Modelling Multiple and Structured Oscillatory Phenomena

A thesis presented for the degree of

Master of Philosophy of the University College London

by

Marko Javorsek

Department of Statistical Science

University College London

Gower Street, London WC1E 6BT

SEPTEMBER, 2015

I, Marko Javorsek confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Abstract

This work deals with the modelling of multiple and structured oscillatory phenomena. The goal of the thesis is to show how stochastic oscillations can be modelled, and define their elliptical structures as a special class of bivariate time-dependant variation. The central part of the research is the introduction of new multivariate elliptical models and the review of existing definitions. The findings are presented in a table, where the classification is made based on whether the definitions are random or deterministic and whether they are defined in time or frequency domains. The previously introduced ellipse definitions for stochastic processes that have been described in the literature are limited to the frequency domain only. The main contribution of this work is in adding to existing time domain models by defining the description of the autocovariance ellipse and the forecast ellipse. Both of these definitions are non-random. The ellipses are defined from either the autocovariance or the forecast functions of the process as one moves forward in lag-time or forecast-time. In order to illustrate these theoretical concepts and show the usefulness of the new definition we investigate these concepts using a parametric model. Univariate and bivariate, real-valued and complex-valued models are considered, and their representation discussed. The richest model proposed is that of a complex-valued bivariate autoregressive process of order one and this is based on modelling using affine transformation matrices. This model results in a stochastic oscillation and the elliptical definitions proposed are explored in this context. The actual behaviour of the proposed stochastic process is also illustrated on simulated data. Some limitations of this approach are discussed and extensions of this model are presented.

Table of contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis introduces new time series models in the realm of multiple stochastic oscillations. We shall unify understanding of such processes from the fields of statistics, signal processing and optics. We will propose new models and develop new understanding of these models that describe the notion of an elliptical oscillation. Therefore the new contributions are: (1) clarifying and summarising the existing elliptical representations; (2) creating a superstructure for describing multivariate elliptical structure, in particular by introducing two new elliptical representations; and (3) introducing a bivariate complex-valued vector autoregressive $\mathbb{C}\mathsf{VAR}(1)$ model and studying its properties in the time and frequency domains.

The classical notion of a time series model is as an aggregation of components. According to Harvey [19] a structural time series model (STSM) is defined as a simple additive model of various components, most notably a trend component, a cyclical component and a seasonal component and can be adapted for both univariate or multivariate data. Such models have been proposed by many authors especially in the field of econometrics and finance (e.g. [19, 21, 54]), but similar models can be found in other applications as well, for example oceanography [13].

One example of such models is defined as an addition of four vector components: the trend component $\boldsymbol{\mu}_t$, the seasonal component $\boldsymbol{\gamma}_t$, and the cyclical component $\boldsymbol{\psi}_t$, and the noise component $\boldsymbol{\epsilon}_t \sim$ i.i.d.$N(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}})$. As such, it is in its multivariate form defined for a vector of n time series, $\mathbf{X}_t = [X_1^{(1)}, \ldots, X_t^{(n)}]^{\mathrm{T}}$. The general form of an STSM for multiple time series is [19]

$$\mathbf{X}_t = \boldsymbol{\mu}_t + \boldsymbol{\gamma}_t + \boldsymbol{\psi}_t + \boldsymbol{\epsilon}_t, \qquad t = 1, \ldots, T. \tag{1.0.1}$$

Generally, it is hard to specify the difference between the seasonal and cyclical components as both are periodic. In this research we understand the seasonal component to have a period governed by the law of nature (e.g. days in a week, months or quarters in a year, etc.), whereas the periodicity of the cyclical component is not constrained to any value. However, in this thesis we are mainly interested in investigating the cyclical component, $\boldsymbol{\psi}_t$, and aim to explain how to model this as a stochastic oscillation. For example, the cyclical component modelling is widely used for modelling of econometric data sets, where one is interested in economic business cycles (e.g. [22, 54, 26]), or in oceanography for the analysis of the oscillatory force of the oceans (e.g. [13]).

We are not only interested in how these cycles move univariately, but how they move together in relation to each other in a multivariate case. This clearly shows the need and logical extension to bivariate and multivariate models, as it allows us to model two or more oscillations in the same model. We wish to describe their cyclical commonality (i.e. frequency of oscillation) and understand the relation between them, whether one is leading the other and what is the phase difference between them.

Elliptical time series is a special class of bivariate time dependent structure and has been used in the physical sciences to understand problems in geophysics and oceanography [6, 13]. Recent developments have formalized this class of signals in the frequency domain [53, 59], advocating the joint analysis using a complex-valued representation of the signal. However, the temporal understanding of such signals is lagging far behind the frequency domain understanding. Here we aim to clarify what an elliptical time series corresponds to, determine its time domain properties and introduce a new family of parametric time series where this structure is important and interpretable.

Correlated bivariate time series are ubiquitous in physical applications. Recent years have seen the development of methods that are based on modelling bivariate time series as univariate complex-valued time series. Several scientific communities use such models to understand oscillations and elliptical signals, especially for inference methods. We can find oscillatory data sets in various applications, such as electromagnetic radiation [71], seismology [55], oceanography [29], econometrics [19] and blood flow [39]. There are also different types of oscillations, the ones with time-varying period are fundamental in neuroscience [58], the physical sciences [13, 30], and econometrics [19, 54]. There are two important classes of models, the time-

varying deterministic oscillation in noise [9] and the stochastic oscillation, such as an autoregressive model AR(2) with complex roots [49].

Many would argue that most of the quantities observed in applications are real-valued and question the need for developing models for complex-valued time series. We argue that complex-valued data are naturally occurring and can be found in many data acquisition settings. It is worth mentioning that in this thesis we are dealing with complex-valued time series, which are not to be understood in the same way as the complex-valued random vectors (RVs). Complex-valued time series are not just a concatenated version of complex-valued RVs, because in the case of time series we are interested in their time evolution and relationship between time points. For time series the direction of this evolution is also important and needs to be kept in consideration, as the natural time flows in one direction.

One of the main advantages of using complex-valued time series is that the relationship between the two series is preserved and analysed when modelled as complex-valued data by using the real and imaginary components. It means that we can convert a bivariate real-valued quantity into a univariate complex-valued quantity, which simplifies mathematical calculations while still preserving the insight into the relationships between the components of the data [62]. This is especially convenient if the quantities come in pairs. In other words, if they are naturally two-dimensional quantities that lie in the plane as orthogonal components, e.g. north-south and east-west components in case of oceanographic data (see [29, 30, 68]). In many sciences, such as electro-magnetics, oceanography and atmospheric science, pairs of real-valued signals coexist together and the researcher is interested in their trajectories in a plane. In the electromagnetic theory the time-varying position of the electric field vector can be viewed as a complex-valued signal in the (u, v)-plane [62, p. 6]. Another way to obtain complex-valued time series is to artificially create them from real-valued data by creating complex analytical signals using the Hilbert transforms.

The simplest way to think about an elliptical signal is to imagine an elliptical trajectory mapped out by the signal over time. Several authors have been analysing elliptical signals and aiming to define what is *the 'ellipse'* of a signal. Early developments have been made for deterministic signals in the optics community where light is observed as two parts of a signal that together map a trajectory of an ellipse [6]. More recent work on the elliptical structure of stochastic processes has been done in the signal processing community [61] as well as the statistical community [52].

In the case of stochastic processes time trajectories are random, so the frequency domain representations have been used to define the elliptical structure. To add to the variety of available definitions, there are also other concepts such as modulated elliptical signals [30], multivariate oscillations [31] and polarisation [24, 56]. Some of these concepts are defined in the time domain of the signals, some in frequency domain, but there is no effective summary or relating the different definitions.

The engineering and signal processing communities are more familiar with analysing a signal in the frequency domain, whereas the statistics and econometrics communities prefer to operate in the time domain. We aim to unify the two viewpoints and clearly define how an elliptical time series can be viewed in both domains. One of the ways to achieve this is by defining the ellipse of a stochastic process in time domain rather than frequency domain. The contribution of our work is the view on an elliptical process through the so-called forecast ellipse and autocovariance ellipse, as they will be defined. Due to the properties of these two qualities the ellipses will be defined in time domain and will be deterministic representations of a stochastic process.

In Chapter 5 of this thesis we aim to summarise the different approaches to defining the ellipse in one coherent framework and propose two new elliptical definitions. However, we start by reviewing the fundamentals of complex-valued random vectors and random processes in Chapter 2, and general theory about time series in Chapter 3. The initial literature review is intended to clarify and emphasize the main concepts needed to understand the differences when dealing with complex-valued signals and processes, and to be able to easily understand the application of these concepts.

In order to illustrate the concepts developed we will present some parametric elliptical models in Chapter 6. On one hand these models should serve as tools to understand the concepts developed in the previous chapters and on the other hand, as general parametric models that can be used to model stochastic oscillations. Parametric models are in general very useful, because they can be easily estimated and have parameters that are interpretable. We start the parametric models chapter by looking at a univariate complex-valued $\mathbb{C}\mathsf{AR}(1)$ and continue by defining a bivariate complex-valued $\mathbb{C}\mathsf{VAR}(1)$ that will be our main focus. At the end we also explore extensions to the bivariate model by introducing a stretching factor. The parametric models are illustrated with simulated examples, which we also use to show the concepts of ellipses developed theoretically.

Chapter 2

Complex random vectors and processes

We shall start by presenting some background theory and main concepts about complex-valued random vectors (RV) and random processes (RP). In many aspects and modelling structure they are no different from real-valued ones, but there is an additional level of complexity that needs to be clear to the reader. In this chapter we introduce complex-valued random vectors and processes and discuss their properties. This description is important, because these concepts will be widely used in the rest of our work and most of us are more used to deal with real-valued random vectors and processes. Since all our work is based on complex-valued processes, it is very important that these concepts are clarified at the beginning.

Mathematically the complex-valued random vector could always be analysed as a pair of real vectors. Although using complex RVs compared to a pair of real RVs has several advantages, mainly because the derivations are simpler and at the same time preserving the physical sense related to the nature of the data [3]. Authors have been arguing whether complex RVs are special or not and whether they can be equally well analysed with bi-variate real random vectors (see for example the discussion in [23, 43]).

2.1 Complex-valued random vectors

In general, complex-valued random-vectors (RVs) can be seen as two real-valued RVs, we provide the definition below. The mathematical theory of complex numbers

will be needed. A complex number can be expressed as $a + bi$, where a is the real part and b is the imaginary part. The imaginary part is multiplied by the imaginary unit i, such that $i^2 = -1$.

Definition 2.1.1 (Complex-valued random vector [3])
The n-dimensional complex-valued RV, $\mathbf{U} \in \mathbb{C}^n$*, is composed of two real-valued RVs,* $\mathbf{X} \in \mathbb{R}^n$ *and* $\mathbf{Y} \in \mathbb{R}^n$*, in the following way,* $\mathbf{U} = \mathbf{X} + i\mathbf{Y}$*.*

In other words, $\mathbf{X} = \Re\{\mathbf{U}\}$ and $\mathbf{Y} = \Im\{\mathbf{U}\}$, where "$\Re$" denotes the real part and "$\Im$" denotes the imaginary part. The complex conjugate of the RV is denoted as $\mathbf{U}^* = \mathbf{X} - i\mathbf{Y}$. It is worth clarifying the notation at this point, we are using capital letters (e.g. $\mathbf{U}$) for random quantities, whereas non-random vectors (such as realisations) will be denoted with small letters, for example $\mathbf{u} = \mathbf{x} + i\mathbf{y}$.

Composite and augmented vecotrs For the purposes of analysis of complex RVs we need to define additional concepts. Based on various literature (e.g. [2, 44, 61]) we define two concepts: (1) the *real composite vector* $\mathbf{V}$, a vector composed of the real and imaginary parts of a complex RV as two real-valued RVs, $\mathbf{V} = [\mathbf{X}^{\mathrm{T}}, \mathbf{Y}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2n}$; and (2) the *complex augmented vector* $\mathbf{W}$, a vector composed of the complex RV and its complex conjugate, $\mathbf{W} = [\mathbf{U}^{\mathrm{T}}, \mathbf{U}^{\mathrm{H}}]^{\mathrm{T}} \in \mathbb{C}^{2n}$. The superscript $^{\mathrm{T}}$ denotes vector transpose and superscript $^{\mathrm{H}}$ denotes conjugate transpose. The algebraic relationships between these sets of vectors are outlined below, along with some useful matrices and concepts that will be used in the continuation. These concepts are needed for the complete picture about complex-valued RVs we need to analyse both the vector itself and its complex conjugate.

Real-to-complex transformation To transform vectors from real composite to complex augmented ones it is useful to introduce a $2n \times 2n$ real-to-complex transformation matrix denoted $\mathbb{T}_n$ [61], as

$$\mathbb{T}_n = \begin{bmatrix} \mathbf{I}_n & i\mathbf{I}_n \\ \mathbf{I}_n & -i\mathbf{I}_n \end{bmatrix}. \tag{2.1.1}$$

The transformation matrix has the following property, $\mathbb{T}_n\mathbb{T}_n^{\mathrm{H}} = \mathbb{T}_n^{\mathrm{H}}\mathbb{T}_n = 2\mathbf{I}_{2n}$, where $\mathbf{I}_n$ is the $n \times n$ identity matrix. This matrix is useful as it transforms the real composite vector of two real-valued RVs into a complex augmented vector of a

single complex-valued RV, whose real and imaginary parts will be the entries of the initial composite RV. Hence, by using the real-to-complex transformation matrix and the composite and augmented vectors mentioned above, we can compose a complex-valued RV $\mathbf{U}$ from two real-valued RVs $\mathbf{X}$ and $\mathbf{Y}$, as follows

$$\mathbf{W} = \begin{bmatrix} \mathbf{U} \\ \mathbf{U}^* \end{bmatrix} = \begin{bmatrix} \mathbf{X} + i\mathbf{Y} \\ \mathbf{X} - i\mathbf{Y} \end{bmatrix} = \mathbb{T}_n \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \mathbb{T}_n \mathbf{V}. \tag{2.1.2}$$

Rotation matrix We introduce the rotation matrix, which will be used in vector valued time series. The rotation matrix $\mathbf{R}_\theta$ rotates the vector by angle θ in the clockwise direction (CW). It is defined as

$$\mathbf{R}_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \tag{2.1.3}$$

and as such is an orthogonal matrix, $\mathbf{R}_\theta^{\mathrm{T}} = \mathbf{R}_\theta^{-1}$ and $\mathbf{R}_\theta \mathbf{R}_\theta^{\mathrm{T}} = \mathbf{I}_2$, with the following properties $\det\{\mathbf{R}_\theta\} = 1$ and $\mathbf{R}_\theta \mathbf{R}_\lambda = \mathbf{R}_{\theta+\lambda}$.

Phase shift matrix A complex-valued phase shift matrix acts as a unitary transformation that shifts the phase of the entries of the vector it operates on by δ in the opposite directions. It can be specified as

$$\mathbf{K}_\delta = \begin{bmatrix} e^{i\delta} & 0 \\ 0 & e^{-i\delta} \end{bmatrix}, \tag{2.1.4}$$

where $\mathbf{K}_\delta \mathbf{K}_\delta^{\mathrm{H}} = \mathbf{I}_2$ and $\mathbf{K}_\delta \mathbf{K}_\nu = \mathbf{K}_{\delta+\nu}$. For example, in time domain it shifts the components of the signal out of phase by 2δ. The effect of a phase shift matrix can be illustrated with an arbitrary complex-valued vector expressed in polar form. The matrix that acts on the vector creates a phase shifted vector as

$$\begin{bmatrix} e^{i\delta} & 0 \\ 0 & e^{-i\delta} \end{bmatrix} \begin{bmatrix} r_1 e^{i\phi_1} \\ r_2 e^{i\phi_2} \end{bmatrix} = \begin{bmatrix} r_1 e^{i(\phi_1+\delta)} \\ r_2 e^{i(\phi_2-\delta)} \end{bmatrix}. \tag{2.1.5}$$

The distribution of complex random vectors

There are many different views on how to understand the distribution of complex RVs and what is needed for the complete specification of their distribution. Picinbono [44] emphasizes that a complex RV, $\mathbf{U} \in \mathbb{C}^n$, can be simply seen as a com-

bination of two real RVs, $\mathbf{X}$ and $\mathbf{Y}$, and as such analysed as two real RVs in $\mathbb{R}^{2n}$. In case of normal distribution, he claims that the complex RV, $\mathbf{U}$, will be normally distributed, if the real and imaginary parts will be jointly normally distributed. He as well concludes that if two complex normal RVs are uncorrelated, they are not necessarily independent, as it would be the case of real random variables.

Whereas, Amblard et al. [3] identify that one must consider both $\mathbf{U}$ and $\mathbf{U}^*$ in order to capture all statistical information contained in the PDF. This is disputed by Olhede [40] as $\mathbf{U}^*$ is only the complex-conjugate of the complex RV, and so it has similar mathematical properties. Two related density functions are possible, but only one can have a proper meaning of a density. The second one just explains the role played by the complex conjugate, which is what has created the most disagreement between the authors in the literature.

Dependence between complex random variables

For the analysis of the dependence between complex RVs, the concepts used for real-valued RVs cannot be directly applied, as one needs to take into consideration the covariances between the real and imaginary parts. In the contrary, the complete second-order statistical information would not be captured. Many authors have recently looked at the definitions of the covariance matrix for complex-valued RVs and have identified that special care needs to be taken (typical examples of such developments include [37, 38, 44, 60, 61]).

The covariance matrix of zero-mean real-valued random vector, $\mathbf{X}$, is usually specified as $\mathbf{\Gamma}_X = \mathrm{cov}(\mathbf{X}, \mathbf{X}) = \mathbb{E}[\mathbf{X}\mathbf{X}^{\mathrm{T}}]$. The covariance matrix for complex-valued RV, $\mathbf{U}$, can be similarly specified as $\mathbf{\Gamma}_U = \mathrm{cov}(\mathbf{U}, \mathbf{U}) = \mathbb{E}[\mathbf{U}\mathbf{U}^{\mathrm{H}}]$, and is referred to as *Hermitian covariance*. As per Schreier and Scharf [61] in case of complex-valued RVs, if the RV $\mathbf{U}$ and its complex conjugate $\mathbf{U}^*$ are correlated, then the covariance matrix as shown above does not describe all the second-order properties of the complex-valued RV. We additionally need to introduce the *complimentary covariance* matrix $\tilde{\mathbf{\Gamma}}_U = \mathrm{cov}(\mathbf{U}, \mathbf{U}^*) = \mathbb{E}[\mathbf{U}\mathbf{U}^{\mathrm{T}}]$ (it is also known by other names as *pseudo-covariance*, *conjugate covariance* matrix or *relation* matrix). Both the Hermitian covariance matrix and the complimentary covariance matrix together describe the complete second-order properties of a complex-valued RV $\mathbf{U}$. For simplicity we introduce the

composite covariance matrix $\mathbf{\Gamma}_V$ of a real composite vector $\mathbf{V}$ [61] as

$$\mathbf{\Gamma}_V = \mathbb{E}[\mathbf{V}\mathbf{V}^{\mathrm{T}}] = \begin{bmatrix} \mathbf{\Gamma}_X & \mathbf{\Gamma}_{XY} \\ \mathbf{\Gamma}_{YX} & \mathbf{\Gamma}_Y \end{bmatrix} \in \mathbb{R}^{2n\times 2n}, \tag{2.1.6}$$

and the augmented covariance matrix $\mathbf{\Gamma}_W$ of the complex augmented vector $\mathbf{W}$ as

$$\mathbf{\Gamma}_W = \mathbb{E}[\mathbf{W}\mathbf{W}^{\mathrm{H}}] = \mathbb{T}\mathbf{\Gamma}_V\mathbb{T}^{\mathrm{H}} = \begin{bmatrix} \mathbf{\Gamma}_U & \tilde{\mathbf{\Gamma}}_U \\ \tilde{\mathbf{\Gamma}}_U^* & \mathbf{\Gamma}_U^* \end{bmatrix} \in \mathbb{C}^{2n\times 2n}. \tag{2.1.7}$$

2.1.1 Propriety and circularity

We aim to clarify the propriety and circularity as concepts inherent to complex RVs. Many authors mix these two terms and incorrectly assume they are interchangeable, We aim to make the distinction clear and illustrate the differences. In general, propriety of a RV means that there is no relation between the real and imaginary part of the RV. That can be desirable because proper complex-valued RVs can be treated in the same way as real-valued RVs. The inference of such RVs is no different to real-valued RVs and the same concepts apply.

Definition 2.1.2 (Propriety [62, p. 35])
A complex-valued zero-mean RV, $\mathbf{U} \in \mathbb{C}^n$, is called proper, if its complimentary covariance matrix is zero, $\tilde{\mathbf{\Gamma}}_U = \mathbf{0}$, otherwise its known as improper.

Additionally for a set of two complex-valued RVs, $\mathbf{U}$ and $\mathbf{Z} \in \mathbb{C}^n$, if the complimentary cross-correlation matrix between them is zero, $\tilde{\mathbf{\Gamma}}_{UZ} = \mathbb{E}\{\mathbf{U}\mathbf{Z}^{\mathrm{T}}\} = \mathbf{0}$, then we say they are *cross-proper*. Whereas, $\mathbf{U}$ and $\mathbf{Z}$ are *jointly proper*, if $\tilde{\mathbf{\Gamma}}_U = \mathbf{0}$, $\tilde{\mathbf{\Gamma}}_{UZ} = \mathbf{0}$ and $\tilde{\mathbf{\Gamma}}_Z = \mathbf{0}$ [60]. Moreover, if $\mathbf{U}$ is proper complex RV then any complex RV $\mathbf{Z}$ obtained from $\mathbf{U}$ by a linear or affine transformation will still be proper [38]. As well as any linear combination of two RVs $\mathbf{Z} = a_1\mathbf{U}_1 + a_2\mathbf{U}_2$ will be proper, only if both $\mathbf{U}_1$ and $\mathbf{U}_2$ are proper.

Definition 2.1.3 (Circularity [43])
A complex random vector is circular, if for all $\alpha \in \mathbb{R}$, $\mathbf{U}$ and $e^{i\alpha}\mathbf{U}$ will have the same probability density functions, that is to say its probability distribution is rotationally invariant.

Circularity is important, because such models have fewer parameters, and so by

the principle of parsimony we should chose simpler model to avid over-fitting. According to Adali et al. [2] for a RV to be proper all the second moments have to be rotationally invariant, whereas for a RV to be circular all the moments have to be rotationally invariant. Thus, it is evident that circularity is a much stronger condition, and implies propriety, but not vice versa. At the level of second-order statistics, the condition for rotational invariance poses all the restrictions on the complimentary covariance matrix only, the Hermitian covariance matrix is unrestricted, as shown below

$$\mathbf{\Gamma}_U = \mathbb{E}\{\mathbf{U}\mathbf{U}^{\mathrm{H}}\} = \mathbb{E}\{e^{i\alpha}\mathbf{U}\mathbf{U}^{\mathrm{H}}e^{-i\alpha}\} = \mathbf{\Gamma}_U,$$

$$\tilde{\mathbf{\Gamma}}_U = \mathbb{E}\{\mathbf{U}\mathbf{U}^{\mathrm{T}}\} = \mathbb{E}\{e^{i\alpha}\mathbf{U}\mathbf{U}^{\mathrm{T}}e^{i\alpha}\} = e^{i2\alpha}\tilde{\mathbf{\Gamma}}_U.$$

From this we can see that the only case when propriety and circularity are the same is when the complex RV is Gaussian and zero-mean, because the Gaussian distribution is completely defined by its first and second order moments.

2.1.2 Affine transformations of complex random vectors

Pre-multiplication of vectors by a matrix leads to a transformation of the vector. One of the most general classes of transformations is the affine transformation, which is composed of several operations. Here we introduce the general topic of widely-linear affine transformations. This will be useful later as we think of our parametric time series models as an affine transformation of the time series vector at every time point.

