spectral density function of a stationary time series is the Fourier transform of its autocovariance function. The discrete Fourier transform components of a linear process within a small bandwidth around a target frequency is approximately complex-Gaussian with mean zero and covariance matrix equaling the spectral density function at the target frequency.

Our goal for this thesis is to model and learn a measured multivariate time series by dynamically transforming a low-dimensional latent time series. The hope is to use classical probabilistic modeling concepts introduced in the next chapter to achieve this goal. Most of those concepts will be based on fitting popular probability density function models on time and lag independent data; but it is time series data that is dealt with. In order to elicit a similar and manageable probability density function that applies to a wide class of time series, the asymptotic theory of discrete Fourier transform components was approached. This is because those components within a small bandwidth may be considered as realizations of a complex-valued Gaussian vector random variable. This enables the possibility of applying standard probabilistic modeling techniques, as reviewed in the next chapter, to multivariate time series modeling.