Definition 2.1.4 (Affine transformation [64, p. 29])
Affine transformation is a general form of linear transformation $\mathbb{R}^p \to \mathbb{R}^p$*, which has the form of* $\mathbf{x}' \to \mathbf{T}\mathbf{x} + \mathbf{a}$*, where* $\mathbf{T}$ *is a* $p \times p$ *matrix and* $\mathbf{x}$ *and* $\mathbf{a}$ *are* $p \times 1$ *column vectors. This can be extended to a widely-linear affine transformation of the form* $\mathbf{x}' \to \mathbf{T}\mathbf{x} + \mathbf{T}'\mathbf{x}^* + \mathbf{a}$*, where* $\mathbf{T}'$ *is a different transformation matrix that acts on the complex conjugate of the vector* $\mathbf{x}^*$*.*

Affine transformations can include any combination of rotation, scaling and/or translation. In a 2-dimensional case we can interpret an affine transformation as

composed of multiple sequential transformations [66]

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \mathbf{B}\ \mathbf{A}\ \mathbf{R}_\theta \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}, \quad \text{where } \mathbf{A} = \begin{bmatrix} \pm 1 & 0 \\ 0 & \mp 1 \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} k_x & k_{xy} \\ k_{yx} & k_y \end{bmatrix}. \tag{2.1.8}$$

In the above from, matrix $\mathbf{R}_\theta$ rotates the original axes by angle θ while preserving the orientation ($\det\{\mathbf{R}_\theta\} = 1$). Matrix $\mathbf{A}$ produces a line-reflection that reverses the orientation ($\det\{\mathbf{A}\} = -1$), if the signs are different, or point-reflection preserving orientation, if both signs are negative. Matrix $\mathbf{B}$ causes scaling, stretching or shear transformation of the axes [66]. The scaling factors, $k.$, control the grade of transformation so that the size is changed, but the parallel lines are preserved. If $k_x = k_y$ we have 'isotopic dilatation' preserving shape and angles, if $k_x \neq k_y$ we get a distorted transformation, and if $k_{xy}, k_{yx} \neq 0$ the transformation is a shear that changes angles and shape.

2.2 Complex-valued stochastic processes

Natural forces are influenced by many factors and as such are random, in addition to making a measurement error which adds to the randomness of the data collected. Stochastic processes exhibit uncertainty and we cannot exactly predict the value of the process at each time point. The possible values of the process have a probability distribution describing their occurrence. In this thesis we will refer to deterministic functions of time as *signals*, whereas to denote stochastic processes we use the term *processes* or *time series*. However, our interest is to mainly research random processes.

In addition, we will mainly focus on complex-valued random processes. According to Olhede and Walden [41] a complex-valued signals or processes can arise in three different ways: (1) from two unrelated signals; (2) as two components of polarised motion that are closely related (such as for example north-south and east-west velocity components in oceanography [29, 68]); and (3) from a complex-valued recording (e.g. in quadrature Doppler ultrasound [14, p. 89]). The main advantage of complex processes is that the relationship between the two time series or two components of a single series is preserved and analysed in a univariate time series, which is composed of complex-valued quantities with real and imaginary parts. In

this chapter we will present complex-valued stochastic processes and talk about their main properties. Just for simplicity we will talk about univariate processes, but the extension to multivariate processes is very simple and natural.

Definition 2.2.1 (Complex-valued stochastic process [35, p. 31])
Let $\{X(t,\omega) \mid t \in T\}$ and $\{Y(t,\omega) \mid t \in T\}$ be two real-valued stochastic processes defined on the same probability space $\mathcal{P} = (\Omega, \mathcal{F}, P)$, where T usually denotes an interval or set of positive integers. Then we have a complex-valued stochastic process $\{U(t,\omega) \mid t \in T\}$ defined on the same probability space $\mathcal{P}$ with $U(t,\omega) = X(t,\omega) + iY(t,\omega)$, so that the real and imaginary parts are P-measurable functions on $\mathcal{P}$ for all $t \in T$.

In the rest of the text, for simplicity, we drop the dependence on ω. We use the notation $\{U(t)\}$ to talk about continuous time stochastic process, and $\{U_t\}$ for discrete time process where t denotes discrete sampling parameter and Δt is the sampling period, such that $U_t = U(t\Delta t)$. Since in time series analysis herein we mainly deal with discretely sampled finite realisations of a time series process, we will mainly use the notation $\{U_t\}$ and reserve the notation $\cdot(t)$ to denote a parameter as a function of time t. In other words the data we talk about are mostly discretely sampled from a continuous processes, since usually it is not possible to observe continuous data directly.

For the purposes of this thesis we establish the notation for discrete complex-valued stochastic time series as $\{U_t\} \in \mathbb{C}$ for $t = 1, \ldots, T$, that is composed from two real-valued time series $\{X_t\}$ and $\{Y_t\}$ in the usual form, $U_t = X_t + iY_t$. For simplicity, we assume that the process is zero-mean, $\mathbb{E}[U_t] = 0, \forall t$. In contrast to the stochastic process, a deterministic signal is denoted with small letters, such as $u_t = x_t + iy_t$. We continue to use the notation of the real composite vector $\mathbf{V}_t = [X_t, Y_t]^{\mathrm{T}} \in \mathbb{R}^2$ and the complex augmented vector $\mathbf{W}_t = [U_t, U_t^*]^{\mathrm{T}} \in \mathbb{C}^2$, for $t = 1, \ldots, T$ in the same way as already introduced earlier in Chapter 2.1.

2.2.1 Autocovariance and propriety

In the case of random processes we are not only interested in the dependence at every time point, but also in the dependence between different points in time. This relationship is captured by the autocovariance function that tells us the dependence of the time series with itself between two points in time. In addition to the usual

autocovariance function we also have to express the complimentary autocovariance function, because we are dealing with complex-valued processes.

Definition 2.2.2 (Autocovariance function [35])
The autocovariance function (ACVF) of a complex-valued, zero-mean process $\{U_t\}$*, is defined as* $\gamma_U(t,r) = cov(U_t, U_r) = \mathbb{E}[U_t U_r^*]$ *and the complimentary autocovariance function (C-ACVF) as* $\tilde{\gamma}_U(t,r) = cov(U_t, U_r^*) = \mathbb{E}[U_t U_r]$*, for all* $t, r \in T$*.*

If the process $\{U_t\}$ is stationary then the ACVF and C-ACVF do not depend on the time t, but only on the difference between the two time points $t-r$, which we will call lag and denote τ. So the ACVF can be expressed as $\gamma_U(\tau) = \mathrm{cov}(U_{t+\tau}, U_t) = \mathbb{E}[U_{t+\tau}U_t^*]$ and the C-ACVF as $\tilde{\gamma}_U(\tau) = \mathrm{cov}(U_{t+\tau}, U_t^*) = \mathbb{E}[U_{t+\tau}U_t]$. The properties of ACVF in case of complex-valued process are [8, p. 115]: (1) $\gamma_U(0) \geq 0$; (2) $|\gamma_U(\tau)| \leq \gamma_U(0)$; (3) $\gamma_U(\tau) = \gamma_U^*(-\tau)$, i.e. ACVF is a Hermitian function; and (4) non-negative definite. By using the Definition 2.1.2 in Chapter 2.1.1 we can say that a complex stochastic process will be *proper*, if its complimentary autocovariance function $\tilde{\gamma}_U(\tau)$ will vanish [38].

For simplicity and in order to capture the full second-order structure we introduce the composite ACV matrix $\mathbf{\Gamma}_V(\tau)$ and the augmented ACV matrix $\mathbf{\Gamma}_W(\tau)$ for a univariate process $\{U_t\}$ by using the real composite and complex augmented vectors $\mathbf{V}_t$ and $\mathbf{W}_t$. These matrices can be expressed as

$$\mathbf{\Gamma}_V(\tau) = \mathbb{E}[\mathbf{V}_{t+\tau}\mathbf{V}_t^{\mathrm{T}}] = \begin{bmatrix} \gamma_X(\tau) & \gamma_{XY}(\tau) \\ \gamma_{YX}(\tau) & \gamma_Y(\tau) \end{bmatrix}, \tag{2.2.1}$$

$$\mathbf{\Gamma}_W(\tau) = \mathbb{E}[\mathbf{W}_{t+\tau}\mathbf{W}_t^{\mathrm{H}}] = \begin{bmatrix} \gamma_U(\tau) & \tilde{\gamma}_U(\tau) \\ \tilde{\gamma}_U^*(\tau) & \gamma_U^*(\tau) \end{bmatrix}, \tag{2.2.2}$$

where $\gamma_U(\tau) = \gamma_U^*(-\tau)$ and $\tilde{\gamma}_U(\tau) = \tilde{\gamma}_U(-\tau)$. The composite and augmented ACV matrices are useful because they capture the full second-order structure in one matrix, which can be further used for analysis. These two matrices are also related by a simple algebraic relationship using the real-to-complex transformation matrix in Equation (2.1.1), as $\mathbf{\Gamma}_W(\tau) = \mathbb{T}\mathbf{\Gamma}_V(\tau)\mathbb{T}^{\mathrm{H}}$.

2.2.2 Stationarity and ergodicity

Here we introduce both the concept of stationarity and ergodicity. The first one is important for general simplicity of analysis and the latter one is more of a generally assumed property of processes.

Definition 2.2.3 (Strict & weak stationarity [8, p. 12])
A real-valued time series $\{X_t\}$ *is* strictly stationary *when the joint distributions of* $[X_{t_1}, \ldots, X_{t_k}]^{\mathrm{T}}$ *and* $[X_{t_1+\tau}, \ldots, X_{t_k+\tau}]^{\mathrm{T}}$ *are the same for all positive integers* k *and for all* $t_1, \ldots, t_k$, $\tau \in \mathbb{Z}$. Weak stationarity, *however, means that the first two moments exist and are invariant under the shift of time. In other words* $\mathbb{E}[X_t] = \mu$ *for all* $t \in \mathbb{Z}$ *and* $cov[X_r, X_s] = cov[X_{r+t}, X_{s+t}]$ *for all* $r, s, t \in \mathbb{Z}$.

In the rest of our research we will consider mainly Gaussian processes as defined by Doob [12], for which strict stationarity is implied by weak stationarity, since the Gaussian distribution is completely defined by the first two moments. Stationary processes are key in time series analysis and most of the theory has been developed for stationary processes. They are especially important to give a meaningful interpretation to the first and second orders of the time series. In the case of complex-valued random processes we need to consider *wide-sense stationarity* (WSS). It is equal in definition to weak stationarity, but in this case we need to consider both the (Hermitian) covariance matrix and the complimentary covariance matrix, with the exception of proper signals. In order to satisfy wide sense stationarity both $\gamma(\tau)$ and $\tilde{\gamma}(\tau)$ need to be independent of t [62, p. 55].

Definition 2.2.4 (Ergodicity [18, p. 46])
We say that the process is ergodic, if the time sample moments converge in probability to the ensemble moments as $T \to \infty$. *For example, in case of the first order, if the sample average* $\overline{X} = \frac{1}{T}\sum_{t=1}^{T} X_t$ *will converge in probability to the expectation* $\mathbb{E}[X_t]$.

We observe several time points of one realisation of a time series process, but nevertheless we would like to calculate its sample moments, such as for example the sample mean, which shoud be taken accross all possible realisations. Unfortunately, from one observation we can only calculate the time moments and not the ensemble moments. That is why the ergodic property is important as it allows us to obtain the ensemble moments by calculating the time moments of a realisation. Most common

is the sample mean, which we calculate as the time average from a realisation. The ergodicity ensures that the time average, $\overline{X} = (1/T)\sum X_t$, will eventually converge to the ensemble first moment, $\mathbb{E}[X_t]$. In general, it is very difficult to check if the ergodicity is satisfied. One such way is by using the property that a process whose autocovariance function $\gamma(\tau)$ goes to zero quickly enough as τ becomes large will be ergodic [18, p. 46-47]. Such a process will satisfy the condition

$$\sum_{\tau=0}^{\infty} |\gamma(\tau)| < \infty. \tag{2.2.3}$$

2.2.3 Spectral analysis of complex processes

The idea of spectral analysis is to see the stochastic process as a distribution of energy on a whole range of frequencies. This view is widely used especially in the signal processing and engineering communities. In this chapter we give a quick overview of the spectral analysis theory important for discrete time stochastic processes with continuous frequencies. In order to do that we need to start with the spectral representation theorem.

Definition 2.2.5 (Spectral representation [8, p. 117])
Every stationary and zero-mean process can be written in terms of its spectral representation as a stochastic integral, or as

$$U_t = \int_{-1/2}^{1/2} e^{i2\pi ft} dZ_U(f), \tag{2.2.4}$$

which is defined for frequencies $f \in [-1/2, 1/2)$ and where $\{dZ_U(f)\}$ is an orthogonal increment process as described below in Definition 2.2.6.

By Hergoltz theorem [8, p. 118], $\gamma_U(\tau) \in \mathbb{C}$ will be the autocovariance function of a stationary process $\{U_t\}$, if and only if it will be (1) non-negative definite, e.g. $\sum_{i,j=1}^{n} a_i \gamma_U(i-j) a_j^* \geq 0$, and (2) will have the following representation

$$\gamma_U(\tau) = \int_{-1/2}^{1/2} e^{i2\pi f\tau} d\mathfrak{S}_U(f). \tag{2.2.5}$$

In the above equation the integration is with respect to a right-continuous, non-decreasing and bounded on $(-\pi, \pi]$ function $\mathfrak{S}(\cdot)$ called *spectral distribution func-*

tion, such that $\mathfrak{S}(-\pi) = 0$. If this function is differentiable, as $\mathfrak{S}(f) = \int_{-1/2}^{1/2} S(f)df$, then $S(f)$ is the *spectral density* and $d\mathfrak{S} = S(f)df$. Using this property and Equation (2.2.5) and if $\gamma_U(\tau)$ is absolutely summable $\sum_{n=-\infty}^{\infty} |\gamma_U(\tau)| < \infty$ [8, p. 120], we can derive its inverse as

$$S_U(f) = \sum_{\tau=-\infty}^{\infty} e^{-i2\pi f\tau}\gamma_U(\tau) \quad \text{for all } f \in [-\tfrac{1}{2}, \tfrac{1}{2}) \text{ and } \tau \in \mathbb{N}. \tag{2.2.6}$$

Definition 2.2.6 (Orthogonal increment process [8, p. 138])
Define a right-continuous complex-valued process $\{Z(f)\}$ on interval $[-1/2, 1/2]$ to be orthogonal increment process if it satisfies the following properties:

i $\mathbb{E}[Z(f)] = 0$ *for* $-1/2 \leq f \leq 1/2$,

ii $\mathbb{E}|Z(f)|^2 < \infty$ *for* $-1/2 \leq f \leq 1/2$,

iii $\mathbb{E}[(Z(f_4) - Z(f_3))(Z(f_2) - Z(f_1))^*] = 0$ *for* $(f_1, f_2] \cap [f_3, f_4) = \emptyset$.

The orthogonal increment process is used to give meaning to the stochastic integral in (2.2.4) of the spectral representation in Definition 2.2.5. It has a unique distribution function $\mathfrak{Z}(\cdot)$ with the following properties: (1) $\mathfrak{Z}(f) = 0$ for $f \leq -1/2$; (2) $\mathfrak{Z}(f) = \mathfrak{Z}(1/2)$ for $f \geq 1/2$; and (3) $\mathfrak{Z}(f_2) - \mathfrak{Z}(f_1) = \mathbb{E}|Z(f_2) - Z(f_1)|^2$ for $-1/2 \leq f_1 \leq f_2 \leq 1/2$ [8, p. 139].

The spectral density function (SDF), $S_U(f)$, and complimentary spectral density function (C-SDF), $\tilde{S}_U(f)$, of a complex-valued stochastic process $\{U_t\}$ are defined through the spectral process with orthogonal increments $\{dZ_U(f)\}$ [53] as

$$\operatorname{var}[dZ_U(f)] = \mathbb{E}[dZ_U(f)dZ_U^*(\nu)] = \begin{cases} S_U(f)df & \text{if } f = \nu, \\ 0 & \text{if } f \neq \nu; \quad \text{and} \end{cases} \tag{2.2.7}$$

$$\operatorname{cov}[dZ_U(f), dZ_U^*(-f)] = \mathbb{E}[dZ_U(f)dZ_U(\nu)] = \begin{cases} \tilde{S}_U(f)df & \text{if } f = -\nu, \\ 0 & \text{if } f \neq -\nu. \end{cases} \tag{2.2.8}$$

In the case of a real-valued process $\{X_t\}$ the SDF, $S_X(f)$, is symmetric, non-negative and has a finite integral on the interval over $[-1/2, 1/2]$. For a complex-valued process $\{U_t\}$ the SDF, $S_U(f)$, is still non-negative and has a finite integral on the interval over $[-1/2, 1/2]$, but is no longer symmetric, whereas the C-SDF $\tilde{S}_U(f)$ is due to stationarity symmetric, but generally complex-valued [45].

Similarly to the composite vector $\mathbf{V}_t$ and the augmented vector $\mathbf{W}_t$ that we have previously introduced, we can also construct the composite vector of the orthogonal processes as $d\mathbf{Z}_V(f) = [dZ_X(f), dZ_Y(f)]^{\mathrm{T}}$ and the complex augmented vector of the orthogonal process as $d\mathbf{Z}_W(f) = [dZ_U(f), dZ_U^*(-f)]^{\mathrm{T}}$. These two vectors allow us to specify the composite SDF matrix $\mathbf{S}_V(f)$ and the augmented SDF matrix $\mathbf{S}_W(f)$ as follows

$$\mathbb{E}[d\mathbf{Z}_V(f)d\mathbf{Z}_V^{\mathrm{T}}(f)] = \begin{bmatrix} S_X(f) & S_{XY}(f) \\ S_{YX}(f) & S_Y(f) \end{bmatrix} df = \mathbf{S}_V(f)df, \tag{2.2.9}$$

$$\mathbb{E}[d\mathbf{Z}_W(f)d\mathbf{Z}_W^{\mathrm{H}}(f)] = \begin{bmatrix} S_U(f) & \tilde{S}_U(f) \\ \tilde{S}_U^*(f) & S_U(-f) \end{bmatrix} df = \mathbf{S}_W(f)df. \tag{2.2.10}$$

The two composite and augmented vectors of the orthogonal increment process can be related by using the real-to-complex transformation matrix, as $d\mathbf{Z}_W(f) = \mathbb{T}d\mathbf{Z}_V(f)$. In the same way we can also relate the composite and augmented SDF matrices, as $\mathbf{S}_W(f) = \mathbb{T}\mathbf{S}_V(f)\mathbb{T}^{\mathrm{H}}$.

Both the augmented and the composite SDF matrices have to be positive semi-definite. In the first case this condition will be satisfied if (1) the SDF is always non-negative $S_U(f) \geq 0$, (2) the property of the complimentary SDF is satisfies the following $\tilde{S}_U(f) = \tilde{S}_U(-f)$, and (3) in relation to each other they satisfy the following property $|\tilde{S}_U(f)|^2 \leq S_U(f)S_U(-f)$ [62, p. 198]. In the second case the (1) both SDFs have to be non-negative $S_X(f) \geq 0$ and $S_Y(f) \geq 0$, (2) the cross SDF satisfies $S_{XY}(-f) = S_{YX}(f) = S_{XY}^*(f)$, (3) and together they satisfy $|\tilde{S}_{XY}(f)|^2 \leq S_X(f)S_Y(-f)$ [5, p. 213].

2.3 Analytic signals

The idea behind analytic signals is to construct a complex-valued signal from a real-valued signal, such that the original real signal would be equal to the the real part of the analytic signal and maintain the same spectral decomposition [15]. In other words we wish to construct a complex-valued signals that has the same frequency decomposition as the original real-valued signal. This is very useful because we can easily switch between real-valued and complex-valued signals without changing any information the signal carries.

Analytic signals are constructed from any real-valued time series by using the

Hilbert transform. The inverse process is very simple, one obtains the original real-valued time series just by taking the real part of the analytic signal. Since the SDF of real-valued signals is symmetric the negative frequencies are redundant and can be discarded, which is exactly the property that the analytic signal exploits to construct a signal with proportional spectral representation.

Definition 2.3.1 (Analytic signal Gabor [15])
Analytic signal of a real-valued signal $\{x_t\}$ *is defined as*

$$x_{+,t} = x_t + i\mathcal{H}\{x_t\} = a_x(t)e^{i\phi_x(t)}, \tag{2.3.1}$$

where $\mathcal{H}\{x_t\}$ *is the Hilbert transform of* $\{x_t\}$*. The amplitude of the signal is* $a_x(t) = |x_{+,t}|$ *and the phase is* $\phi_x(t) = \arctan\left(\frac{\Im\{x_{+,t}\}}{\Re\{x_{+,t}\}}\right)$.

The Discrete Hilbert Transform (DHT) for a real-valued signal $\{x(t)\}$ is defined by Kak [25] as

$$\mathcal{H}\{x_t\} = \begin{cases} \frac{2}{\pi} \sum\limits_{\text{n odd}} \frac{x_u}{t-u} & \text{if } t \text{ is even,} \\ \frac{2}{\pi} \sum\limits_{\text{n even}} \frac{x_u}{t-u} & \text{if } t \text{ is odd.} \end{cases} \tag{2.3.2}$$

Whereas, the continuous version of the Hilbert transform for a real-valued signal $\{x(t)\}$ is defined by Cohen [9, p. 31] as

$$\mathcal{H}\{x(t)\} = x(t) * (\pi t)^{-1} = \tfrac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(u)}{t-u} du, \tag{2.3.3}$$

where "$*$" denotes convolution. The Hilbert transform is a convolution between $\{x(t)\}$ and $(\pi t)^{-1}$, and the above integral is defined as the Cauchy principal value, in order to ensure that the integral is defined also when $t = u$.

The analytic signal is a well known representation of a univariate and real-valued signal viewed as an amplitude/frequency modulated signal. We can see, from the above Equation (2.3.1), that an analytic signal can always be represented in the polar form, which gives us the natural link to the phase, instantaneous frequency and amplitude of a signal [9]. The concept of analytic signals enables us to calculate the envelope of the signal. As we introduced earlier in many cases we may prefer to use complex-valued time series models over real-valued ones. The analytic signal theory gives us the tool to create a complex-valued process for a observed real-valued

time series that will have the same spectrum as the original process.

We would like to point out two useful properties of the analytic signals. The first being that the original real-valued signal is obtained simply by taking the real part of the analytic signal, $x_t = \Re\{x_{t,+}\}$. The second one, is that the Fourier transform (or the spectrum) of the analytic signal on the positive frequencies is the same shape as the one of the original real-valued time series. The frequency contributions for all the positive frequencies are doubled, but the frequency contributions of the negative frequencies are zero [30]

$$S_{x_+}(f) = \mathfrak{H}(f)S_x(f), \quad \mathfrak{H}(f) = \begin{cases} 2 & \text{if } f > 0, \\ 1 & \text{if } f = 0, \\ 0 & \text{if } f < 0; \end{cases} \tag{2.3.4}$$

where $\mathfrak{H}(f)$ is the 'unit step function' [1, p. 1020].

The analytic signals are very useful, for example, in radio communications, where complex baseband information signal is modulated and only the real part is transmitted. The receiver reconstitutes the original signal by forming an analytic signal and demodulating [47]. There are also other practical uses in signal processing and time series analysis, where one creates a complex valued signal from real valued data, e.g. speech analysis [67].

Chapter 3

Time series analysis

The main aim of this chapter is to link the complex-valued random vectors and processes as described in the previous chapter to time series analysis. We provide the basic time series theory and talk about time series models. We also introduce the concept of pseudo-periodic behaviour that we will use in the modelling part of our work. In this thesis we focus only on stationary processes, because stationarity is important for many of the concepts we rely on. Additionally, this chapter also summarises the literature on multivariate time series, which we will use later on for the construction of multivariate time series models.

Time series analysis can be viewed as analysis of sequential data [13, p. 371]. In time series analysis one is not only concerned with the multivariate dependence, but also temporal relationships between data points. This brings us to the argument that time series data have a specific direction. In other words time evolves only forward, so the direction of the data cannot be freely interchanged, as it can in the case of a random vector. Because we are interested in the relationship between the data points evolution through time, the standard random vector theory will not be sufficient for the analysis of time series data.

One of the most studied processes and the basic building block of time series analysis is the *autoregressive moving average model* (ARMA). Many authors have analysed it and many new models are based on and derived from ARMA models (see [7, 8, 18]). Most of this chapter is derived in terms of a complex-valued process $\{U_t\}$, but in some parts where we talk about real-valued process, $\{X_t\}$. However, this theory equally applies to both a real and complex-valued process.

Definition 3.0.2 (ARMA [8, p. 78])
An autoregressive moving average processes ARMA(p,q) is defined using linear difference equations with constant coefficients, in the different forms, as

$$\begin{aligned} U_t - \varphi_1 U_{t-1} - \ldots - \varphi_p U_{t-p} &= \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}, \\ U_t - \sum_{k=1}^{p} \varphi_k U_{t-k} &= \sum_{l=0}^{q} \theta_l \epsilon_{t-l}, \\ \Phi(B) U_t &= \Theta(B) \epsilon_t, \end{aligned} \tag{3.0.1}$$

where B is the lag operator and $\{\epsilon_t\} \sim N(0, \sigma^2)$ is a white noise process.

In the above definition, $\Phi(B)$ and $\Theta(B)$ are called characteristic polynomials and are evaluated in the lag operator B. By Brockwell and Davis [8, p. 84-88] any ARMA process whose polynomials $\Phi(\cdot)$ and $\Theta(\cdot)$ have no common zeros is (1) *stationary and causal*, if and only if, all the zeros of $\Phi(z)$ lie outside the unit circle, i.e. $\Phi(z) \neq 0$ for all $z \in \mathbb{C}$ such that $|z| \leq 1$; and (2) *invertible*, if and only if, all the zeros of $\Theta(z)$ lie outside the unit circle, i.e. $\Theta(z) \neq 0$ for all $z \in \mathbb{C}$ such that $|z| \leq 1$.

Stationarity of the process is a desirable property, because then the process characteristics, such as the mean and the variance, are constant over time. Whereas, causality means that the process depends only on its past values and not the future values, and as such can be expressed as an infinite sum of past innovation terms. Invertibility means that the process can be inverted and expressed as an infinite sum of the past values. Any stationary and invertible ARMA, as defined above, can be represented with polynomials $\Psi(\cdot)$ and $\Pi(\cdot)$, respectively

$$U_t = \frac{\Theta(B)}{\Phi(B)} \epsilon_t = \Psi(B) \epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, \tag{3.0.2}$$

$$\epsilon_t = \frac{\Phi(B)}{\Theta(B)} U_t = \Pi(B) U_t = \sum_{j=0}^{\infty} \pi_j U_{t-j}. \tag{3.0.3}$$

Time invariant linear filters Time invariant filters are generally, due to their many uses, very important in time series analysis. By applying a linear filter to a signal we can obtain a new signal with certain desired characteristics. For example, in spectral analysis of time series they help us keep or suppress certain frequencies from a source signal and obtain a time series with desired shape of spectral

distribution function.

Based on Brockwell and Davis [8, p. 153], if a time-invariant linear filter $H = \{h_i, i = 0, \pm 1, \ldots\}$, that does not depend on time and is absolutely summable $\sum_{n=-\infty}^{\infty} |h_n| < \infty$, is applied to series $\{U_t\}$ we obtain time series $\{Y_t\}$, in the following way

$$Y_t = \sum_{k=-\infty}^{\infty} h_k U_{t-k}. \tag{3.0.4}$$

The linear filter H is causal, if the new process $\{Y_t\}$ can be expressed only with past values of the process $\{U_t\}$, i.e. $h_k = 0$ for all $k < 0$. Additionally, if the process $\{U_t\}$ has spectral representation as defined in (2.2.4) and the filter H converges $\sum_{j=-\infty}^{\infty} h_j e^{-ij\cdot} = h(e^{-ij\cdot})$ in $\mathcal{L}^2$, then we can write the spectral distribution function and the spectral representation of process $\{Y_t\}$ in the following way

$$\mathfrak{S}_Y(f) = \int_{-1/2}^{1/2} |h(e^{-i2\pi f})|^2 d\mathfrak{S}_U(f), \tag{3.0.5}$$

$$Y_t = \int_{-1/2}^{1/2} e^{i2\pi ft} h(e^{-i2\pi f}) dZ_U(f). \tag{3.0.6}$$

This relationship is important for any stationary ARMA time series as it simplifies the way to obtain its spectral density function. We notice that the polynomial $\Psi(\cdot)$ in Equation (3.0.2) acts as a linear filter of the random noise ϵ_t. We also know that the spectral density of the white noise process $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ is just its variance, $S_\epsilon(f) = \sigma_\epsilon^2$. Thus, by using the concept of linear filters we can obtain the spectral density function of an ARMA model just by knowing its parameters and the noise variance. We can use this relationship in Equation (3.0.5) to obtain the spectral density function of an ARMA model [8, p. 123], as

$$S_U(f) = \frac{\sigma_\epsilon^2 |\Theta(e^{-i2\pi f})|^2}{|\Phi(e^{-i2\pi f})|^2}. \tag{3.0.7}$$

3.1 Autoregressive time series models

In this chapter we will look more in detail on the autoregressive time series models of general order p, AR(p). This class of models is a restricted version of the above ARMA(p,q) model, such that $q = 0$ or $\Theta(z) = 1$. A general AR(p) process can be

described by the equation

$$U_t = \sum_{j=1}^{p} \varphi_j U_{t-j} + \epsilon_t, \tag{3.1.1}$$

which holds for both real- and complex-valued AR(p) time series [8, p. 79]. When process $\{U_t\}$ is complex-valued, the φ_j are complex and ϵ_t is complex white noise. Moreover, a stationary discrete-time $\mathbb{C}$AR(p) model can also be viewed as an output of an $\mathbb{C}$AR(p) filter with the input being a discrete-time (complex-valued) white noise [46].

Especially interesting is the *first order* autoregressive model, AR(1), because it is dependant only on one time step in the past. That is to say, the process at time t depends only on time $(t-1)$ [48, p. 116]. This type of models exhibits the *Markov property*, because the conditional distribution depends only on values one step back in time, $p(U_t \mid U_{t-1}, U_{t-2}, U_{t-3}, \ldots) = p(U_t \mid U_{t-1})$.

Here we are interested in an $\mathbb{C}$AR(1) that satisfies the difference equation, $U_t = \varphi U_{t-1} + \epsilon_t$. The properties of such models are easy to define and mathematically simple to deal with. The stationarity of this model will be easily achieved simply by restricting the values of the parameter to $|\varphi| < 1$. Under stationarity and normal distribution of the error term, $\epsilon_t \sim N(0, \sigma_\epsilon^2)$, the process will also be zero-mean $\mathbb{E}[U_t] = 0$, and its ACVF will be given by $\gamma(\tau) = \varphi^{|\tau|} \frac{\sigma_\epsilon^2}{1-|\varphi|^2}$ [48, p. 116]. Moreover, another nice property is that this model is relatively easy to extend to the multivariate case.

3.1.1 Spectral density function of autoregressive models

Based on the theory introduced in Chapter 2.2.3 and the general formula for the spectral density function of an ARMA(p,q) in Equation (3.0.7), the SDF of an AR(p) process is easy to write as

$$S_U(f) = \frac{\sigma_\epsilon^2}{|\Phi(e^{-i2\pi f})|^2}, \qquad \text{for } -1/2 \leq f \leq 1/2. \tag{3.1.2}$$

In the above equation $\Phi(e^{-i2\pi ft})$ denotes the characteristic polynomial of the AR(p) process with the lag operator B changed for $e^{-i2\pi ft}$ and σ_ϵ^2 the variance of the white

noise process. In case of an AR(1) the above simplifies to

$$S_U(f) = \frac{\sigma_\epsilon^2}{|1 - \varphi e^{-i2\pi f}|^2} = \frac{\sigma_\epsilon^2}{1 - 2\varphi\cos(2\pi f) + \varphi^2}, \qquad \text{for} -1/2 \leq f \leq 1/2. \quad (3.1.3)$$

3.1.2 Pseudo-periodic behaviour

According to Priestley [48, p. 131] pseudo-periodic behaviour can be observed for real-valued processes $\{X_t\}$ both in their realisation and the autocovariance function (ACVF). The same pseudo-periodic behaviour can also be modelled with certain AR(2) models, i.e. $X_t + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} = \epsilon_t$, if both of the roots of the characteristic polynomial $\Phi(B)$ are complex-valued. The easiest way to understand why this happens for these type of models is through the behaviour of the ACVF. In the extreme case when the ACVF is an exactly periodic function, which happens when $\varphi_2 \to 1$, we can conclude that the process $\{X_t\}$ will also be exactly periodic. In the same way, if the ACVF function will be pseudo-periodic, damped periodic function, the process will also be periodic to some extent, some sort of 'distorted periodicity'.

The intuition behind this lies in looking at the continuous time equivalent of AR(2) family models. The analogue of a second-order difference equation is the second-order differential equation in the form $\ddot{X}(t) + \alpha\dot{X}(t) + \beta X(t) = \epsilon(t)$. This equation represents a well known harmonic motion, with α being the 'damping factor' and $\{\epsilon(t)\}$ the 'driving force' which sustains the motion of the harmonic oscillator. From this we can conclude that the pseudo-periodic behaviour of AR(2) is the definition of a stochastic oscillation.

Similar to the differential equation theory where we can transform a second-order differential equation to a system of first-order differential equations, we will be able to transform a real-valued AR(2) to a bivariate VAR(1) model and in some cases also to a univariate $\mathbb{C}$AR(1). This means that we can achieve the same pseudo-periodic behaviour with bivariate VAR(1) instead. We will exploit this property later on to propose multivariate models for stochastic oscillations.

3.1.3 Vector autoregressive models

The multivariate equivalent of an AR(p) model is the vector autoregressive model denoted as VAR(p). For most of the concepts introduced previously, general characteristics and properties are preserved. In the multivariate case one needs to

create a vector of time series processes, multiple time series, $U_t^{(1)}, U_t^{(2)}, \ldots, U_t^{(n)}$, will be stacked up into a vector of dimension n to form the vector time series $\mathbf{U}_t = [U_t^{(1)}, U_t^{(2)}, \ldots, U_t^{(n)}]^{\mathrm{T}}$, which will be used in the analysis.

In this case our multivariate time series are vector-valued so the parameters of the model need to be matrices. This, in turn, increases the number of parameters needed to be estimated. Based on Reinsel [51, p. 27] an n-dimensional time series VAR(p) model satisfies the following equation

$$\mathbf{U}_t - \sum_{j=1}^{p} \mathbf{\Phi}_j \mathbf{U}_{t-j} = \boldsymbol{\epsilon}_t, \quad \text{or} \quad \mathbf{\Phi}(B)\mathbf{U}_t = \boldsymbol{\epsilon}_t. \tag{3.1.4}$$

In the above equation $\mathbf{\Phi}(B) = \mathbf{I}_n - \mathbf{\Phi}_1 B - \ldots - \mathbf{\Phi}_p B^p$ is the characteristic polynomial, the $\mathbf{\Phi}_j$ are $n \times n$ dimensional matrices, and $\boldsymbol{\epsilon}_t = [\epsilon_t^{(1)}, \ldots, \epsilon_t^{(n)}]^{\mathrm{T}}$ is an n-dimensional vector white noise process with $\mathbb{E}[\boldsymbol{\epsilon}_t] = \mathbf{0}$ and $\mathbb{E}[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t^{\mathrm{H}}] = \Sigma_\epsilon, \forall t$.

The stationarity condition for VAR(p) models is defined in terms of its characteristic polynomial. The process will be stationary if the roots of $\det\{\mathbf{\Phi}(B)\} = 0$ are greater than one in absolute value [51, p. 27]. The autocovariance function will be an $n \times n$ ACV matrix denoted $\mathbf{\Gamma}_U(\tau) = \mathbb{E}[\mathbf{U}_{t+\tau}\mathbf{U}_t]$. The SDF will be an $n \times n$ spectral matrix that can be expressed as

$$\mathbf{S}_U(f) = \mathbf{H}(e^{-i2\pi f})\mathbf{\Sigma}_{\boldsymbol{\epsilon}}\mathbf{H}^{\mathrm{H}}(e^{-i2\pi f}), \tag{3.1.5}$$

where $\mathbf{H}(e^{-i2\pi f})$ is the *transfer function matrix* calculated as the inverse of the characteristic polynomial $\mathbf{H}(B) = \mathbf{\Phi}(B)^{-1}$ by substituting B with $e^{-i2\pi f}$ [48, p. 688].

3.2 Forecasting of time series

Hamilton [18, p. 73] defines the forecast for one period in the future, $\widehat{U}_t(1)$, as the conditional expectation of the value of the time series in that period conditioned on the values of time series today and in the past $\widehat{U}_t(1) = \mathbb{E}[U_{t+1} \mid U_t, U_{t-1}, U_{t-2}, \ldots]$. He also proves that this is the best estimate of U_{t+1} in respect of minimizing the mean squared error, $\mathrm{mse}\{\widehat{U}_t(1)\} = \mathbb{E}[U_{t+1} - \widehat{U}_t(1)]^2$. We can generalise this result to any l-step ahead forecast, where l is a positive integer, $l \in \mathbb{Z}_+$. The forecast of the value of the time series U_t l-periods in the future will then be

$$\widehat{U}_t(l) = \mathbb{E}[U_{t+l} \mid U_t, U_{t-1}, U_{t-2}, \ldots]. \tag{3.2.1}$$

The intuitive interpretation of this forecast is that it is the mean of the different values U_{t+l} can take, based on the time series generated by the particular realisation we have observed in the past [48, p. 728]. In order to evaluate the conditional expectation we need to know the joint PDF, so for the prediction to be optimal we usually assume multivariate normal structure. By doing this the expectation in Equation (3.2.1) becomes a linear function of the past values of $\{U_t\}$.

The expression of the forecast function of an AR(1), of the form $U_t = \varphi U_{t-1} + \epsilon_t$, is relatively simple due to the Markov property. The l-step ahead forecast will depend only on the today's value of the time series and the coefficient of the AR(1) model as

$$\begin{aligned} \widehat{U}_t(l) &= \mathbb{E}[U_{t+l} \mid U_t, U_{t-1}, U_{t-2}, \ldots] \\ &= \mathbb{E}[\varphi U_{t+l-1} + \epsilon_t \mid U_t, U_{t-1}, U_{t-2}, \ldots] \\ &= \varphi \widehat{U}_{t+l-1} = \ldots = \varphi^l U_t. \end{aligned} \tag{3.2.2}$$

The same property applies to both univariate and multivariate AR(1) models. This means that we can easily forecast the values of any AR(1) if we know its todays value and its parametric form - the model coefficient. We will exploit this useful property in the continuation to propose a definition of an elliptical model.

Chapter 4

Introduction to elliptical models

In this chapter we introduce elliptical models by reviewing the literature on elliptical signals, presenting the elliptical representations and specifying the various parameters of an ellipse. We start by defining multivariate oscillations. This chapter is intended to provide background understanding of elliptical models, which will be further elaborated in the following chapter. Since oscillations can be found in many natural phenomena, it is possible to establish a link between these models and applied research in other fields.

4.1 Oscillations

Periodic phenomena are present in everyday life, such as for example the rotation of earth around its axis or around the sun. Many such phenomena are deterministic, which means that the we exactly know the period and we are also able to predict its value in the future. On the other hand some natural phenomena, such as for example the velocity and movements of ocean currents, are random. They are influenced by other stochastic forces and we are not able to exactly predict their period in the future. The realisations of these processes will still be periodic (oscillatory), but their period will not be exact and the trajectories will not follow nice and smooth curves.

It is worth noting that in our work we use the term signals for deterministic functions of time, such as a simple sinusoid, but we use the term time series or process for a stochastic process with a certain distribution. Let us also clarify that because we usually observe only one realisation of the time series, we assume it is a discrete processes or a discrete sample of a continuous random process.

Let us begin by motivating oscillations from the physics literature. The usual way to motivate oscillations is with a simple harmonic motion [16, sec. 6.2]. Let us imagine a particle that rotates on the circumference of a circle with a certain angular velocity ω. At any point in time t, this particle will have a perpendicular projection onto the x-axis denoted as A, in a system where the intersection of the x and y axes is in the centre of the circle O. As the particle rotates on the circle, point A will move left and right on the x-axis. The distance between the origin of the coordinate system O and the projection of the particle A will be given by the function $x = a\cos(\omega t)$, where a is the radius of the circle. Another projection we can observe is the projection of the particle onto the y-axis, denoted as B. Now the distance from the origin of the system O to the point B will be given by $y = a\sin(\omega t)$. Together these movements define the oscillatory movement and can be described by the differential equation of a harmonic oscillator

$$\frac{d^2x}{dt^2} + \omega^2 x = 0, \tag{4.1.1}$$

with a general solution to the differential equation $x = A\sin(\omega t) + B\cos(\omega t)$ [16, p. 6.5].

The sinusoid is the most common periodic function and is also the solution to the differential equation of a simple harmonic oscillator as seen above. In general terms we can write a possible equation for a sinusoid curve as

$$x_t = A\cos(\omega t + \varphi), \tag{4.1.2}$$

where A is the amplitude, ω the angular frequency and φ the phase shift. The period of the above sinusoid is $2\pi/\omega$. In the above equation we can see that the frequency and amplitude are constant. This means that, if we were to observe a plot of this function in time t, we could easily spot the exact frequency and amplitude just by observing it. Contrary to that we can allow the frequency and/or the amplitude to be smooth functions of time, which we denote as $\omega(t)$ and $A(t)$. This means that the sinusoid will have a time-varying amplitude and/or frequency, and the change will be visible on the plot of the signal over time.

It is worth mentioning at the beginning that most of the theory that we introduce here is based on discrete time signals, but can be extended to continuous time processes as well. We decided to talk about the general theory in discrete time

only, because it follows naturally into the time series analysis. Time series is usually seen as a discrete sample of a continuous process $x(t)$ at regular time intervals Δt, such that $x_t = g(t\Delta t)$. Hence our notation follows the standard convention of x_t for discrete time signals and $x(t)$ for continuous time signals. However, both continuous time signals (e.g. $x(t)$) and time-varying parameters (e.g. $A(t)$), both will have arguments (t) and the context will clarify whether we refer to a signal or a parameter. But nevertheless, we will rarely talk about continuous time signals, so this notation will be mainly used for time-varying parameters.

In some instances we can allow to have more than one frequency and/or amplitude present in the signal. We can extend the simple oscillation in Equation (4.1.2) to allow a multicomponent oscillation [70]. Such models mix more than one sinusoid curves, each with one, potentially different, frequency and/or amplitude. This can be generalised to [48, p. 147]

$$x_t = \mu + \sum_{j=1}^{r} A^{(j)} \cos(\omega^{(j)} t + \varphi^{(j)}), \tag{4.1.3}$$

where $r \in \mathbb{N}$ is the number of sinusoidal terms present and $^{(j)}$ refers to the amplitude, frequency or phase of the j-th term. In a multicomponent model with various different sinusoids, it becomes very difficult to identify the various frequencies and amplitudes present just by looking at the trajectories of such a signal.

Fourier theory is very useful to identify the individual frequency components of a multicomponent signals. Signals, such as the ones in (4.1.3), are composed of several different sinusoids, each with a different frequency and amplitude. The Fourier transform of a signal decomposes this signal into its constituent sinusoids. That helps us to understand each sinusoid with its respective frequency and amplitude present in the signal. The Fourier theory entails all combinations of discrete or continuous time and frequency (for more detail see [42, p. 87]), here we will talk only about the case of discrete time and continuous frequency. Based on Percival & Walden [42, p. 87] we can write the Fourier transform $\tilde{x}(f)$ of a deterministic signal $\{x_t\}$, as

$$\tilde{x}(f) = \Delta t \sum_{t=-\infty}^{\infty} x_t e^{-i2\pi f t \Delta t}, \tag{4.1.4}$$

and the inverse Fourier transform as

$$x_t = \int_{-1/(2\Delta t)}^{1/(2\Delta t)} \tilde{x}(f) e^{i2\pi f t \Delta t} df. \tag{4.1.5}$$

The above class of multicomponent models in (4.1.3) were one of the first models to be studied in time series analysis [48, p. 148]. Initially these models for $\{x_t\}$ were treated as deterministic functions of t, and as such not a very interesting non-stationary process for time series analysis. To change the deterministic sinusoidal signal into a stochastic process one has two options: (1) let the phase $\varphi^{(j)}$ be treated as a random variable, usually as uniformly distributed on the interval $(-\pi, \pi)$; or (2) add a random noise ϵ_t with a distribution, for example $\epsilon_t \sim N(0, \sigma^2)$ [49, p. 5]. The latter is more useful for modelling real applications, where we usually include the noise also due to measurements error. Such extensions apply to both the simpler one component signals in (4.1.2) and the multicomponent models in (4.1.3), as shown below

$$X_t = A\cos(\omega t + \varphi) + \epsilon_t, \tag{4.1.6}$$

$$X_t = \sum_{j=1}^{r} A^{(j)} \cos(\omega^{(j)} t + \varphi^{(j)}) + \epsilon_t. \tag{4.1.7}$$

The above mentioned stochastic oscillations are basically deterministic oscillatory functions that have some of their parameters random or a noise term added. In time series, however, it would be more common to look at stochastic oscillations in terms of random time series models. In the former case, the trajectory would be a sinusoidal curve with noise, which would distort the smoothness of the function depending on its variance. Whereas in the latter case, the trajectory would not resemble a sinusoid function and we would not be able to see the period just by looking at its plot. As described already in Chapter 3.1.2 an AR(2) model with complex roots will have pseudo-periodic properties. We can construct such an AR(2) model that will always exhibit stochastic oscillations, such as for example

$$X_t = 2\frac{\cos(\zeta)}{\alpha} X_{t-1} - \frac{1}{\alpha} X_{t-2} + \epsilon_t, \tag{4.1.8}$$

where ϵ_t is white noise and ζ, $\alpha > 0$ (and for stationary processes $|\alpha| > 1$). The AR(2) in the above equation will have complex roots and thus observe pseudo-

periodic behaviour for any value of α and any ζ on the interval $(-\pi/2, \pi/2)$.

There are many examples of use of both deterministic and stochastic oscillations in natural and social sciences. Oscillations are very useful models because of their mean reverting properties. We can observe them in many applications, such as electromagnetic radiation [71], seismology [55], oceanography [29], econometrics [19], bloodflow [39], etc. For example in macroeconomic phenomena we find periodic time series, such as the business cycles. They tend to move together and models which use stochastic cycles can be used to model cyclical dynamics [54]. Such multivariate time series models are used to detect phase shifts among the cycles, i.e. the lead and lag indicators of each individual cycle.

4.2 Elliptical signals

One of the aims of our research is to observe the movement and orientation of oscillations, which in general can be elliptical, circular or linear. The elliptical polarisation as shown in Figure 4.1. We can see two components, the upper one (dotted) and lower one (dashed), which can either represent two elements of a bivariate vector or the real and imaginary parts of a complex-valued signal. Their common evolution is represented by the black thick line along the horizontal axis, which maps out an ellipse in the $x - y$ plane on the left-hand side of the figure. In other words waves that oscillate with more than one orientation trace a trajectory of an ellipse, which is called the *polarisation* of the wave. Under certain conditions the elliptical polarisation degenerates to either a circle or a straight line.

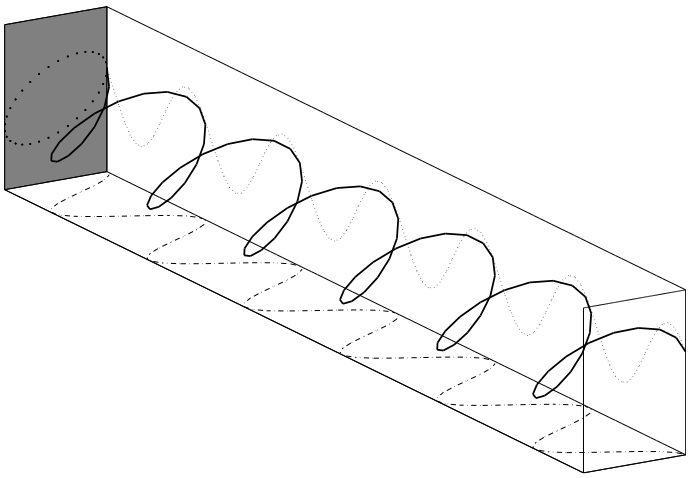


Figure 4.1: Elliptical polarisation

We are mainly interested in the *mapped ellipse*, which is the ellipse mapped out by the trajectory of the signal as we progress through the natural time of the signal. By the natural time of the signal (usually denoted t) we mean the time that the signal has been sampled in. The study of the ellipse trajectory has become an important research field also in signal processing and time series analysis, because it represents a simple parametrisation of a signal. Most of this research has been conducted for deterministic signals. In Figure 4.2 we show a sketch of a mapped ellipse as a geometric object with its parameters that can be seen from a geometric representation. In the continuation we will discuss similar ellipses and define their parameters. We mainly aim to define the mapped ellipse for stochastic time series, that is to say the intrinsic ellipse parametrization of the time series as such, but we will not explore complex geometry in detail.

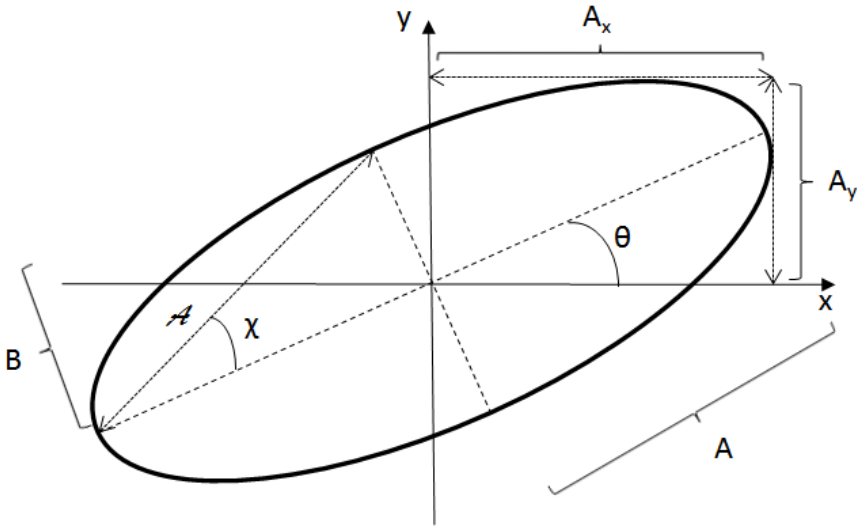


Figure 4.2: The ellipse and its parameters

Researchers in physics and optics fields have been studying deterministic waves for a long time. The easiest way to describe a mapped ellipse is, as the object mapped out by the trajectory of a monochromatic (single-frequency) deterministic signal. This signal is a non-random smooth function of time, and has an oscillation at only one specified frequency, e.g. a sinusoid with frequency ω of the form $\cos(\omega t)$ is a monochromatic deterministic signal. On one hand bivariate real-valued elliptical signals are such that, as we move through time the trajectory that this signal maps out in the $x - y$ plane has the shape of an ellipse as a geometric object (as shown

above in Figure 4.2). On the other hand a univariate complex-valued elliptical signal also maps out a trajectory of an ellipse, but in the complex plane, where the x-axis represents the real part and y-axis the imaginary part of the signal. A possible extension of the deterministic elliptical signals are modulated elliptical signals, whose ellipse parameters change with time (e.g. [29]). The trajectory of a modulated elliptical signal in time is not a perfect ellipse, because its properties change over time. Such signals have to be seen as ellipses that are frozen in time at every time step, which means that we cannot observe them in the 'global' time, see e.g. [30].

In case of stochastic processes, one cannot observe an ellipse mapped out by the trajectory of a stochastic time series, which is due to the randomness of the process and the fact that they are zero-mean. In order to define the ellipse of a stochastic signal authors have resorted to the frequency domain (see e.g. [53], [59]). In the frequency domain it is possible to define a random ellipse of the process at every frequency. The downside of this approach is that the frequency domain is less intuitive then the time domain for researchers and users in statistics, who prefer to look at time series in the time domain rather than the frequency domain.

Based on the above we can summarise that it is possible to define ellipses in both time and frequency domains [59]. Ellipses have been defined in time domain for deterministic signals and in frequency domain for both deterministic signals and stochastic processes. After reviewing the available literature and the ellipses that have already been defined, we see an advantage in proposing new definitions of non-random ellipses for stochastic processes in the time domain. In continuation we will first, for clarity, define what we mean by the object of ellipse and define all its parameters. In the main part of our work will present the various definitions of ellipses, the ones that exist in the literature and the ones we propose. First we will look through the ellipse definitions for deterministic signals and then through the ellipse definitions for stochastic time series, proposing two new concepts.

4.2.1 Ellipse representation

In this chapter we aim to provide a general representation of a modulated ellipse mainly with the intention of illustrating its geometric properties. According to Lilly [28] a particle φ on the interval $[-\pi, \pi]$ in 2-dimensions can be seen to circle an

ellipse (or its degenerations), if it can be represented in a parametric form in $\mathbb{R}^2$ as

$$\mathbf{x}_t(\varphi) = \boldsymbol{R}_{\theta(t\varepsilon)} \begin{bmatrix} A(t\varepsilon) & 0 \\ 0 & B(t\varepsilon) \end{bmatrix} \boldsymbol{R}_{\phi(t)} \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}. \tag{4.2.1}$$

In the above equation $\boldsymbol{R}_{\cdot}$ is a rotation matrix as described in Equation (2.1.3), and the parameters $A(t\varepsilon)$, $B(t\varepsilon)$ and $\theta(t\varepsilon)$ are time-varying, where ε controls the rate of modulation. Usually ε is chosen to be small, $\varepsilon \ll 1$, to ensure the smoothness of the modulated oscillation.

In order to assign a unique set of ellipse parameters to the above 2-dimensional real-valued signal we construct a complex-valued analytic signal $\mathbf{x}_+(t)$ (see Chapter 2.3), where the original signal is the real part of the analytic signal, $\mathbf{x}_t = \Re\{\mathbf{x}_{+,t}\}$. The ellipse parameters are then given by the 'normal form' ([27] and [28]), as

$$\mathbf{x}_{+,t} = e^{i\phi(t)} \boldsymbol{R}^T_{\theta(t\varepsilon)} \begin{bmatrix} A(t\varepsilon) \\ -iB(t\varepsilon) \end{bmatrix}, \tag{4.2.2}$$

where $A(t\varepsilon)$ and $B(t\varepsilon)$ are the semi-major and semi-minor of the modulated ellipse, $\theta(t\varepsilon)$ is the orientation (the angle between the semi-major and the x-axis), and the ε controls the rate of modulation. If $\varepsilon = 0$ it means that A, B and θ do not depend on t and so the ellipse is non-modulated. The only movement is the rotation around the trajectory of the ellipse with frequency $\omega(t) = \phi'(t)$. In the continuation of this chapter we will look at ellipses with fixed parameters (non-modulated ellipses with $\varepsilon = 0$), so we will drop the t from the arguments where not needed, apart from in the argument $\phi(t)$ where the time t makes the motion of mapping out the ellipse.

4.2.2 Parameters and measures of ellipticity

We use the above representation from (4.2.2) and the notation introduced in Figure 4.2 to discuss the parameters and measures of the ellipse. The aim is to present a comprehensive list of possible parameters and measures that can be observed for an elliptical signal. Parameters are geometric representations of the ellipse. Whereas measures define the size of the ellipse and its ellipticity, that is to say they help us distinguish the difference between an ellipse and its degenerate forms (circle and straight line). Table 4.1 summarises all the geometric parameters and measures, which are split into measures of size and ellipticity.

All of the geometric parameters can also be represented on a sketch. We can

Geometric	A	semi-major axis	$A = \kappa\sqrt{1+\lvert\lambda\rvert}$
	B	semi-minor axis	$\lvert B\rvert = \kappa\sqrt{1-\lvert\lambda\rvert}$
	θ	orientation angle (or azimuth)	$-\pi/2 \leq \theta \leq \pi/2$
	χ	ellipticity angle	$-\pi/4 \leq \chi \leq \pi/4$
Size	-	area	$AB\pi$
	κ	root-mean-square amplitude	$\kappa = \sqrt{\frac{A^2+B^2}{2}}$
	$\mathcal{A}$	amplitude	$\mathcal{A} = \sqrt{A^2+B^2} = \kappa/\sqrt{2}$
Ellipticity	$\mathcal{E}$	eccentricity	$\mathcal{E} = \pm\sqrt{1-\frac{B^2}{A^2}}$
	λ	ellipse parameter	$\lambda = \pm\frac{A^2-B^2}{A^2+B^2} = \pm\frac{\mathcal{E}^2}{2-\mathcal{E}^2}$
	$\mathcal{AR}$	signed aspect ratio	$\mathcal{AR} = \text{sgn}\{B\}\frac{B}{A}$

Table 4.1: Parameters and measures of an ellipse

see a sketch of an ellipse in Figure 4.2. The geometric parameters are very intuitive quantities to describe an ellipse and can be seen in the sketch. The lengths of the *semi-major* A and *semi-minor* B axes are directly linked to the size and shape of the ellipse. They will also provide an intuition into the shape of the ellipse, for example, if the lengths of the major and minor will be equal we will have a circle versus if the lengths are very different we will most likely have an ellipse. The *orientation angle* θ tells us the angle between the x-axis and the major axis of the ellipse. The *ellipticity angle* χ is very useful in distinguishing the ellipse form its degenerate cases of circle and straight line. It characterizes scale-invariant shape of the ellipse characterised by its values presented in Table 4.2.

The measures of size and ellipticity give us a more relative and comparable measure. Among the parameters of size we are interested in the *area* of the ellipse,

Shape	χ
CW polarised circle	$\chi = -\pi/4$
CW (right-handed) polarised ellipse	$-\pi/4 < \chi < 0$
Linearly polarised (straight line)	$\chi = 0$
CCW (left-handed) polarised ellipse	$0 < \chi < \pi/4$
CCW polarised circle	$\chi = \pi/4$

Table 4.2: Shape of the ellipse based on values of the ellipticity angle χ

which can be calculated as $AB\pi$. We can use the *root-mean-square amplitude* κ to compare between ellipses [30], which provides a convenient simplification. The *amplitude* $\mathcal{A}$ is one of the size parameters that is also visible in a sketch of an ellipse (Figure 4.2) and its square is sometimes also referred to as *intensity* [53]. Both κ and $\mathcal{A}$ will give us information about the size, the bigger the value of these parameters the larger the oscillation will be.

We can use the parameters of the ellipse to characterise what type of elliptical signal we are dealing with, an ellipse, a circle or a straight line. It is clear that the circle will be a special kind of ellipse such that the semi-major and semi-minor are equal, $A = B$. Whereas, the ellipse will degenerate to a straight line if the semi-minor will be infinitely small. Nevertheless, it is useful to have concrete measures that will tell us the ellipticity of a signal and will also provide the direction of circulation, i.e. clockwise (CW) or counter-clockwise (CCW). In the above table we listed the *eccentricity*, $\mathcal{E}$ [27], and the *ellipse parameter*, λ [29], both measure the deviation from circularity. If they are equal to zero (0) we have a purely circularly polarised signal and if they are equal to 1 the signal is linearly polarised. Moreover, we can also define the signed *aspect ratio* [53], which is positive (+) for counter-clockwise (CCW) rotation and negative (-) for clockwise (CW) rotation.

4.2.3 The complex-valued vector as an ellipse

Above we have been focusing on the mapped ellipse, or in other words, an ellipse that is being mapped by the trajectory of a signal through time. Nevertheless, because we are working with complex-valued vectors and processes, we need to point out that there also exists another concept of ellipse. This ellipse is mainly present in the physics literature, and follows from the inherent properties of complex-valued vectors. We mention it here only for completeness purposes and to highlight the difference between such ellipses. Whereas in the remaining text we will be dealing

only with the mapped ellipse.

As per Lindell [32] any complex-valued vector $\mathbf{u}$ has a one-to-one correspondence with time-harmonic vectors $\mathbf{F}(t')$, as $\mathbf{u} \leftrightarrow \mathbf{F}(t') = \Re\{\mathbf{u}e^{i\omega t'}\} = \mathbf{F}_1 \cos\omega t' + \mathbf{F}_2 \sin\omega t'$, where $\Re\{\mathbf{u}\} = \mathbf{F}_1$ and $\Im\{\mathbf{u}\} = -\mathbf{F}_2$. The time-harmonic vector traces a trajectory of an ellipse, which degenerates to a line or a circle under certain conditions. The vector is linearly polarised, if $\mathbf{F}_1$ and $\mathbf{F}_2$ are parallel or at least one of them is null ($\mathbf{F}_1 \times \mathbf{F}_2 = 0$), and circularly polarised, if $\mathbf{F}_1^2 = \mathbf{F}_2^2$. In any other combination it is elliptically polarised, which means that every complex-valued vector $\mathbf{u}$ can be represented as an ellipse so that the real part $\Re\{\mathbf{u}\}$ defines the time origin value and the imaginary part $\Im\{\mathbf{u}\}$ defines the direction of the rotation.

From the above we can see that every complex-valued vector is represented as an ellipse in n-dimensional space that is defined by vectors of its real and imaginary parts. The ellipse trajectory in this sense is mapped out in time t' in the above definition, but we are not interested in this time, which does not have physical interpretation for the purposes of our analysis. Based on this any vector-valued complex time series has a an ellipse representation in this sense at every time point. However, these ellipses are unrelated and are just an instantaneous representation of the time series at every time point. In our work we are not interested in this instantaneous representations, but rather in time evolution of the time series which is represented by the mapped ellipses as introduced earlier. As already mentioned this particular way of looking on any complex-valued vector is not the subject of our research and will not be explored any further.

Chapter 5

Definitions of elliptical models

In this chapter we describe the definition of elliptical models already existing in the literature, and propose new ones. The chapter is divided into deterministic signals and stochastic processes and for each of these we look at the elliptical definitions in the time or frequency domains. This chapter includes our theoretical contribution to the research theory, by adding the definitions of the autocovariance ellipse and the forecast ellipse. We also provide a synthesis of the ellipse definitions at the end of this chapter, which we have not been able to find in any other literature.

5.1 Deterministic elliptical signals

5.1.1 Time domain definition

We need to understand the notion of a deterministic elliptical oscillations first and then we move to models for stochastic oscillations. One possible understanding of the deterministic elliptical oscillation is that such a signal maps out an ellipse in time, and we shall investigate how this concept extends to stochastic processes. These types of signals have been studied by different authors [6, 29, 59], and we will summarize our understanding of existing models in a single framework. We start by formalizing the notion of an elliptical trajectory.

Definition 5.1.1 (Elliptical trajectory [6])
The elliptical trajectory over time points $t \in \{0, \dots, T\}$ is the set of points $\{(x_t, y_t)_t\}$ satisfying the equation

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} A_x \cos(\omega t + \varphi_x) \\ A_y \cos(\omega t + \varphi_y) \end{bmatrix}, \tag{5.1.1}$$

and the set $\{(x_t, y_t)_t\}$ *maps out an ellipse in the* $x - y$ *plane.*

The above is a general definition of a deterministic signal that exhibits an elliptical trajectory as it evolves through time. The size and axes of the ellipse will be determined by the parameters A_x and A_y and the orientation will be set by the phase angles φ_x and φ_y. This signal will map out an ellipse in the $x-y$ plane, called the *mapped ellipse*. By eliminating the component t in the above Equation (5.1.1) one can obtain the equation of an ellipse as a geometric object

$$\frac{x_t^2}{A_x^2} + \frac{y_t^2}{A_y^2} - 2\frac{x_t y_t}{A_x A_y}\cos\varphi = \sin^2\varphi, \tag{5.1.2}$$

where $\varphi = \varphi_y - \varphi_x$. The above equation is a constrained relationship between x_t and y_t that is time homogeneous, unless φ, A_x and/or A_y change in time.

Here we need to distinguish between the terms elliptical trajectory and mapped ellipse. Elliptical trajectory is the trajectory of a signal through time that maps out an ellipse (as per Definition 5.1.1), whereas a mapped ellipse is an ellipse in $x - y$ plane that is mapped by the signal through time. The distinction can be nicely seen in Figure 4.1, where the elliptical trajectory is the trajectory of the signal as it evolves through time, the thick black curve. Whereas the mapped ellipse is the ellipse mapped by this signal onto the shaded 2-dimensional plane on the left-hand side.

We prefer to represent the bivariate real-valued signal as a univariate complex-valued signal. This yields a parametrisation (see below) that is similar to the one given in Equation (5.1.1), but the complex parametrisation simplifies the analysis. This parametrisation is in terms of the so-called rotary components, which rotate in the opposite directions and together map out an ellipse in the complex plane [59]. The parametric model that we will use can be easily obtained from Equation (5.1.1)

as

$$\begin{aligned} u_t &= x_t + iy_t \\ &= \underbrace{A_x \cos(\omega t + \varphi_x)}_{x_t} + i \underbrace{A_y \cos(\omega t + \varphi_y)}_{y_t} \\ &= \underbrace{\tfrac{1}{2}(A_x e^{i\varphi_x} + iA_y e^{i\varphi_y})}_{A_+ e^{i\theta_+}} e^{i\omega t} + \underbrace{\tfrac{1}{2}(A_x e^{-i\varphi_x} + iA_y e^{-i\varphi_y})}_{A_- e^{-i\theta_-}} e^{-i\omega t} \\ &= \underbrace{A_+ e^{i\theta_+} e^{i\omega t}}_{u_{+,t}} + \underbrace{A_- e^{-i\theta_-} e^{-i\omega t}}_{u_{-,t}} \\ &= u_{+,t} + u_{-,t}, \end{aligned} \tag{5.1.3}$$

which is a sum of a counterclockwise (CCW) $u_{+,t}$ (analytic) and a clockwise (CW) $u_{-,t}$ (anti-analytic) rotating phasors. Phasors or rotary components are two counter rotating elliptical motions, where each has its amplitude A_+ and A_- and phase θ_+ and θ_- [36]. The rotary component corresponding to positive frequencies has counterclockwise motion, and the one corresponding to negative frequencies has a clockwise motion. We will use the above parametric representation of an elliptical signal also to show that a certain signal has elliptical polarisation.

Parameters of the ellipse From the above parametrisations in both (5.1.1) and (5.1.3) we can easily obtain the main parameters of the ellipse and establish relationships between both of the parametrisations introduced. The equations of the axes and expressions for parameters are listed in Table 5.1. The semi-major axis A and the semi-minor axis B are defined in Equations (5.1.4) and (5.1.5), respectively. The equations relating both parametrisations are presented in Equations (5.1.6) to (5.1.9), together with some relationships between them in (5.1.10) and (5.1.11) [29]. Equation (5.1.12) denotes the amplitude, which together with the ellipticity angle (defined in Equation (5.1.15)) characterise the semi-major and semi-minor as shown in Equation (5.1.13). Further derivations of additional measures of size and ellipticity as described in Chapter 4.2.2 are shown in the Appendix B.

Table 5.2 outlines the main angles of the ellipse and the relationships between both parametrisations. The angle between the x-axis and the major axis is called the orientation or the azimuth of the ellipse, denoted as θ on the interval $(-\pi/2, \pi/2]$. It can be easily obtained from any of the two parametrisations as shown in Equation (5.1.14). The ellipticity angle $\chi \in (-\pi/4, \pi/4]$, which is very useful in helping

Axes

$$A = A_+ + A_- \tag{5.1.4}$$
$$B = |A_+ - A_-| \tag{5.1.5}$$

Relations between the different parametrisations

$$A_x^2 = (A_+^2 + A_-^2 + 2A_+A_- \cos 2\theta) \tag{5.1.6}$$
$$A_y^2 = (A_+^2 + A_-^2 - 2A_+A_- \cos 2\theta) \tag{5.1.7}$$
$$A_+^2 = ||A_+ e^{i\theta_+}||^2 = \tfrac{1}{4}(A_x^2 + A_y^2 + 2A_xA_y \sin\varphi) \tag{5.1.8}$$
$$A_-^2 = ||A_- e^{i\theta_-}||^2 = \tfrac{1}{4}(A_x^2 + A_y^2 - 2A_xA_y \sin\varphi) \tag{5.1.9}$$
$$A_+^2 + A_-^2 = \frac{A_x^2 + A_y^2}{2} = \frac{A^2 + B^2}{2} \tag{5.1.10}$$
$$A_+^2 - A_-^2 = A_xA_y \sin\varphi = \pm AB \tag{5.1.11}$$

Amplitude

$$\mathcal{A} = \sqrt{A_x^2 + A_y^2} = \sqrt{2A_+^2 + 2A_-^2} \tag{5.1.12}$$
$$A = \mathcal{A}\cos\chi \qquad \text{and} \qquad B = \mathcal{A}\sin\chi \tag{5.1.13}$$

Table 5.1: Definition of ellipse axes and their relationships

us to understand the polarisation and the direction of rotation of the ellipse, can be calculated as shown in Equation (5.1.15) [62]. Both of the above angles and expressions can be further simplified by introducing an auxiliary angle $\alpha \in (-\pi/2, \pi/2]$ [10], which does not have physical interpretation, but is useful for the analytic simplicity in Equation (5.1.16). For the purposes of practical calculations of the ellipse parameters and measures from any complex-valued signal we introduce the phase angle ϕ (Equation (5.1.17)), the average phase angle ϕ_a and the difference phase angle ϕ_d (both in Equation (5.1.18)). By doing this we can express the relationships between the phases defined in Equations (5.1.19) and (5.1.20) [29].

Orientation

$$2\theta = \theta_+ - \theta_- = \arctan \frac{\Im\{A_+ e^{i\theta_+}\}}{\Re\{A_+ e^{i\theta_+}\}} + \arctan \frac{\Im\{A_- e^{-i\theta_-}\}}{\Re\{A_- e^{-i\theta_-}\}}$$
$$= \arctan\left(\frac{2A_xA_y}{A_x^2 - A_y^2}\cos\varphi\right) = \arcsin\left(\frac{A_+^2 - A_-^2}{2A_+A_-}\cot\varphi\right) \tag{5.1.14}$$

Ellipticity angle

$$2\chi = \pm\arctan\frac{B}{A} = \pm\arcsin\frac{2AB}{A^2+B^2} = \arcsin\frac{2A_xA_y\sin\varphi}{A_x^2+A_y^2} = \arcsin\frac{A_+^2 - A_-^2}{A_+^2 + A_-^2} \tag{5.1.15}$$

$$\tan\alpha = \frac{A_y}{A_x} \quad \text{so that} \quad \begin{cases} \tan 2\theta = (\tan 2\alpha)\cos\varphi \\ \sin 2\chi = (\sin 2\alpha)\sin\varphi \end{cases} \tag{5.1.16}$$

Other angles and their relationships

$$\phi = (\theta_+ + \theta_-)/2 \tag{5.1.17}$$

$$\phi_a = (\varphi_x + \varphi_y + \pi/2)/2 \qquad \text{and} \qquad \phi_d = (\varphi_x - \varphi_y - \pi/2)/2 \tag{5.1.18}$$

$$\varphi_x = \phi + \Im\{A_+ e^{i\theta} + A_- e^{-i\theta}\} \quad \text{and} \quad \varphi_y = \phi + \Im\{A_+ e^{i\theta} - A_- e^{-i\theta}\} - \pi/2 \tag{5.1.19}$$

$$\theta_+ = \phi_a + \Im\{\ln(A_x e^{i\phi_d} + A_y e^{-i\phi_d})\} \quad \text{and} \quad \theta_- = \phi_a + \Im\{\ln(A_x e^{i\phi_d} - A_y e^{-i\phi_d})\} \tag{5.1.20}$$

Table 5.2: Definition of ellipse angles and their relationships

In Table 5.3 we present a concise way to calculate the parameters of an arbitrary elliptical signal. With this schematic representation we aim to facilitate the understanding and follow-up.

1. Take complex-valued signal $u_t = x_t + iy_t$, by using the Hilbert transform calculate analytic signals of its real x_t and imaginary y_t parts to obtain $x_{+,t}$ and $y_{+,t}$ (see Chapter 2.3).
2. Calculate the magnitudes (A_x and A_y) and phase angles (φ_x and φ_y) of the analytic signals $x_{+,t}$ and $y_{+,t}$, then using Equations (5.1.18) calculate the average phase ϕ_a and difference phase ϕ_d angles.
3. Calculate θ_+ and θ_- using Equations (5.1.20) in order to obtain the ellipse orientation as $\theta = (\theta_+ - \theta_-)/2$.
4. Calculate A_+ and A_- using Equations (5.1.8) and (5.1.9) in order to obtain the axes $A = A_+ + A_-$ and $B = |A_+ - A_-|$.
5. The remaining ellipse measures κ and λ (and other) can be obtained using above mentioned equations in Table 4.1.

Table 5.3: Procedure to calculate ellipse parameters from an arbitrary complex-valued signal [28]

Polarisation states In this chapter we talk about the 'shape' polarisation only (see definition of polarisation and discussion in Chapter 5.2.4). The trajectory of any signal defined by either of the parametrisations above, (5.1.1) and (5.1.3), will generally be an ellipse. There are three polarization states, and the most general corresponds to the elliptical polarization. Other states derived from the general are the linear and the circular (right or left) polarisation states, which can also be viewed as degenerate states of the elliptical polarisation. In Figure 4.1 we can see an elliptical polarisation, but we can easily imagine how can the ellipse become a straight line or a circle. Different combinations of the above mentioned parameters will result in the three different polarisation states. The polarisation will be linear (straight line), if the ellipticity angle is $\chi = 0$. The straight line will be horizontal if $A_y = 0$, vertical line if $A_x = 0$, or a 45° line if $A_x = A_y$ and $\varphi = \{0, \pi\}$. In terms of the parametrisation in (5.1.3) we will obtain a linear polarisation if $A_+ = A_-$.

The signal will be circularly polarised when the ellipticity angle is $\chi = \pm\pi/4$. The direction of the polarisation will be left circular (CCW), if $\varphi > 0$, and right circular (CW), if $\varphi < 0$. In terms of parameters in (5.1.1) we need the constraint $A_x = A_y$ and $\varphi = \pm\pi/2$, whereas in terms of parametrisation in (5.1.3) the parameters need to be $A_+ = 0$ or $A_- = 0$ in order to have circular polarisation state. In all other cases

the polarisation will be elliptical as the most general polarisation state. Elliptical polarisation can be left elliptical (CCW), if $\varphi > 0$, and right elliptical (CW), if $\varphi < 0$.

5.1.2 Frequency domain definition

We can also define the ellipse of a deterministic signal through its frequency representation, more precisely through the power and cross-power of the spectral matrix [6, 59]. For the definition of the ellipse of deterministic signals in the frequency domain we use the definitions of the composite and augmented SDF matrices (Equations (2.2.9) and (2.2.10)), and the parametrisation of the signal in (5.1.3). Since the signal is monochromatic (single-frequency) and deterministic the spectral density matrix will be non-zero only for f equal to frequency of oscillation ω and will sometimes be referred to as the coherency matrix ([6, 62]),

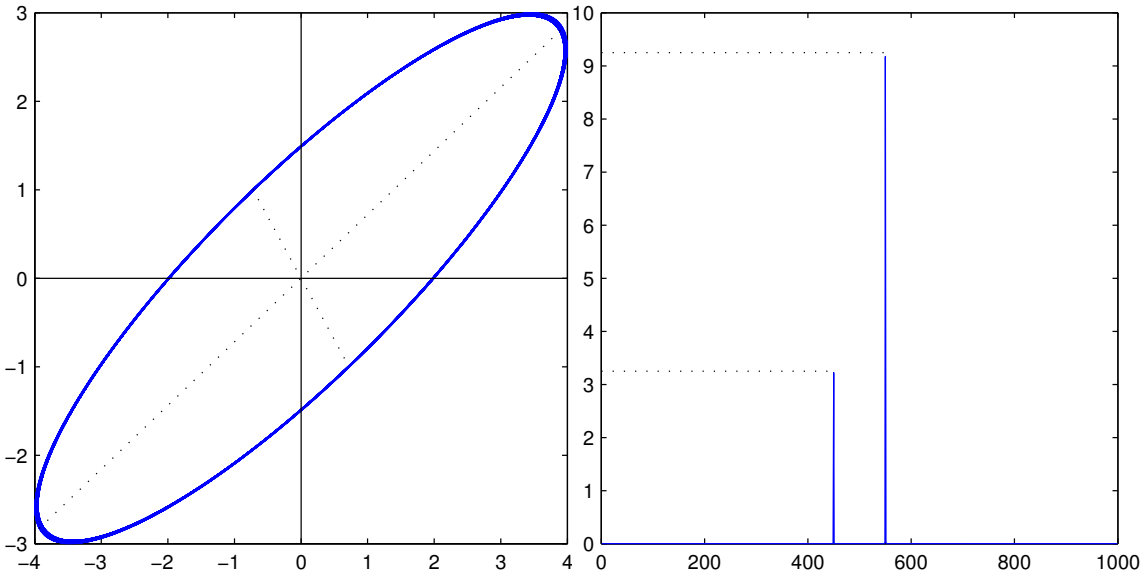
$$\mathbf{S}_v(\lambda) = \begin{bmatrix} S_{xx}(\lambda) & S_{xy}(\lambda) \\ S^*_{yx}(\lambda) & S_{yy}(\lambda) \end{bmatrix} \delta(\lambda - f) = \frac{1}{4} \begin{bmatrix} A_x^2 & A_x A_y e^{i\varphi} \\ A_x A_y e^{-i\varphi} & A_y^2 \end{bmatrix}, \tag{5.1.21}$$

$$\mathbf{S}_w(\lambda) = \begin{bmatrix} S_s(\lambda) & \tilde{S}_s(\lambda) \\ \tilde{S}^*_s(\lambda) & S_s(-\lambda) \end{bmatrix} \delta(\lambda - f) = \begin{bmatrix} A_+^2 & A_+ A_- e^{i2\theta} \\ A_+ A_- e^{-i2\theta} & A_-^2 \end{bmatrix}. \tag{5.1.22}$$

The above tells us that in the case of monochromatic deterministic signals we can find the parameters of the mapped ellipse the signal traces in time, in its spectral density matrix at the frequency of its single oscillation. In other words the ellipse is defined by its spectral density matrix at the frequency of oscillation and it is a non-random representation, which means that its parameters (A, B and θ) are constant. This is an interesting direct link between the time and the frequency domains, as these signals have the same ellipses in both domains. Unfortunately, as we will see later, this only holds in the case of deterministic signals.

5.1.3 Example

For illustration purposes we provide an example of an elliptical signal $\{s_t\}$ as defined by the parametrisation in (5.1.1) with parameters: $f = 50$, $A_x = 4$, $A_y = 3$, $\varphi_x = \pi/3 \approx 1.05$ and $\varphi_y = \pi/6 \approx 0.52$. These parameters are equivalent to the complex representation, the parametrisation in (5.1.3), with the following parameters: $A_+ = 3.04$, $A_- = 1.80$, $\theta_+ = 1.49$ and $\theta_- = 0.24$. Such signal will have an elliptical



(a) Trajectory of the signal in time (b) Fourier transform of the signal

Figure 5.1: Plots of the trajectory and the spectral density of a deterministic elliptical signal

trajectory and will trace a mapped ellipse with parameters: semi-major $A = 4.84$, semi-minor $B = 1.24$, orientation angle $\theta = 0.623$ and ellipticity angle $\chi = 0.25$ (indicating CCW polarisation). Its measures of size and ellipticity will be: root-mean-square amplitude $\kappa = 3.54$, amplitude $\mathcal{A} = 5$, Area $= 18.85$, ellipse parameter $\lambda = 0.88$, ellipticity $\mathcal{E} = 0.97$, aspect ratio $\mathcal{AR} = 0.26$. The mapped ellipse traced by the signal trajectory is plotted in Figure 5.1a.

Fourier transform of the above simulated signal $\{s_t\}$ is shown in Figure 5.1b. The coherency matrix of the composite vector v_t is calculated based on the Equation (5.1.21). The coherency matrix of the augmented vector w_t is calculated based on the Equation (5.1.22)). Both have a direct relationship to the parametrisation in (5.1.3) and are easily calculated using the parameters, if known, or otherwise using the procedure described in Table 5.3. Both of the coherency matrices of the simulated signal are shown below

$$\mathbf{S}_v(\lambda) = \begin{bmatrix} 4.0 & 2.6 + 1.5i \\ 2.6 - 1.5i & 2.25 \end{bmatrix}, \qquad \mathbf{S}_w(\lambda) = \begin{bmatrix} 9.25 & 1.75 + 5.2i \\ 1.75 - 5.2i & 3.25 \end{bmatrix}.$$

From these matrices we can clearly see that the ellipse parameters in the time domain are the same as the parameters we can obtain from the frequency domain.

This shows the direct link between the ellipse trajectory definition in the time domain and in the frequency domain ellipse definition of monochromatic deterministic signals [62]. We can conclude that for a monochromatic deterministic signal the relationship between the time domain mapped ellipse and the frequency domain ellipse is perfect, mapping out the same ellipse. From time series perspective the issue is that the elliptical trajectory is only visible for deterministic signals. For their stochastic counterparts the trajectory of the ellipse is not visible, because of the the fact that it is zero-mean and also because it is distorted at every step, due to the effect of the innovation noise.

5.2 Stochastic stationary elliptical time series

5.2.1 Frequency domain ellipse

In the previous chapter we discussed deterministic elliptical signals, which proved to be a useful parametrisation of the signal. The elliptical parametrisation of a signal provides us with a convenient summary of deterministic signals, just by using a few parameters. We would like to find such parametrisation also for random processes, mainly for stationary complex-valued random processes denoted by $\{U_t\}$, where $t = 1, \ldots, T$. This will require extending the concept from deterministic signals to stochastic processes. Unfortunately, in the case of stationary stochastic processes the elliptical definition is not so straight forward as in the deterministic case.

We are dealing with time series data where the time sequence of the process (trajectory) is important, but the trajectory of a zero-mean stochastic process will in general be just a cloud of points (see Figure 6.2), as opposed to a ring of points or any other structure. Two properties of the process, randomness and zero-mean, will prevent any structure in the time evolution of the process. In other words, it is not possible to describe an ellipse in the time trajectory of a stochastic process similar to what we did in Definition 5.1.1 for deterministic signals. Hence most authors, so far, have described the ellipse of stochastic processes in the frequency domain. The downside of this approach is that it is a random ellipse representation and is less intuitive [59, 53, 69].

In order to provide a definition of the ellipse for stochastic processes we need to resort to the basic frequency domain theory, as already introduced in Chapter 2.2.3. Every zero-mean stationary random processes can be written with its

spectral representation, as established in Definition 2.2.5 and Equation (2.2.4). The spectral representation of a stationary stochastic process is defined for frequencies $f \in [-1/2, 1/2)$ as

$$U_t = \int_{-\infty}^{\infty} e^{i2\pi ft} dZ_U(f), \tag{5.2.1}$$

where $Z_s(f)$ is a complex-valued spectral process of U_t with orthogonal increments $dZ_U(f)$ (see Chapter 2.2.3 for more details) [8, 42]. From this spectral representation we can derive the frequency domain ellipse, as per Definition 5.2.1 below.

Definition 5.2.1 (Frequency domain ellipse [53, 59])
Spectral representation can simply be rewritten in terms of non-negative frequencies $f \geq 0$ only, which will result in representation as follows

$$U_t = \int_0^{\infty} \Big(dZ_U(f) e^{i2\pi ft} + dZ_U(-f) e^{-i2\pi ft} \Big). \tag{5.2.2}$$

The above equation can be seen as a superposition of ellipses, or in other words, the summation of two opposite rotating phasors ($e^{i2\pi ft}$ and $e^{-i2\pi ft}$) with potentially different amplitudes ($dZ_U(f)$ and $dZ_U(-f)$). At each frequency, $f \geq 0$, the contribution to $\{U_t\}$ can then be expressed as

$$\begin{aligned} U_t(f) &= dZ_U(f) e^{i2\pi ft} + dZ_U(-f) e^{-i2\pi ft} \\ &= \underbrace{|dZ_U(f)| e^{i \arg\{dZ_U(f)\}} e^{i2\pi ft}}_{U_t(f_+)} + \underbrace{|dZ_U(-f)| e^{i \arg\{dZ_U(-f)\}} e^{-i2\pi ft}}_{U_t(f_-)}. \end{aligned} \tag{5.2.3}$$

This shows that at every frequency there is a random ellipse defined by the orthogonal increment process through the rotating phasors. The orthogonal increment process at positive frequency $\{dZ_U(f)\}$ defines the CCW phasor, and the orthogonal increment process at negative frequency $\{dZ_U(-f)\}$ defines the CW phasor. Since there is an ellipse for each frequency, $f \geq 0$, the whole set of ellipses across all frequencies yields a family of curves.

We can easily see the similarity between the parametric elliptical model for deterministic signals in Equation (5.1.3) and the frequency domain ellipse in (5.2.3). The random amplitude $|dZ_U(f)|$ has a similar interpretation to the deterministic A_+ ($|dZ_U(f)|$ similar to A_-), and the random phase $\arg\{dZ_U(f)\}$ has a similar interpretation to the deterministic orientation of the signal θ_+ ($\arg\{dZ_U(f)\}$ similar

to θ_-). Taking parallels with the deterministic signal in the previous chapter we can write the *random* parameters of the ellipse at each frequency $f \geq 0$ in terms of the orthogonal increment process, which clearly shows the random nature of the representation [53]:

- the random semi-major axis $A(f) = |dZ_U(f)| + |dZ_U(-f)|$,
- the random semi-minor axis $B(f) = ||dZ_U(f)| - |dZ_U(-f)||$,
- the random area of the ellipse $A(f)B(f)\pi = ||dZ_U(f)|^2 - |dZ_U(-f)|^2|\pi$, and
- the random angle between the major axis and the x-direction (orientation) $\theta(f) = \arg\{dZ_U(f)dZ_U(-f)\}/2$.

The above shows that the phase and orientation, and so the geometrical properties of the ellipse, vary at each frequency and are random. The frequency domain ellipse needs to be understood in terms of a family of ellipses, because there is a different curve at each frequency. Whereas in the case of the deterministic monochromatic process in previous chapter, the spectral (coherency) matrix in (5.1.22) was non-zero only at the single frequency. That gave constant quantities that described the mapped ellipse of the signal and so it was non-random.

5.2.2 Forecast ellipse

The above definition of the frequency domain ellipse (Definition 5.2.1), varies across frequencies and is random in itself. Both of these characteristics are limitations to the usefulness of such representation. Such quantity needs to be estimated, as it is based on the orthogonal increment process of which we cannot obtain samples [65]. Thus, we would prefer to define an ellipse as a deterministic representation.

In order to achieve that, we look at the forecast of the stochastic process. As introduced in Chapter 3.2, the forecast can be defined as the conditional expectation of the future values knowing the present and all past values. Using the forecast we will define what we call forecast ellipse for some models, such as first-order autoregressive AR(1) models. The forecast ellipse is the elliptical trajectory of the forecast, as a function of l forecast-time steps ahead in the future, which maps out an ellipse. The basic intuition behind the forecast ellipse is the expected trajectory of the process in the future. In other words, in the future we expect the trajectory of the process to map out an ellipse, but due to randomness this trajectory will be

distorted. For some models we will be able to observe an elliptical trajectory of the forecast function so that the expected trajectory of the process in the future will trace a mapped ellipse.

Definition 5.2.2 (Forecast ellipse)
The forecast ellipse is the set of points $\{(x_l, y_l)_l\}$ *over forecast-time* $l \in \{0, \ldots, L\}$ *that will exhibit an elliptical trajectory (Definition 5.1.1). Where* x_l *and* y_l *are the real and imaginary parts of the forecast function* $\widehat{U}_t(l)$ *(or of one entry of the forecast vector in multivariate case) defined in Equations* (3.2.1) *and* (3.2.2)*, as*

$$\begin{aligned} \widehat{U}_t(l) &= E[U_{t+l} \mid U_t, U_{t-1}, U_{t-2}, \ldots] \\ &= E[\phi U_{t+l-1} + \epsilon_t \mid U_t, U_{t-1}, U_{t-2}, \ldots] \\ &= \phi \widehat{U}_{t+l-1} = \phi^l U_t. \end{aligned} \tag{5.2.4}$$

The set $\{(x_l, y_l)\}$ *maps out an ellipse in the complex plane for the forecast function of the time series, or for every entry of the forecast vector in multivariate case.*

The above forecast ellipse definition of a process can be seen as the elliptical trajectory of the forecast function $\{\widehat{U}_{t+l}\}$ that traces a mapped ellipse as we move through forecast-time l, with non-random parameters. This means that as we advance through forecast-time l the trajectory that the forecast functions maps is of elliptical shape. For complex-valued multivariate processes we have a forecast ellipse in the complex-plane for every entry of the forecast vector separately. The forecast is a deterministic quantity, because it is observed after expectation has been taken. That means that also the trajectory of the forecast function in non-random. In other words, the forecast is a deterministic function of time that is determined by the parameters of the process. Some, but not all, stochastic processes will have a forecast ellipse, as not all the forecast functions of every process will map out an ellipse. Moreover, the forecast function is not easy to express in a nice form for all stochastic processes. We need a parametric model for that and the easiest way is to express it for an AR(1) model. Later we will illustrate this with a concrete example of a parametric model that will exhibit these properties.

5.2.3 Autocovariance ellipse

Similar to the forecast ellipse above, we can define the autocovariance ellipse by observing the autocovariance function (ACVF) of a process as a function of lag τ. In some cases the ACVF will exhibit an elliptical trajectory that traces a mapped ellipse in the complex plane. In this definition we will observe the trajectory of each entry of the autocovariance matrix as a function of lag τ. For complex-valued process the autocovariance function is defined in Definition 2.2.2 and the autocovariance matrix in Equation (2.2.1). The autocovariance ellipse means that the separate entries of the autocovariance matrix will exhibit elliptical trajectories as we move forward through lag τ. Again, this is the characteristic only of certain models, and is most conveniently shown with a parametric model.

Definition 5.2.3 (Autocovariance ellipse)
The autocovariance ellipse is the set of points $\{(x_\tau, y_\tau)_\tau\}$ *over lag-time* $\tau \in \{0, \dots, T\}$ *that will exhibit an elliptical trajectory (Definition 5.1.1). Where* x_τ *and* y_τ *are the real and imaginary parts of one entry of the autocovariance matrix* $\mathbf{\Gamma}_W(\tau)$*, as defined in Equation* (2.2.1)*. The set* $\{(x_\tau, y_\tau)\}$ *maps out an ellipse in the complex plane for every entry of the autocovariance matrix.*

For illustration we take a univariate complex-valued time series $\{U_t\}$ and use the augmented vector $\mathbf{W}_t = [U_t, U_t^*]^{\mathrm{T}}$ to define the autocovariance matrix as,

$$\mathbf{\Gamma}_W(\tau) = E[\mathbf{W}_{t+\tau}\mathbf{W}_t^{\mathrm{H}}] = \begin{bmatrix} \gamma_U(\tau) & \tilde{\gamma}_U(\tau) \\ \tilde{\gamma}_U^*(\tau) & \gamma_U^*(\tau) \end{bmatrix}. \tag{5.2.5}$$

In this matrix the autocovariance ellipse will be represented by the elliptical trajectory that traces a mapped ellipse in the complex plane, defined by the real and imaginary parts for every entry of the autocovariance matrix. This means that each entry of this matrix (e.g. $\gamma_U(\tau)$, $\tilde{\gamma}_U(\tau)$, etc.) yields an autocovariance ellipse. These ellipses are parametrised by the process parameters and tell us the evolution of the dependence as we move through lags τ.

The two main advantages of both Definitions 5.2.2 and 5.2.3 of ellipses of a stochastic time series processes is that they are deterministic quantities and are defined in time domain. On one hand, they have a very simple and intuitive interpretation and are very useful in the estimation or analysis of a process. On the

other hand, the disadvantage is that these definitions are specific to certain parametric models and not all processes will exhibit the forecast and/or autocovariance ellipse. The forecast or autocovariance will have an elliptical structure only for certain stationary stochastic processes. Since both, the forecast and the autocovariance ellipses, are derived from the shape of the trajectory of the forecast function or the ACVF, it is evident that not all such functions we will be able to parametrise in the form of an ellipse as in Equation (5.1.3). Thus in order to present this and visualise we need to develop a parametric model that will exhibit these properties.

5.2.4 Polarisation and coherence

The last way of looking at elliptical processes is through determining the polarisation of a wave [57]. Based on Collett [10] the definition of polarization comes from optics and physics, where it has been defined for light. Light consists of two oppositely polarised rays with opposite behaviour of intensity, in other words that the two rays of light are polarised. The polarisation of any wave field can be determined by examining its spectral matrix [24], which tells us the degree of polarisation.

Here we should point out that there are (at least) two uses of the term polarisation. Based on the discussion by Schreier & Scharf [62, p. 211] one use refers to the definition of polarisation itself and the other to the 'shape polarisation'. The latter characterises the shape of ellipse we observe, whereas a polarised wave as defined below in Definition 5.2.4 does not need to have a shape of an ellipse.

Definition 5.2.4 (Polarisation ellipse [34, 62])
The polarisation ellipse is defined as the degree of polarisation by comparing the power of the completely polarised part to the total power of the process. The degree of polarisation is defined from the augmented spectral matrix $\mathbf{S}_W(f)$ *(Equation* (2.2.10)*) for a complex-valued process* $\{U_t\}$*, or composite spectral matrix* $\mathbf{S}_V(f)$ *(Equation* (2.2.9)*) for a bivariate real-valued process* $\mathbf{V}_t = [X_t, Y_t]^{\mathrm{T}}$*, as*

$$\mathcal{P}(f) = \frac{\lambda_1(f) - \lambda_2(f)}{\lambda_1(f) + \lambda_2(f)}, \tag{5.2.6}$$

where $\lambda_1(f)$ *and* $\lambda_2(f)$ *are the two eigenvalues of the augmented spectral matrix.*

The degree of polarization is defined on $0 \leq \mathcal{P}(f) \leq 1$, where $\mathcal{P}(f) = 0$ means that the signal is unpolarised. $\mathcal{P}(f) = 1$ means that it is completely polarised

(pure state) at frequency f, and so the spectral matrix will have only one non-zero eigenvalue [57]. Elliptical polarisation occurs when the off-diagonal elements of the spectral matrix are complex-valued ($\Im\{\tilde{S}_U(f)\} \neq 0$), whereas circular polarisation occurs if additionally the diagonal elements are equal ($S_U(f) = S_U(-f)$) [34].

Coherence, is similar to a correlation coefficient and should not be confused with the degree of polarisation. In general the coherence is defined for multivariate signals and can be seen as a correlation coefficient in the frequency domain. In this case of polarisation analysis it also tells us the relationship between the counter-clockwise and clockwise phasors [62, p. 211] and is defined as

$$\mathcal{C}^2(f) = \frac{|\tilde{S}_U(f)|^2}{S_U(f)S_U(-f)}, \quad \text{or} \quad \mathcal{C}^2(f) = \frac{|S_{XY}(f)|^2}{S_X(f)S_Y(f)}. \tag{5.2.7}$$

Coherence is bounded by $\mathcal{C}^2(f) \leq 1$, and is defined to be equal to 1, if either $S_U(f) = 0$ or $S_U(-f) = 0$ (or $S_X(f) = 0$ or $S_Y(f) = 0$). If the signal is completely polarised, the coherence will be same to the polarisation $\mathcal{C}^2(f) = 1$, as $|\tilde{S}_U(f)|^2 = S_U(f)S_U(-f)$ (or $|S_{XY}(f)|^2 = S_X(f)S_Y(f)$). The difference between polarisation and coherence is that the degree of polarisation separates the spectral matrix into the polarised and unpolarised part, whereas coherence separates is onto coherent and incoherent parts [24]. The relationship is of the form $\mathcal{P}^2(f) \geq \mathcal{C}^2(f)$, where the equality holds if $S_U(f) = S_U(-f)$.

5.3 Summary of ellipse definitions

We have seen several different definitions of 'the ellipse'. Table 5.4 summarises the possible ellipse definitions characterising both deterministic signals and stochastic processes. Some of them were presented form the literature and some of them were proposed in this thesis, i.e. the forecast and the autocovariance ellipse. The table presents a view on one hand based on whether the definitions are in the time or frequency domain, and on the other hand if the ellipse definitions are deterministic or random. The distinction is made based on the nature of the ellipse representation, rather than the nature of the signal. So it does not provide the distinction between deterministic signals and random process. Certain concepts apply only to deterministic signals or stochastic processes, or to both.

In general a deterministic ellipse definition is more useful, as the parameters of such a representation are non-random. Time domain definitions are generally

more natural compared to definitions in the frequency domain. For this reason the non-random elliptical trajectory in time domain is most probably the most intuitive definition of an elliptical signal out of all the definitions analysed. Unfortunately, only deterministic signals will exhibit a deterministic elliptical trajectory. Stochastic processes will have trajectories that do not map out ellipses. In other words time trajectories of stochastic process will be a cloud of points and will not have an elliptical structure in time.

The frequency domain ellipse, that is to say, the representation through orthogonal increments, is random in itself, because it is based on a random orthogonal increment process $\{dZ_U(f)\}$. It gives us a view on a process through a family of random ellipses, where the random process is an aggregation of all the ellipses at each frequency. Such definition is not that practical and easy to understand. On the other hand among the spectral definitions we also have the definition of polarisation and coherence. The polarization ellipse tells us about the degree of polarization of the spectral matrix, but not about the 'shape' polarization of the process. By the shape polarisation we mean that the process exhibits an elliptical trajectory in either its time trajectory or some other function of the process. Thus, the polarization ellipse is less useful, because what we are interested in is the shape polarisation of a process.

From this discussion we can see the benefit of proposing two new, non-random time domain definitions, the forecast ellipse and the autocovariance ellipse. Both are deterministic representations, since they are functions of the process after the expectation has been taken. They are as well defined in time domain and have a very intuitive interpretation. The forecast ellipse is the expected elliptical trajectory of the process as we move forward in forecast-time into the future. The autocovariance ellipse is the elliptical trajectory of the autocovariance function as we move through lags τ. In other words, it tells us about the elliptical structure of the dependence in time. Both of these ellipses are functions of forecast-time or lag-time, and will depend on the parametrisation of the model, which means that not all such functions will exhibit an elliptical trajectory. That is to say, we need certain parametric model that will produce a function with elliptical trajectory in its ACVF and forecast function.

We conclude that the definition of 'the ellipse' of a process is not completely trivial and that there is not a single concept that would fit best with the notion of recovering the elliptical properties of the time series. We have presented several

definitions and each tells a different story. We have also seen that the distinction between deterministic signals and stochastic processes is very important. The ellipse for deterministic signals is relatively easy to define, whereas the ellipse for stochastic processes is much more intricate.

Ellipse representation vs. domain	Deterministic	Random
Time	• Elliptical trajectory (Definition 5.1.1) • Forecast ellipse (Definition 5.2.1) • Autocovariance ellipse (Definition 5.2.3)	Random trajectories which do not map ellipses due to stochasticity and zero-mean (Figure 6.2)
Frequency	Polarization ellipse (Definition 5.2.4)	Frequency domain ellipse (Definition 5.2.1)

Table 5.4: Summary of ellipse definitions

Chapter 6

Stochastic elliptical parametric models

In this chapter we will talk about different parametric time series models for modelling stochastic oscillations. We will look at models that model the cyclical component part of a structural time series model (STSM), which is used to model stochastic oscillations. We also aim to illustrate the concepts developed in the previous chapter. The main scope of this chapter is to construct stationary stochastic oscillation models and to illustrate the definitions of ellipses proposed in the previous chapter. We propose a bivariate complex-valued vector autoregressive model that will fully illustrate the concepts of autocovariance and forecast ellipses.

6.1 Introduction to stochastic oscillation models

We can model stochastic oscillations by using the pseudo-periodic behaviour of time series models described in Chapter 3.1.2. As described, the usual way to do that is by using an AR(2) process with complex roots. The downside of AR(2) models is that the concepts such as the ACVF, SDF and forecast are more intricate to work with and harder to extend to higher dimensions, compared to AR(1) models. Hence, we will explore opportunities to obtain an oscillation with a simpler structure. One possible alternative is to construct a bivariate model (e.g. [20, 54]). A VAR(1) model will exhibit an oscillation, if the coefficient matrix is composed of a rotation matrix, such as $\mathbf{R}_\theta$ from Equation (2.1.3). Such model has been, for example, developed by Harvey and Koopman [20] and presented below in Equation (6.1.1).

The nice property of a stochastic oscillation defined in this way is that the peak of the spectral density will be at the angular frequency same to the rotation angle used in the coefficient matrix of the VAR(1) model.

Harvey [19] and Harvey & Koopman [20] have developed a way to model cyclical time series and Rünstler [54] has further expanded the model to take into account the shifts among stochastic cycles. The model is a bivariate real-valued VAR(1) for n-dimensional vertor, as

$$\begin{bmatrix} \psi_t^{(j)} \\ \breve{\psi}_t^{(j)} \end{bmatrix} = \rho^{(j)} \begin{bmatrix} \cos(\lambda^{(j)}) & \sin(\lambda^{(j)}) \\ -\sin(\lambda^{(j)}) & \cos(\lambda^{(j)}) \end{bmatrix} \begin{bmatrix} \psi_{t-1}^{(j)} \\ \breve{\psi}_{t-1}^{(j)} \end{bmatrix} + \begin{bmatrix} \kappa_t^{(j)} \\ \breve{\kappa}_t^{(j)} \end{bmatrix}, \tag{6.1.1}$$

where $0 < \rho^{(j)} < 1$ is the damping factor in $\mathbb{R}$, $0 < \lambda^{(j)} < \pi$ is the frequency in $\mathbb{R}$ for series $j = 1, \dots n$, and the innovations $\boldsymbol{\kappa}_t^{(j)} = [\kappa_t^{(j)}, \breve{\kappa}_t^{(j)}]^{\mathrm{T}}$ are i.i.d. $N(\mathbf{0}, \sigma_\kappa^{(j)2}\mathbf{I}_2)$. Only the damping factor $\rho^{(j)}$ influences the stationarity of the model in (6.1.1), because the determinant of the rotation matrix is equal to zero. So we limit the damping factor to the interval $(0, 1)$, which ensures stationarity.

In the framework proposed by Harvey and Rünstler this model is not meant to be used for bivariate data, because the element $\{\breve{\psi}_t^{(j)}\}$ is just as an auxiliary process. It is introduced as a way to make the model bivariate, VAR(1), and so obtain pseudo-periodic behaviour for $\{\psi_t^{(j)}\}$. That is to say, the process of interest is only the first element of the bivariate series $\{\psi_t^{(j)}\}$, which means that for every process of interest j one needs to model a bivariate system.

The second aim of this chapter is to illustrate and support the theoretical elliptical definitions from the previous chapter with a parametric model. We aim to show how one can observe the forecast ellipse (Definition 5.2.2) and the autocovariance ellipse (Definition 5.2.3) in a parametric representation of a time series model. We also believe that the complex-valued family of models allow for more flexibility, so we will attempt to look at defining stochastic oscillations with complex-valued models.

We will begin by looking at univariate complex-valued $\mathbb{C}$AR(1) models that exhibit a single non-zero frequency peak and demonstrably possess oscillatory structure. Later we will extend the model into a bivariate complex-valued $\mathbb{C}$VAR(1), which exhibits two peaks in its SDF. This model is more flexible in terms of the behaviour of its SDF, ACVF and forecast functions, as well as exhibits some interesting elliptical representations that we will discuss later. Additional possible extension that we will explore is to introduce *stretching* or an anisotropic transformation of

the coefficient matrix of the $\mathbb{C}$VAR(1). In this case the model is unevenly rotated at every step, which introduces some additional structure into the SDF, ACVF and forecast, but unfortunately becomes mathematically very complex. In this case the ACVF and forecast will not exist in closed form solution, which makes these models hard to work with.

All the models below are interconnected and we should see them as extended or restricted version of each other. Our main model that we propose is the the bivariate $\mathbb{C}$VAR(1) model in Chapter 6.3. The most general model described here is the $\mathbb{C}$VAR(1) with the stretching factor in Chapter 6.4.1, which is an extension of our bivariate $\mathbb{C}$VAR(1), by adding the extension factor. The real-valued models described here are both restrictions of the previous complex-valued models, such that they accommodate only real-valued time points. The bivariate real-valued VAR(1) in Chapter 6.4.2 is a restriction of the $\mathbb{C}$VAR(1) with the stretching factor, whereas the model in Equation (6.1.1) above is a restriction of the general bivariate $\mathbb{C}$VAR(1) model.

6.2 Univariate $\mathbb{C}$AR(1)

As said earlier in order to obtain periodic behaviour and a stochastic oscillation model, the roots of the characteristic polynomial need to be complex-valued. In real-valued models we can obtain complex-roots only with an AR(2), but in case of complex-valued models we can have one complex-valued root of the characteristic polynomial already with an $\mathbb{C}$AR(1) model. The difference is that the real-valued AR(2) has two complex roots that are conjugate pairs, but the $\mathbb{C}$AR(1) can only have one complex root. This is the reason that its SDF has only one peak at one side of the frequency spectrum.

We write a simple $\mathbb{C}$AR(1) model as

$$\psi_t = \varphi_\lambda \psi_{t-1} + \epsilon_t, \tag{6.2.1}$$

where ψ_t, ϵ_t and $\varphi_\lambda \in \mathbb{C}, \forall t$, and the characteristic polynomial is $\Phi(B) = (1 - \varphi_\lambda B)$. In order for the model to exhibit an oscillation at frequency λ, we propose the coefficient to be $\varphi_\lambda = \rho_1 \cos\lambda + i\rho_2 \sin\lambda$, where ρ_1 and ρ_2 are so called damping factors. For the process to be stationary we need to ensure that $|\varphi_\lambda| < 1$. This form of the coefficient φ_λ allows us to control the frequency peak through the parameter

λ, which is very convenient. Since the characteristic polynomial has only one root, the frequency peak will be only on one, at positive or negative frequency range.

6.2.1 Autocovariance function of univariate $\mathbb{C}$AR(1)

We are mainly interested to see if the behaviour of the ACVF is same to the one described in the definition of the autocovariance ellipse (Definition 5.2.3). For that reason we first analyse the form of the ACVF as a function of lag τ. We can write the ACVF of the $\mathbb{C}$AR(1) time series model in (6.2.1) for positive and negative τ as

$$\tau \geq 0: \qquad \gamma_\psi(\tau) = \mathbb{E}\{\psi_{t+\tau}\psi_t^*\} = \varphi_\lambda^\tau \frac{\sigma_\epsilon^2}{1-|\varphi_\lambda|^2}, \tag{6.2.2}$$

$$\tau < 0: \qquad \gamma_\psi(\tau) = \gamma_\psi^*(|\tau|) = \varphi_\lambda^{*|\tau|} \frac{\sigma_\epsilon^2}{1-|\varphi_\lambda|^2}. \tag{6.2.3}$$

Remark 6.2.1. *We want to understand how the ACVF, $\gamma_\psi(\tau)$, behaves as a function of lag-time τ (in light of the Definition 5.2.3). W.l.o.g. we look only at the case where $\tau \geq 0$. In general, we see that τ is in the power of the complex parameter φ_λ. The ACVF can be expressed in the following way*

$$\begin{aligned} \gamma_\psi(\tau) &= (\rho_1 \cos\lambda + i\rho_2 \sin\lambda)^\tau \gamma_\psi(0) \\ &= \left(\sqrt{\rho_1^2 \cos^2\lambda + \rho_2^2 \sin^2\lambda}\right)^\tau e^{i\chi\tau}\gamma_\psi(0) \\ &= |\varphi_\lambda|^\tau e^{i\chi\tau}\gamma_\psi(0), \end{aligned} \tag{6.2.4}$$

where $\chi = tan^{-1}\left\{\frac{\rho_2 \sin\lambda}{\rho_1 \cos\lambda}\right\} = tan^{-1}\left\{\frac{\rho_2}{\rho_1} tan\lambda\right\}$. The above shows that the autocovariance ellipse will always trace the shape of a damped circle as we move through the lags. The $e^{i\chi\tau}$ part generates the motion of moving around a circle, this means that the angle χ indicates where the circle starts, but does not influence the trajectory of the ACVF in respect of the movement through lags. The $|\varphi_\lambda|^\tau$ part, for stationary processes ($|\varphi_\lambda| < 1$), causes the damping of the circle. It means that the circle is damped proportionally to the magnitude of the coefficient of the $\mathbb{C}$AR(1) process, $|\varphi_\lambda|$. If $|\varphi_\lambda|$ close to the unity then the damping is slower and the dependence is more persistent, whereas if $|\varphi_\lambda|$ is close to zero then the circle dampens very quickly and the dependence disappears quickly as we move in lag-time τ.

6.2.2 Forecast function of univariate $\mathbb{C}$AR(1)

We will perform a similar analysis also by looking at the l-step ahead forecast $\{\widehat{\psi}_t(l)\}$ of the time series $\{\psi_t\}$ in Equation (6.2.1). We want to analyse the behaviour of the forecast and are interested, if this process exhibits a forecast ellipse. We define the 1-step ahead forecast as the conditional expectation of the process by knowing the past values until present as defined in Chapter 3.2, which is similarly extended for any l-steps ahead. The forecast function of the $\mathbb{C}$AR(1) time series model in (6.2.1) for $l \in \mathbb{N}$, is

$$\widehat{\psi}_t(l) = \mathbb{E}[\psi_{t+l} \mid \psi_t, \psi_{t-1}, \psi_{t-2}, \ldots] = \varphi_\lambda^l \psi_t. \tag{6.2.5}$$

Remark 6.2.2. *We want to understand the trajectory of the forecast function as we move through forecast-time in the future (in light of the Definition 5.2.2). We look at the forecast function* $\widehat{\psi}_t(l)$ *in* (6.2.5) *for* $l \in \mathbb{N}$ *steps ahead as a function of the* l*-steps*

$$\widehat{\psi}_t(l) = \varphi_\lambda^l \psi_t = |\varphi_\lambda|^l e^{i \arg\{\varphi_\lambda\} l} \psi_t. \tag{6.2.6}$$

From the above equation we can see that the forecast as a function of l *will map a trajectory of a damped circle, similar to the ACVF analysed earlier in Remark 6.2.1. The difference between Equations* (6.2.4) *and* (6.2.6) *is only in the constants (*$\gamma_\psi(0)$ *vs* ψ_t*), which influence the initial radius of the circle. The coefficient* $|\varphi_\lambda|^l$ *makes the circle to be damped, if the magnitude is restricted to be smaller than one as is the case in the stationary case,* $|\varphi_\lambda| < 1$*. The exponential* $e^{i \arg\{\varphi_\lambda\} l}$ *causes the circular movement. In other words we can conclude that the forecast ellipse is a damped circle in the case of the* $\mathbb{C}$*AR(1) time series in Equation* (6.2.1).

6.2.3 Spectral analysis of univariate $\mathbb{C}$AR(1)

The spectral density function of a complex-valued $\mathbb{C}$AR(1) time series can be easily obtained using the spectral theory introduced in Chapter 3.1.1. The SDF of the time series model in (6.2.1) with complex-valued parameter φ_λ can be, by using the Equation (3.1.2), expressed as

$$S(f) = \frac{\sigma_\epsilon^2}{|1 - \varphi_\lambda e^{-i2\pi f}|^2} = \frac{\sigma_\epsilon^2}{1 - 2\rho_1 \cos(2\pi f) - 2\rho_2 \sin(2\pi f) + |\varphi_\lambda|^2}, \tag{6.2.7}$$

where $\varphi_\lambda = \rho_1 \cos\lambda + i\rho_2 \sin\lambda$, with ρ_1 and ρ_2 are such that $|\varphi_\lambda| < 1$ to satisfy stationarity of the model, and $f \in [-1/2, 1/2]$.

We can find the frequency at which the SDF above will achieve its maximum so that we differentiate the denominator with respect to the frequency f and set it to zero, to obtain

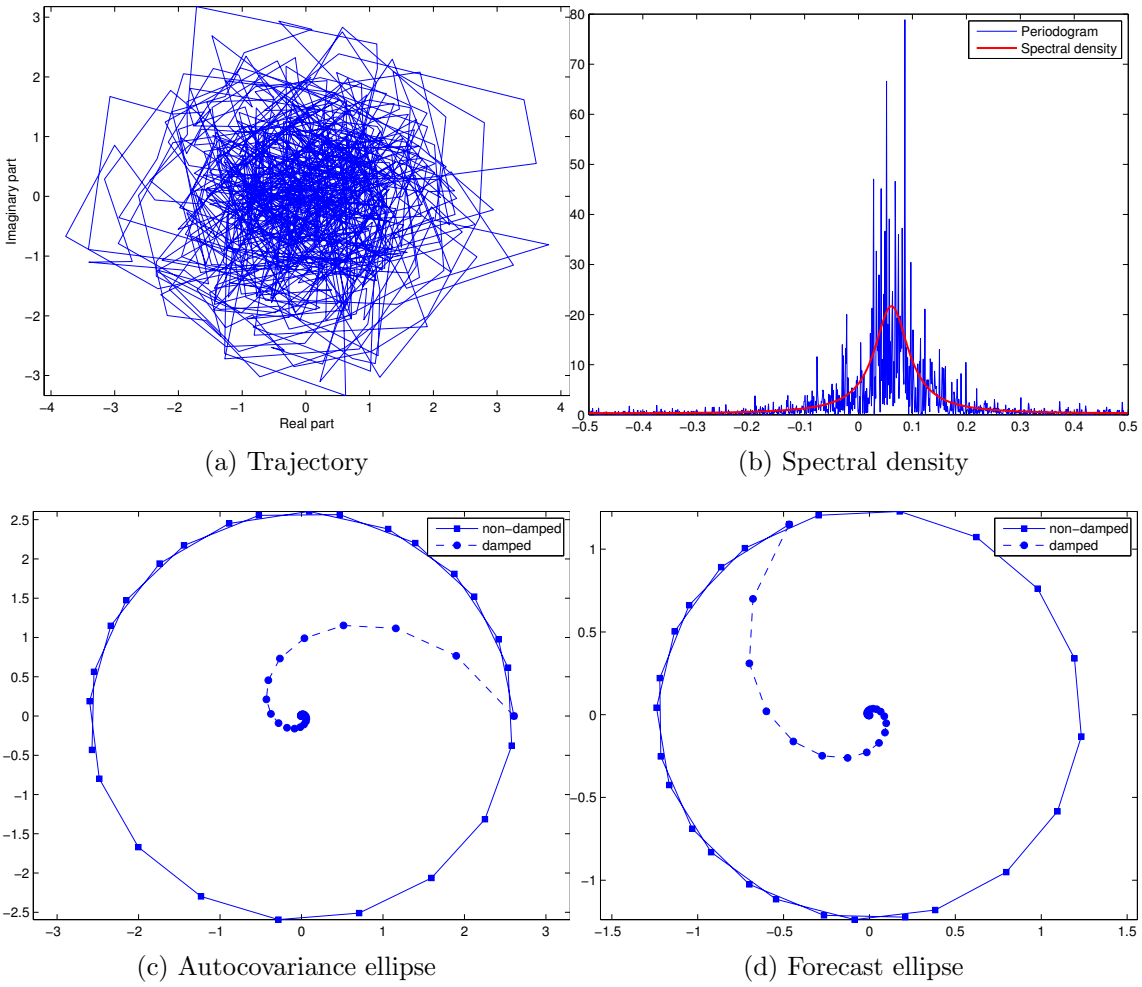
$$f = \frac{1}{2\pi} \arctan\left\{\frac{\rho_2}{\rho_1} \tan\lambda\right\}. \tag{6.2.8}$$

This shows that there will only be one peak in SDF for such a process. If $\rho_1 = \rho_2$ Equation (6.2.8) gives us the expected relationship between frequency and angular frequency, $f = \lambda/2\pi$.

The analysis of the SDF reinforces our previous findings. There is only one peak in the SDF, which indicated that the process structure is circular rather than elliptical. This was also seen previously in both the autocovariance and forecast structures, which were circular and not elliptical. The circle is just a degenerative form of an ellipse, or in other form a restricted version, where the axes are of equal lengths. So in order to achieve an elliptical structure in the time series parametric model we would need to observe two frequency peaks, on positive and negative frequencies. This can be achieved both by extending the complex-valued model into a widely linear model, or extending it into a bivariate model. In later parts of our research we will analyse the extension of this ℂAR(1) to the bivariate ℂVAR(1) model, because it is more interesting and simpler to parametrise.

6.2.4 Simulated example of univariate ℂAR(1)

In order to illustrate the above discussion we present a simulated example. We simulate one realization of length 1000 from the specified time series model $\{\psi_t\}$ in (6.2.1), by using the following values of parameters $f = 0.1$, $\lambda = 2\pi f = 0.6283$, $\rho_1 = 0.9$, $\rho_2 = 0.5$ and $\epsilon_t \sim N(0, 1)$. In Figure 6.1a we can see the trajectory of the simulated time series $\{\psi_t\}$ through time in the complex plane. As we can see, the trajectory of $\{\psi_t\}$ is a cloud of points without any structure, which is due to randomness and zero-mean. In Figure 6.1b we can see the periodogram (blue solid line) and the theoretical SDF (red solid line), which has a peak at frequency $f_{max} = 0.0611$ (angular frequency $\lambda_{max} = 2\pi f_{max} = 0.3836$), corresponding to the calculation in Equation (6.2.8).



(a) Trajectory (b) Spectral density

(c) Autocovariance ellipse (d) Forecast ellipse

Figure 6.1: Plots for the simulated $\mathbb{C}$AR(1) time series: (a) process trajectory over time; (b) periodogram and the theoretical spectral density (SDF); (c) non-damped and damped ACVF as a function of lag τ; and (d) non-damped and damped forecast function as a function of forecast-time l.

In Figure 6.1c we can see the plot of the theoretical ACVF in its original form, damped with the τ power of the parameter φ_λ, and a non-damped form. For the non-damped form we divide the ACVF with the τ power of the parameter φ_λ. We plot the non-damped one just because it is easer to see that it maps out a trajectory of a circle, which is rather difficult to see in the trajectory of the ACVF itself, because of the damping effect. Similar picture can be seen in Figure 6.1d where we plot the forecast function as it evolves through forecast-time l in its original form, damped with the l power of the parameter $|\varphi_\lambda|$, and a non-damped form, where we divide the forecast with the l power of the parameter $|\varphi_\lambda|$. Also in this case we can

nicely see that the non-damped forecast maps out a trajectory of a circle, whereas in the original version it is harder to see that we are dealing with a damped circle as opposed a damped ellipse. In other words, we process the functions to remove the damping only with the aim to see its trajectory more clearly.

6.3 Bivariate ℂVAR(1)

In the previous chapter we defined our initial stochastic oscillation model in the form of a univariate ℂAR(1). We have seen that it exhibits only one frequency peak at a desired frequency, either on the positive or the negative range. Moreover, also the autocovariance and forecast ellipses will always be circular. This means that the ℂAR(1) is a constrained version of a general elliptical model. We want to explore possibilities for a more general parametric model that would allow for a less constrained structure. In order to achieve a more general structure, we shall extend the simple complex univariate model to a bivariate complex-valued autoregressive model, which we will call a ℂVAR(1) model. In this chapter we will propose a general bivariate stochastic oscillation model that will produce an oscillation at the desired frequency and will at the same time be very useful to illustrate the concepts developed in the previous chapters.

The ℂVAR(1) is defined for a bivariate vector of time series elements, $\boldsymbol{\psi}_t = [\psi_t, \eta_t]^{\mathrm{T}}$. In order to produce a stochastic oscillation ℂVAR(1) model and induce the frequency of oscillation at a desired value we use the rotation matrix $\boldsymbol{R}_\lambda$ as described in Equation (2.1.3). As a basis to construct such a model we take the cyclical model as proposed by Harvey and Koopman [20] and introduced in Equation (6.1.1). However, we adapt their model to make the parameter complex-valued and express it in the following way

$$\boldsymbol{\psi}_t = \rho \boldsymbol{M}_{\delta,\lambda} \boldsymbol{\psi}_{t-1} + \boldsymbol{\epsilon}_t = \rho \underbrace{\boldsymbol{K}_\delta \boldsymbol{R}_\lambda \boldsymbol{K}_\delta^{\mathrm{H}}}_{\boldsymbol{M}_{\delta,\lambda}} \boldsymbol{\psi}_{t-1} + \boldsymbol{\epsilon}_t \tag{6.3.1}$$

$$\begin{aligned} \begin{bmatrix} \psi_t \\ \eta_t \end{bmatrix} &= \rho \begin{bmatrix} e^{i\delta} & 0 \\ 0 & e^{-i\delta} \end{bmatrix} \begin{bmatrix} \cos(\lambda) & \sin(\lambda) \\ -\sin(\lambda) & \cos(\lambda) \end{bmatrix} \begin{bmatrix} e^{-i\delta} & 0 \\ 0 & e^{i\delta} \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \eta_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t^{(1)} \\ \epsilon_t^{(2)} \end{bmatrix} \\ &= \rho \begin{bmatrix} \cos(\lambda) & \sin(\lambda) e^{i2\delta} \\ -\sin(\lambda) e^{-i2\delta} & \cos(\lambda) \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \eta_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t^{(1)} \\ \epsilon_t^{(2)} \end{bmatrix}. \end{aligned} \tag{6.3.2}$$

This model belongs to the class of ℂVAR(1) models [33, p. 13], but not every

$\mathbb{C}$VAR(1) model can be written in the form of (6.3.1), because of the parameter matrix that needs to be expressed as $\mathbf{M}_{\delta,\lambda} = \mathbf{K}_\delta \mathbf{R}_\lambda \mathbf{K}_\delta^{\mathrm{H}}$. The phase-shift matrix $\mathbf{K}_\delta$ is introduced in Equation (2.1.4) and the rotation matrix $\mathbf{R}_\lambda$ in Equation (2.1.3) both in Chapter 2.1. Also here will use the complex augmented vectors and express the model in its augmented form. The complex augmented vectors of bivariate vectors will now become 4×1 vectors

$$\underline{\boldsymbol{\psi}}_t = [\boldsymbol{\psi}_t^{\mathrm{T}}, \boldsymbol{\psi}_t^{\mathrm{H}}]^{\mathrm{T}} = [\psi_t, \eta_t, \psi_t^*, \eta_t^*]^{\mathrm{T}}, \tag{6.3.3}$$

and for the innovation processes

$$\underline{\boldsymbol{\epsilon}}_t = [\boldsymbol{\epsilon}_t^{\mathrm{T}}, \boldsymbol{\epsilon}_t^{\mathrm{H}}]^{\mathrm{T}} = [\epsilon_t^{(1)}, \epsilon_t^{(2)}, \epsilon_t^{(1)*}, \epsilon_t^{(2)*}]^{\mathrm{T}}. \tag{6.3.4}$$

Using these vectors we can express the time series model in (6.3.1) in its *augmented form*, which will be useful for estimation with the maximum likelihood method or direct definition of the autocovariance function, as

$$\underline{\boldsymbol{\psi}}_t = \rho \underline{\boldsymbol{M}}_{\delta,\lambda} \underline{\boldsymbol{\psi}}_{t-1} + \underline{\boldsymbol{\epsilon}}_t, \quad \text{where} \quad \underline{\boldsymbol{M}}_{\delta,\lambda} = \begin{bmatrix} \boldsymbol{M}_{\delta,\lambda} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{M}_{\delta,\lambda}^* \end{bmatrix}. \tag{6.3.5}$$

Parameters As we can see the coefficient of the model is composed of several elements. The matrix $\boldsymbol{M}_{\delta,\lambda}$ can be uniquely expressed as a product of phase shift and rotation matrices as $\boldsymbol{K}_\delta \boldsymbol{R}_\lambda \boldsymbol{K}_\delta^{\mathrm{H}}$. The following list provides explanation and intuition for the parameters used in our model:

- ρ is the *damping factor* in $\mathbb{R}$; in order to maintain stationarity of the model it is restricted to values $|\rho| < 1$.
- $\lambda \in (0, \pi)$ is the *frequency* in $\mathbb{R}$, which is associated with the clockwise rotation matrix $\boldsymbol{R}_\lambda$, and makes the model exhibit a SDF peak at the specified frequency, $f = \lambda/2\pi$.
- $\delta \in [-\pi/2, \pi/2]$ is the *phase shift factor* in $\mathbb{R}$, which is associated with the complex phase shift matrix $\boldsymbol{K}_\delta$, and makes the coefficient complex-valued.
- The *random noise* $\boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon)$ is bivariate and complex-valued, $\boldsymbol{\epsilon}_t \in \mathbb{C}^2$, sometimes called "doubly white noise" [52]. It can either be correlated or uncorrelated, and equally proper or improper. Direct observation verifies that

assuming a covariance between the elements of $\boldsymbol{\epsilon}_t$, comparing to $\text{cov}(\epsilon_t^{(1)}, \epsilon_{t'}^{(2)}) = \Sigma_\epsilon \delta_{t,t'}$, directly modifies the process $\{\boldsymbol{\psi}_t\}$ non-trivially, such that the height of the SDF frequency is altered. Incorporating such events allows us to reach a larger set of processes.

- Without loss of generality we are assuming a *zero-mean process*, such that $\mathbb{E}[\boldsymbol{\psi}_t] = 0$.

Interpretation We can interpret the above bivariate model as an affine transformation of the time series vector from the previous time point, $\boldsymbol{\psi}_{t-1}$. There are three deterministic transformations: rotation, complex-phase shift forward and backward and damping; and one stochastic transformation: translation (for more on affine transformations see Chapter 2.1.2 and [66, 64]). The complex-phase shift matrix shifts the phase of the two components of the time series angle by δ apart before rotation and back after rotation. This creates a "mixing effect" between the two components of the time series. Scaling could also be introduced as a possible extension, but that will be dealt with separately at a later point. The rotation of the axes, complex-phase shift and damping, in the absence of the noise $\boldsymbol{\epsilon}_t$, would produce a damped elliptical trajectory through time, where the damping depends on factor ρ. Due to the stochastic noise the values of the time series are at each step translated, which causes that the trajectory of the model not to be an ellipse any longer. This means that in the plot of the trajectory of the time series through time we can only see a scattered cloud of points without any pattern, see Figure 6.3 for example.

The bivariate model in (6.3.1) is in general defined for complex-valued time series vector $\boldsymbol{\psi}_t = [\psi_t, \eta_t]^\text{T} \in \mathbb{C}^2, \forall t$ and its coefficient matrix is generally complex-valued. Nevertheless, if the phase shift factor takes values as $\delta \in \{0, \pm\pi/2\}$, then the model coefficient will be real-valued and so the whole model will not be much different to the model described in Equation (6.1.1). In order for them to be equivalent, also the innovations vector would need to be real-valued $\boldsymbol{\epsilon}_t = [\epsilon_t^{(1)}, \epsilon_t^{(2)}]^\text{T} \in \mathbb{R}^2, \forall t$. We can say that the real-valued case is a constrained version of the general model in (6.3.1).

Normality assumption In this parametric model we assume that the innovation term is Gaussian, $\boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon)$. For many applications this might be a reasonable assumption and in many cases there is no need to doubt normality. In addition, normality is very useful since both the conditional and marginal distributions of the

general model in Equation (6.3.1) will be normal, see Section 6.3.2 below for more on distributions of this proposed model.

However, we might want to consider cases when normality cannot be assumed or the assumed distribution changed to a more flexible one. To relax Guassianity we can either specify a different innovation distribution $\epsilon_t \sim p(m, \boldsymbol{\theta})$, and marginal or conditional distributions of U_t. However, by doing this we will rarely have the same marginal, conditional or innovation distributions, or even, in some cases some of them may not be available analytically.

The simplest way is to specify a broad class of models, so called innovation class models, $U_t = \phi U_{t-1} + \epsilon_t$ where $\epsilon_t \sim p(m, \boldsymbol{\theta})$. Hence $U \sim p(m, \boldsymbol{\theta})$ will be a random variable with distribution p, which has mean m and other parameters characterizing the distributions in vector $\boldsymbol{\theta}$ [17]. Several authors have proposed $\mathsf{AR}(1)$ models with different innovation distributions, for example, using univariate distribution [4], logistic distribution [63], hyperbolic secant [50], or Laplace [11], etc.

The change of distribution of the innovation noise, in our model in Equation (6.3.1), would mainly impact the estimation of its parameters. If the conditional distribution is known, then the model can be estimated using maximum likelihood method. However, depending on the density function, the likelihood function might be discontinuous or not robust enough for estimation. On the other hand, it would not impact the analysis of the ACVF of forecast ellipses as presented below.

6.3.1 Innovation process of bivariate $\mathbb{C}\mathsf{VAR}(1)$

We find it beneficial to spend a little of time talking about the innovation process. Here we are dealing with a bivariate $\mathbb{C}\mathsf{VAR}(1)$ model so the definition of the "doubly white noise" is not that straight-forward. In general the innovation process, $\{\boldsymbol{\epsilon}_t\}$, will be complex-valued vector white noise with multivariate complex normal distribution that is independent across time. As with the analysis of second-order properties of any complex-valued process, also here we need to take into consideration all the relationships, including the one between real and imaginary parts.

Each of the elements of the random innovation vector at each time point is a complex-valued random number, such as $\epsilon_t^{(1)} = x_t + iy_t$ and $\epsilon_t^{(2)} = u_t + iv_t$. This indicates the need to analyse correlations between the real parts (x_t with u_t), imaginary parts (y_t with v_t), and the cross-correlation between terms combining real and imaginary parts of the vector (x_t with v_t and y_t with u_t). We analyse the com-

plex innovation process through its complex augmented vector, which will have four terms, $\underline{\boldsymbol{\epsilon}}_t = [\boldsymbol{\epsilon}_t^{\mathrm{T}}, \boldsymbol{\epsilon}_t^{\mathrm{H}}]^{\mathrm{T}} = [\epsilon_t^{(1)}, \epsilon_t^{(2)}, \epsilon_t^{(1)*}, \epsilon_t^{(2)*}]^{\mathrm{T}}$. The augmented covariance matrix can thus be written as

$$\underline{\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}} = \mathbb{E}[\underline{\boldsymbol{\epsilon}}_t \underline{\boldsymbol{\epsilon}}_t^{\mathrm{H}}] = \mathbb{E}\begin{bmatrix} \epsilon_t^{(1)}\epsilon_t^{(1)*} & \epsilon_t^{(1)}\epsilon_t^{(2)*} & \epsilon_t^{(1)}\epsilon_t^{(1)} & \epsilon_t^{(1)}\epsilon_t^{(2)} \\ \epsilon_t^{(2)}\epsilon_t^{(1)*} & \epsilon_t^{(2)*}\epsilon_t^{(2)*} & \epsilon_t^{(2)}\epsilon_t^{(1)} & \epsilon_t^{(2)}\epsilon_t^{(2)} \\ \epsilon_t^{(1)*}\epsilon_t^{(1)*} & \epsilon_t^{(1)*}\epsilon_t^{(2)*} & \epsilon_t^{(1)*}\epsilon_t^{(1)} & \epsilon_t^{(1)*}\epsilon_t^{(2)} \\ \epsilon_t^{(2)*}\epsilon_t^{(1)*} & \epsilon_t^{(2)*}\epsilon_t^{(2)*} & \epsilon_t^{(2)*}\epsilon_t^{(1)} & \epsilon_t^{(2)*}\epsilon_t^{(2)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} & \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} \\ \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}^* & \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}^* \end{bmatrix}. \tag{6.3.6}$$

We can see that the augmented covariance matrix above can, by symmetry, be split into two matrices and their complex conjugates. That simplifies the analysis to be able to deal with two 2×2 matrices only. The covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$ and the complimentary covariance matrix $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}$ can be expressed in terms of the complex innovation processes $\epsilon_t^{(1)} = x_t + iy_t$ and $\epsilon_t^{(2)} = u_t + iv_t$, as

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} = \mathbb{E}[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t^{\mathrm{H}}] &= \begin{bmatrix} \sigma_x^2 + \sigma_y^2 + i(\rho_{yx} - \rho_{xy}) & \rho_{xu} + \rho_{yv} + i(\rho_{yu} - \rho_{xv}) \\ \rho_{ux} + \rho_{vy} - i(\rho_{uy} - \rho_{vx}) & \sigma_u^2 + \sigma_v^2 + i(\rho_{vu} - \rho_{uv}) \end{bmatrix}, \\ \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} = \mathbb{E}[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t^{\mathrm{T}}] &= \begin{bmatrix} \sigma_x^2 - \sigma_y^2 + i(\rho_{xy} + \rho_{yx}) & \rho_{xu} - \rho_{yv} + i(\rho_{xv} + \rho_{yu}) \\ \rho_{ux} - \rho_{vy} + i(\rho_{vx} + \rho_{uy}) & \sigma_u^2 - \sigma_v^2 + i(\rho_{uv} + \rho_{vu}) \end{bmatrix}, \end{aligned} \tag{6.3.7}$$

where $\sigma_x^2 = \mathrm{var}(x_t)$, $\rho_{xy} = \mathrm{cov}(x_t, y_t)$, etc.

6.3.2 Likelihood estimation and identifiability of bivariate ℂVAR(1)

In this section we discuss the maximum likelihood technique for estimating the parameters of the model in Equation (6.3.1) as one of the most used estimation techniques in time series analysis. To estimate the parameters of a ℂAR(p) model, we use the concept of *conditional likelihood*, where we condition the likelihood on the past values of the time series [7, p. 226].

For the above general time series model the conditional distribution of its augmented vector $\underline{\boldsymbol{\psi}}_t$ can be expressed as

$$\underline{\boldsymbol{\psi}}_t \mid \underline{\boldsymbol{\psi}}_{t-1}, \underline{\boldsymbol{\psi}}_{t-2}, \ldots \sim N(\rho \underline{\boldsymbol{M}}_{\delta,\lambda} \underline{\boldsymbol{\psi}}_{t-1}, \underline{\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}}). \tag{6.3.8}$$

Due to the Markov property of first-order autoregressive models, the value at time t will depend only on the value at time $t-1$, and so we can express the above

conditional distribution with values at $t-1$ only

$$\underline{\boldsymbol{\psi}}_t \mid \underline{\boldsymbol{\psi}}_{t-1} \sim N(\rho \underline{\boldsymbol{M}}_{\delta,\lambda} \underline{\boldsymbol{\psi}}_{t-1}, \underline{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}), \tag{6.3.9}$$

We assume Gaussianity and use the general PDF of a complex normal RV $\mathbf{U} \in \mathbb{C}^n$, which can be written by using its complex augmented vector $\mathbf{W} = [\mathbf{U}^{\mathrm{T}}, \mathbf{U}^{\mathrm{H}}]^{\mathrm{T}}$ [2], as

$$p(\mathbf{U}) = \frac{1}{\pi^n |\boldsymbol{\Gamma}_W|^{1/2}} \exp\Big\{ -\tfrac{1}{2}(\mathbf{W} - \boldsymbol{\mu}_W)^{\mathrm{H}} \boldsymbol{\Gamma}_W^{-1} (\mathbf{W} - \boldsymbol{\mu}_W) \Big\}, \tag{6.3.10}$$

where $|\boldsymbol{\Gamma}_W|$ denotes the matrix determinant of the augmented covariance matrix, $\boldsymbol{\Gamma}_W = \mathbb{E}[(\mathbf{W} - \boldsymbol{\mu}_W)(\mathbf{W} - \boldsymbol{\mu}_W)^{\mathrm{H}}]$. Together with the above conditional distribution of the $\mathbb{C}$VAR(1) model in (6.3.9) we can write the conditional PDF of our time series model, as

$$p(\underline{\boldsymbol{\psi}}_t \mid \underline{\boldsymbol{\psi}}_{t-1}; \boldsymbol{\theta}) = \frac{1}{\pi^n |\underline{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}|^{1/2}} \exp\Big\{ -\tfrac{1}{2}(\underline{\boldsymbol{\psi}}_t - \rho \underline{\boldsymbol{M}}_{\delta,\lambda} \underline{\boldsymbol{\psi}}_{t-1})^{\mathrm{H}} \underline{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}{}^{-1} (\underline{\boldsymbol{\psi}}_t - \rho \underline{\boldsymbol{M}}_{\delta,\lambda} \underline{\boldsymbol{\psi}}_{t-1}) \Big\}. \tag{6.3.11}$$

In the above conditional PDF, $\boldsymbol{\theta}$ is a vector of parameters of interest, which in pur case of the $\mathbb{C}$VAR(1) model is $\boldsymbol{\theta} = [\rho, \delta, \lambda]^{\mathrm{T}}$. From the above the conditional likelihood $L(\boldsymbol{\theta})$, we derive the log-likelihood of our time series model as

$$\begin{aligned} l(\boldsymbol{\theta}) &= \sum_{t=2}^{T} \log p(\underline{\boldsymbol{\psi}}_t \mid \underline{\boldsymbol{\psi}}_{t-1}; \boldsymbol{\theta}) \\ &= -\frac{T-1}{2} \log |\underline{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}| - \frac{1}{2} \sum_{t=2}^{T} \Big\{ (\underline{\boldsymbol{\psi}}_t - \rho \underline{\boldsymbol{M}}_{\delta,\lambda} \underline{\boldsymbol{\psi}}_{t-1})^{\mathrm{H}} \underline{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}{}^{-1} (\underline{\boldsymbol{\psi}}_t - \rho \underline{\boldsymbol{M}}_{\delta,\lambda} \underline{\boldsymbol{\psi}}_{t-1}) \Big\} + C. \end{aligned} \tag{6.3.12}$$

To estimate the parameters in $\boldsymbol{\theta}$ we maximise the log-likelihood as a function of the parameters $\boldsymbol{\theta}$.

6.3.3 Autocovariance function of bivariate $\mathbb{C}$VAR(1)

The autocovariance function (ACVF) of the bivariate $\mathbb{C}$VAR(1) time series model in Equation (6.3.1) is derived by using the complex augmented vector $\underline{\boldsymbol{\psi}}_t = [\boldsymbol{\psi}_t^{\mathrm{T}}, \boldsymbol{\psi}_t^{\mathrm{H}}]^{\mathrm{T}} = [\psi_t, \eta_t, \psi_t^*, \eta_t^*]^{\mathrm{T}}$. Since the process is zero-mean, we can write the augmented auto-

covariance (ACV) matrix as

$$\begin{aligned}\underline{\boldsymbol{\Gamma}_{\boldsymbol{\psi}}}(\tau) &= \mathbb{E}[\underline{\boldsymbol{\psi}}_{t+\tau}\underline{\boldsymbol{\psi}}_{t}^{\mathrm{H}}] \\ &= \begin{bmatrix} \gamma_{\psi\psi}(\tau) & \gamma_{\psi\eta}(\tau) & \tilde{\gamma}_{\psi\psi}(\tau) & \tilde{\gamma}_{\psi\eta}(\tau) \\ \gamma_{\eta\psi}(\tau) & \gamma_{\eta\eta}(\tau) & \tilde{\gamma}_{\eta\psi}(\tau) & \tilde{\gamma}_{\eta\eta}(\tau) \\ \tilde{\gamma}^{*}_{\psi\psi}(\tau) & \tilde{\gamma}^{*}_{\psi\eta}(\tau) & \gamma^{*}_{\psi\psi}(\tau) & \gamma^{*}_{\psi\eta}(\tau) \\ \tilde{\gamma}^{*}_{\eta\psi}(\tau) & \tilde{\gamma}^{*}_{\eta\eta}(\tau) & \gamma^{*}_{\eta\psi}(\tau) & \gamma^{*}_{\eta\eta}(\tau) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Gamma}_{\boldsymbol{\psi}}(\tau) & \tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}(\tau) \\ \tilde{\boldsymbol{\Gamma}}^{*}_{\boldsymbol{\psi}}(\tau) & \boldsymbol{\Gamma}^{*}_{\boldsymbol{\psi}}(\tau) \end{bmatrix}, \end{aligned} \tag{6.3.13}$$

where are the bottom two blocks ($\tilde{\boldsymbol{\Gamma}}^{*}_{\boldsymbol{\psi}}(\tau)$ and $\boldsymbol{\Gamma}^{*}_{\boldsymbol{\psi}}(\tau)$) just conjugates of the two blocks on the top ($\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(\tau)$ and $\tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}(\tau)$). This simplifies the analysis and we can focus only on the Hermitian ACV matrix $\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(\tau) = \mathbb{E}[\boldsymbol{\psi}_{t+\tau}\boldsymbol{\psi}_{t}^{\mathrm{H}}]$, and the complimentary autocovariance (C-ACV) matrix, $\tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}(\tau) = \mathbb{E}[\boldsymbol{\psi}_{t+\tau}\boldsymbol{\psi}_{t}^{\mathrm{T}}]$.

The easiest way to obtain the ACVF is to use the Yule-Walker equations [51]. In order to obtain the expression for the Hermitian ACV matrix we iteratively solve the Yule-Walker equation for different values of τ

$$\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(\tau) = \mathbb{E}[\boldsymbol{\psi}_{t+\tau}\boldsymbol{\psi}_{t}^{\mathrm{H}}] = \mathbb{E}[(\rho\boldsymbol{M}_{\delta,\lambda}\boldsymbol{\psi}_{t+\tau-1} + \boldsymbol{\epsilon}_{t})\boldsymbol{\psi}_{t}^{\mathrm{H}}]. \tag{6.3.14}$$

The ACV matrix, for different values of τ, then becomes

$$\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(\tau) = \begin{cases} \rho^{2}\boldsymbol{M}_{\delta,\lambda}\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0)\boldsymbol{M}_{\delta,\lambda}^{\mathrm{H}} + \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} & \text{for } \tau = 0, \\ \rho^{\tau}\boldsymbol{K}_{\delta}\boldsymbol{R}_{\lambda\tau}\boldsymbol{K}_{\delta}^{\mathrm{H}}\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0) & \text{for } \tau > 0, \\ \rho^{|\tau|}\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0)\boldsymbol{K}_{\delta}\boldsymbol{R}_{\lambda|\tau|}^{\mathrm{T}}\boldsymbol{K}_{\delta}^{\mathrm{H}} & \text{for } \tau < 0. \end{cases} \tag{6.3.15}$$

These equations cannot determine the covariance matrix at $\tau = 0$, as the term $\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0)$ appears on both sides of the equal sign, $\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0) = \rho^{2}\boldsymbol{M}_{\delta,\lambda}\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0)\boldsymbol{M}_{\delta,\lambda}^{\mathrm{H}} + \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$. According to Lütkepohl [33, p. 27] we can use the *vec* operator[1] to rearrange the terms of (6.3.15) when $\tau = 0$. By doing this we obtain an expression where the

[1]The vec operator, denoted as $\mathrm{vec}\{\cdot\}$ transforms an $m \times n$ matrix by stacking its columns into an $mn \times 1$ vector. The property of the vec operator used in this derivation is $\mathrm{vec}\{ABC\} = (C^{T} \otimes A)\mathrm{vec}\{B\}$, where A, B and C are matrices with appropriate dimensions and $\otimes$ denotes Kronecker product.

elements of the variance matrix of the model are all stacked up in a 4×1 vector

$$\begin{aligned} \text{vec}\{\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0)\} &= \rho^2 \text{vec}\{\boldsymbol{M}_{\delta,\lambda}\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0)\boldsymbol{M}_{\delta,\lambda}^{\text{H}}\} + \text{vec}\{\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}\} \\ &= \rho^2[(\boldsymbol{M}_{\delta,\lambda}^{\text{H}})^{\text{T}} \otimes \boldsymbol{M}_{\delta,\lambda}]\text{vec}\{\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0)\} + \text{vec}\{\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}\} \\ &= [\mathbf{I}_4 - \rho^2((\boldsymbol{M}_{\delta,\lambda}^{\text{H}})^{\text{T}} \otimes \boldsymbol{M}_{\delta,\lambda})]^{-1}\text{vec}\{\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}\}. \end{aligned} \tag{6.3.16}$$

Similarly as above, in order to obtain the expressions for the C-ACV matrix we use the Yule-Walker equation $\tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}(\tau) = \mathbb{E}[\boldsymbol{\psi}_{t+\tau}\boldsymbol{\psi}_t^{\text{T}}] = \mathbb{E}[(\rho\boldsymbol{M}_{\delta,\lambda}\boldsymbol{\psi}_{t+\tau-1} + \boldsymbol{\epsilon}_t)\boldsymbol{\psi}_t^{\text{T}}]$ and solve it iteratively for different values of τ. The expression for the C-ACV matrix for the different values of τ is

$$\tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}(\tau) = \begin{cases} \rho^2\boldsymbol{M}_{\delta,\lambda}\tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}(0)\boldsymbol{M}_{\delta,\lambda}^{\text{T}} + \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} & \text{for } \tau = 0, \\ \rho^{\tau}\boldsymbol{K}_{\delta}\boldsymbol{R}_{\lambda\tau}\boldsymbol{K}_{\delta}^{\text{H}}\tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}(0) & \text{for } \tau > 0, \\ \rho^{|\tau|}\tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}^{\text{H}}(0)\boldsymbol{K}_{\delta}\boldsymbol{R}_{\lambda|\tau|}^{\text{T}}\boldsymbol{K}_{\delta}^{\text{H}} & \text{for } \tau < 0. \end{cases} \tag{6.3.17}$$

Due to the same issue, the above cannot be used to solve for the complimentary covariance matrix. We use the same property of the vec operator to yield an expression for the complimentary covariance matrix $\tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}(0)$ with its elements all stacked up in a 4×1 vector

$$\text{vec}\{\tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}(0)\} = [\mathbf{I}_4 - \rho^2(\boldsymbol{M}_{\delta,\lambda} \otimes \boldsymbol{M}_{\delta,\lambda})]^{-1}\text{vec}\{\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}\}. \tag{6.3.18}$$

ACVF as a function of lag τ We are interested in analysing the behaviour of the ACVF as a function of lag-time τ in order to see if this parametric model exhibits an autocovariance ellipse (as per Definition 5.2.3). We look at the trajectory of each of the elements of the ACVF as we progress through lags τ and observe their shape. In other words, we plot the trajectory of the ACVF in the complex plane. For simplicity we will analyse only the hermitian ACV matrix $\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(\tau)$, but the analysis would be the same for the C-ACV matrix, mutatis mutandis. Same applies to the positive and negative values of τ. We will look only at the case when $\tau \geq 0$, because for $\tau < 0$ the ACVF is just its conjugate transpose, $\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(\tau) = \boldsymbol{\Gamma}_{\boldsymbol{\psi}}^{\text{H}}(-\tau)$.

Remark 6.3.1. *We analyse the autocovariance function of this time series model in* (6.3.1) *as function of lag τ. From such analysis we can see that each entry of the ACV matrix will represent a damped ellipse, based on the univariate parametric*

representation in (5.1.3),

$$
\begin{aligned}
&\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(\tau) = \rho^{\tau}\boldsymbol{M}_{\delta,\lambda}^{\tau}\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0) = \rho^{\tau}\boldsymbol{K}_{\delta}\boldsymbol{R}_{\lambda\tau}\boldsymbol{K}_{\delta}^{\mathrm{H}}\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0)\\
&= \rho^{\tau}\begin{bmatrix} \cos(\lambda\tau) & \sin(\lambda\tau)e^{i2\delta} \\ -\sin(\lambda\tau)e^{-i2\delta} & \cos(\lambda\tau) \end{bmatrix}\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0)\\
&= \rho^{\tau}\begin{bmatrix} \cos(\lambda\tau) & \sin(\lambda\tau)e^{i2\delta} \\ -\sin(\lambda\tau)e^{-i2\delta} & \cos(\lambda\tau) \end{bmatrix}\begin{bmatrix} \gamma_1 & \gamma_2 \\ \gamma_2^* & \gamma_4 \end{bmatrix}\\
&= \frac{\rho^{\tau}}{2}\begin{bmatrix} \alpha_1 e^{i\lambda\tau} + \alpha_2 e^{-i\lambda\tau} & -ie^{i2\delta}(\alpha_3 e^{i\lambda\tau} - \alpha_4 e^{-i\lambda\tau}) \\ ie^{-i2\delta}(\alpha_1 e^{i\lambda\tau} - \alpha_2 e^{-i\lambda\tau}) & \alpha_3 e^{i\lambda\tau} + \alpha_4 e^{-i\lambda\tau} \end{bmatrix};
\end{aligned}
\tag{6.3.19}
$$

where

$$
\begin{aligned}
\alpha_1 &= \gamma_1 - i\gamma_2^* e^{i2\delta}, \qquad & \alpha_3 &= \gamma_4 + i\gamma_2 e^{-i2\delta},\\
\alpha_2 &= \gamma_1 + i\gamma_2^* e^{i2\delta}, \qquad & \alpha_4 &= \gamma_4 - i\gamma_2 e^{-i2\delta}.
\end{aligned}
$$

From Remark 6.3.1 we can clearly see that each entry of the autocovariance matrix follows a parametric model of an elliptical trajectory as discussed earlier in Equation (5.1.3). This proves that the elements of the autocovariance matrix will exhibit elliptical trajectories and in some instances one of the degenerate cases (circle or line). For example, if we compare the first entry of the matrix in (6.3.19) to the parametric ellipse representation from (5.1.3), we can see that α_1 acts as $A_+e^{i\theta_+}$ and α_2 acts as $A_-e^{-i\theta_-}$, which define the amplitude and phase of the rotating phasors. As a result the autocovariance function of this model produces four autocovariance ellipses through lag-time τ. These ellipses are damped with exponential decay due to the stationarity condition, $\rho < 1$. The theoretical ACVF can be divided by ρ^{τ} to achieve a constant (non-damped) mapped ellipse for every entry of the matrix. The first and fourth one have orientation angle of zero, because the variance at lag zero is real-valued.

Moreover, we see the root-mean-square amplitude is the same between ellipses in the columns. They are pairwise proportionately same in size, but with different orientations. This indicates that we therefore need to study only two of these four entries.

$$
\kappa_{\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(\tau)} = \rho^{\tau}\begin{bmatrix} \sqrt{(\gamma_1^2 + |\gamma_2|^2)/2} & \sqrt{(\gamma_4^2 + |\gamma_2|^2)/2} \\ \sqrt{(\gamma_1^2 + |\gamma_2|^2)/2} & \sqrt{(\gamma_4^2 + |\gamma_2|^2)/2} \end{bmatrix} \tag{6.3.20}
$$

6.3.4 Forecast of bivariate $\mathbb{C}$VAR(1)

In Remark 6.3.1 we saw that the bivariate $\mathbb{C}$VAR(1) model in (6.3.1) exhibits an autocovariance ellipse. We now want to perform the same analysis on the forecast of the model and analyse its behaviour. The forecast function of time series was introduced in Chapter 3.2 and expressed for AR(1) models in Equation (3.2.2). The forecast ellipse per Definition 5.2.2 is interpreted as: if we know the value of the process today and in the past, we will be able to make our best estimate for its future based on the elliptical trajectory of the forecast function. The main aim is to understand what kind of trajectory of the signal we can expect in the future conditionally to knowing the present and past values.

Remark 6.3.2. *The forecast ellipse of the model in* (6.3.1) *can be viewed as a set of points (real and imaginary parts) of each entry of the forecast vector that will map an elliptical trajectory in the complex plane as we progress through forecast-time l. We represent this by deriving the forecast function of the model and expressing it in a parametric form of an ellipse as*

$$\begin{aligned}
\widehat{\boldsymbol{\psi}}_t(l) &= \begin{bmatrix}\widehat{\psi}_t(l)\\ \widehat{\eta}_t(l)\end{bmatrix} = \mathbb{E}[\boldsymbol{\psi}_{t+l} \mid \boldsymbol{\psi}_t, \boldsymbol{\psi}_{t-1}, \ldots] \\
&= \rho \boldsymbol{M}_{\delta,\lambda} \widehat{\boldsymbol{\psi}}_t(l-1) = \ldots = \rho^l \boldsymbol{M}^l_{\delta,\lambda} \boldsymbol{\psi}_t \\
&= \rho^l \boldsymbol{K}_\delta \boldsymbol{R}_{\lambda l} \boldsymbol{K}^{\mathrm{H}}_\delta \boldsymbol{\psi}_t \\
&= \rho^l \begin{bmatrix} \cos(\lambda l) & \sin(\lambda l) e^{i2\delta} \\ -\sin(\lambda l) e^{-i2\delta} & \cos(\lambda l) \end{bmatrix} \begin{bmatrix}\psi_t \\ \eta_t\end{bmatrix} \\
&= \rho^l \begin{bmatrix} \psi_t \cos(l\lambda) + \eta_t e^{i2\delta} \sin(l\lambda) \\ -\psi_t e^{-i2\delta} \sin(l\lambda) + \eta_t \cos(l\lambda) \end{bmatrix} \\
&= \rho^l \begin{bmatrix} \frac{1}{2}(\psi_t - i\eta_t e^{i2\delta}) e^{i\lambda l} + \frac{1}{2}(\psi_t + i\eta_t e^{i2\delta}) e^{-i\lambda l} \\ \frac{1}{2}(\eta_t + i\psi_t e^{-i2\delta}) e^{i\lambda l} + \frac{1}{2}(\eta_t - i\psi_t e^{-i2\delta}) e^{-i\lambda l} \end{bmatrix}.
\end{aligned} \tag{6.3.21}$$

From Remark 6.3.2 we see that each of the elements of the forecast function of the model will map a trajectory of a damped ellipse as we move through forecast-time l forward with an exponential decay. If we are dealing with multivariate time series, each of the entries of the forecast vector in (6.3.21) is in a parametric form that will map an ellipse in the complex plane, similar to the signal in (5.1.3). On the contrary, not all forecast functions will result in a function that would map an ellipse as we move through the forecast-time l. This also means that not every parametric

model will exhibit a forecast ellipse, but it depends on the parametrisation itself. Moreover, we can show that the two 2-dimensional ellipses analysed through the forecast vector are same in magnitude. The root-mean-square amplitude κ is same for both ellipses

$$\kappa = \sqrt{(|\psi_t|^2 + |\eta_t|^2)/2}. \tag{6.3.22}$$

6.3.5 Spectral analysis of bivariate ℂVAR(1)

Here we analyse the time series model in (6.3.1) in terms of the spectral analysis as discussed in Chapter 2.2.3. In order to express the spectral density function of the model we need to define the spectral processes $\{Z_\psi(f)\}$ and $\{Z_\eta(f)\}$ for each of the parts of our process $\boldsymbol{\psi}_t = [\psi_t, \eta_t]^{\mathrm{T}}$, with the respective orthogonal increments $\{dZ_\psi(f\}$ and $\{dZ_\eta(f)\}$. For convenience and comprehensive analysis we stack the orthogonal increment processes into a complex augmented vector as

$$\underline{\mathbf{dZ}_{\boldsymbol{\psi}}}(f) = [d\mathbf{Z}_{\boldsymbol{\psi}}^{\mathrm{T}}(f), d\mathbf{Z}_{\boldsymbol{\psi}}^{\mathrm{H}}(-f)]^{\mathrm{T}} = [dZ_\psi(f), dZ_\eta(f), dZ_\psi^*(-f), dZ_\eta^*(-f)]^{\mathrm{T}}. \tag{6.3.23}$$

Using the augmented vector of the orthogonal increment process we can derive the augmented spectral density (SDF) matrix using the Equation (2.2.10), as

$$\begin{aligned}\underline{\mathbf{S}_{\boldsymbol{\psi}}}(f) &= \mathbb{E}[\underline{\mathbf{dZ}_{\boldsymbol{\psi}}}(f)\underline{\mathbf{dZ}_{\boldsymbol{\psi}}}^{\mathrm{H}}(f)] \\ &= \begin{bmatrix} S_{\psi\psi}(f) & S_{\psi\eta}(f) & \tilde{S}_{\psi\psi}(f) & \tilde{S}_{\psi\eta}(f) \\ S_{\eta\psi}(f) & S_{\eta\eta}(f) & \tilde{S}_{\eta\psi}(f) & \tilde{S}_{\eta\eta}(f) \\ \tilde{S}^*_{\psi\psi}(f) & \tilde{S}^*_{\psi\eta}(f) & S^*_{\psi\psi}(-f) & S^*_{\psi\eta}(-f) \\ \tilde{S}^*_{\eta\psi}(f) & \tilde{S}^*_{\eta\eta}(f) & S^*_{\eta\psi}(-f) & S^*_{\eta\eta}(-f) \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{\boldsymbol{\psi}}(f) & \tilde{\mathbf{S}}_{\boldsymbol{\psi}}(f) \\ \tilde{\mathbf{S}}^*_{\boldsymbol{\psi}}(f) & \mathbf{S}_{\boldsymbol{\psi}}(-f) \end{bmatrix}.\end{aligned} \tag{6.3.24}$$

The above equation is composed of four block matrices, but only two are essential for our analysis. Those two are the spectral matrix (SDF), $\boldsymbol{S}_{\boldsymbol{\psi}}(f) = \mathbb{E}[d\mathbf{Z}_{\boldsymbol{\psi}}(f)d\mathbf{Z}_{\boldsymbol{\psi}}^{\mathrm{H}}(f)]$ and the complimentary spectral matrix (C-SDF), $\tilde{\boldsymbol{S}}_{\boldsymbol{\psi}}(f) = \mathbb{E}[d\mathbf{Z}_{\boldsymbol{\psi}}(f)d\mathbf{Z}_{\boldsymbol{\psi}}^{\mathrm{T}}(-f)]$.

To derive the augmented spectral matrix of the parametric ℂVAR(1) model we can use the Equation (3.1.5) presented in Chapter 3.1.3. We need to take into consideration the complex-valued property of our model and so need to use the augmented version of our model $\{\underline{\boldsymbol{\psi}}_t\}$ as per Equation (6.3.5). So the augmented spectral matrix can be expressed in the following way

$$\underline{\mathbf{S}_{\boldsymbol{\psi}}}(f) = \underline{\mathbf{H}}(e^{-i2\pi f})\underline{\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}}\underline{\mathbf{H}}^{\mathrm{H}}(e^{-i2\pi f}), \tag{6.3.25}$$

where $\underline{\mathbf{H}}(e^{-i2\pi f})$ is the augmented transfer function matrix calculated as the inverse of the characteristic polynomial $\underline{\mathbf{H}}(B) = \underline{\Phi}(B)^{-1} = [I_2 - \rho\underline{\mathbf{M}}_{\delta,\lambda}B]^{-1}$, by substituting the lag operator B with $e^{-i2\pi f}$. Unfortunately the above matrix multiplication and the inversion of the characteristic polynomial is not a simple operation and so we are not able to express the spectral matrix in closed form.

From the equation above we can see that the spectral matrix will depend on the parameters of the frequency λ, damping factor ρ and scaling factor δ. Nevertheless, using data simulation from our model, we observe the SDF and notice that we are able to control the peaks of the oscillation with a single parameter, λ. The parameter λ places the peaks of the spectral density at the angular frequencies of $-\lambda$ and λ on the interval $[-\pi, \pi]$, or in terms of frequencies at $f = \lambda/(2\pi)$ on the interval $[-1/2, 1/2]$. The angular frequencies $\omega = \lambda$ and the frequencies f are the same concept and have a straight forward relationship $\omega = 2\pi f$. The damping factor ρ has a dispersion effect on the shape of the spectral density. The phase-shift parameter δ also influences the shape of the curve and heights of the peaks, but not the location of the peaks. The height of the SDF will also be influenced by the noise term, if the innovation process is uncorrelated the peaks will be of same heights, whereas if it is correlated across terms the peaks will have different heights.

6.3.6 Simulated example of bivariate $\mathbb{C}$VAR(1)

We continue with presenting a simulated example from our time series model in (6.3.1). On one hand we use this example to present how does the stochastic oscillation model work and on the other to illustrate the concepts of the autocovariance and forecast ellipse developed earlier. We have simulated time series realisation of length 1000 with parameter values: $\rho = 0.8$, $\delta = \pi/5$ and $\lambda = 0.6283$ (corresponding to $f = 0.1$). For the generation of values we used the recursive model Equation (6.3.1) initiated at a random value generated from the same distribution as the random noise. The innovation process used are bivariate complex-valued and improper independent normally distributed RVs, such that $\underline{\boldsymbol{\epsilon}}_t \sim N(0, \underline{\boldsymbol{\Sigma}}_\epsilon)$, where the ACV and C-ACV matrices are chosen to be

$$\boldsymbol{\Sigma}_\epsilon = \begin{bmatrix} 1.4 & 0.6 \\ 0.6 & 1.1 \end{bmatrix}, \quad \text{and} \quad \tilde{\boldsymbol{\Sigma}}_\epsilon = \begin{bmatrix} 0.2 + 0.8i & 0.2i \\ 0.2i & 0.1 + 0.6i \end{bmatrix}. \tag{6.3.26}$$

Figure 6.2 shows the time trajectory of the bivariate time series $\{\boldsymbol{\psi}_t\}$ separately

for the real and imaginary parts, whereas Figure 6.3 shows the trajectory of both entries of the time series vector in the complex plane. As we can see these trajectories are just a cloud of points without any structure, mainly due to randomness and zero-mean of the process. In Figure 6.4 we show the periodograms of the two components of the simulated time series and their theoretical SDF (red solid line). Figure 6.5 shows the plots of the autocovariance ellipses of each of the four entries of the ACVF of our simulated model, and Figure 6.6 shows the plots of the forecast ellipses of each of the two entries of the forecast vector of the simulated model.

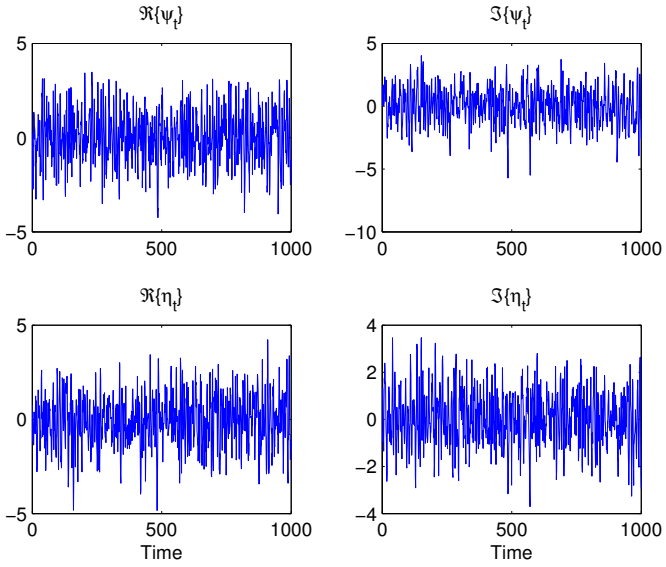


Figure 6.2: Plots of the simulated ℂVAR(1) time series trajectory in time for the real (left) and the imaginary (right) parts of the two components, ψ_t (above) and η_t (below)

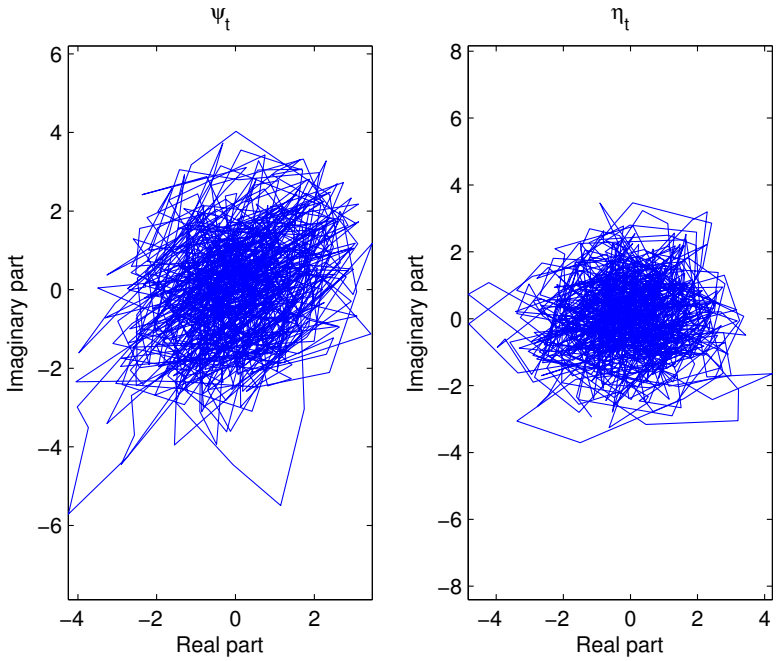


Figure 6.3: Plots of the simulated $\mathbb{C}$VAR(1) time series trajectory in the complex plane of the two components, ψ_t (left) and η_t (right)

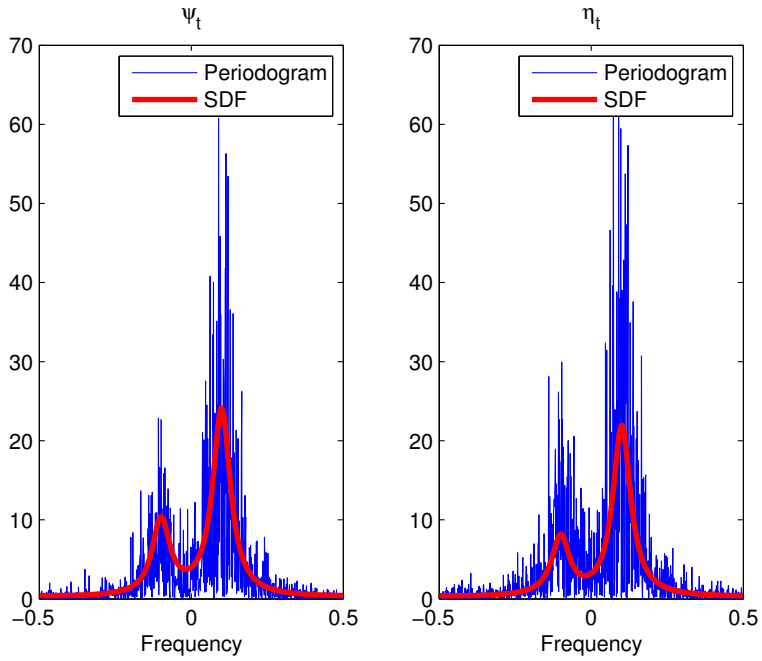


Figure 6.4: Periodograms and the SDF of the simulated $\mathbb{C}$VAR(1) time series ψ_t (left) and η_t (right)

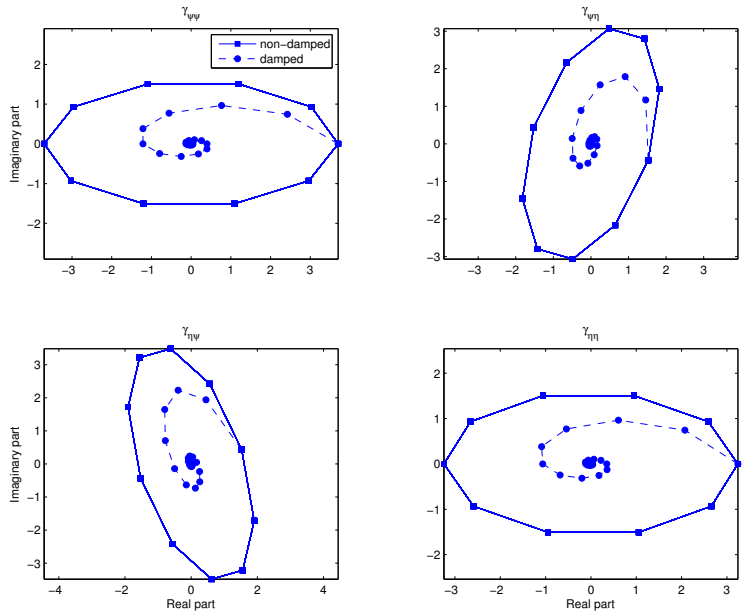


Figure 6.5: Plots of the autocovariance ellipse of the simulated ℂVAR(1) time series, showing the damped and non-damped ellipses from each of the entries of the ACV matrix through lags τ

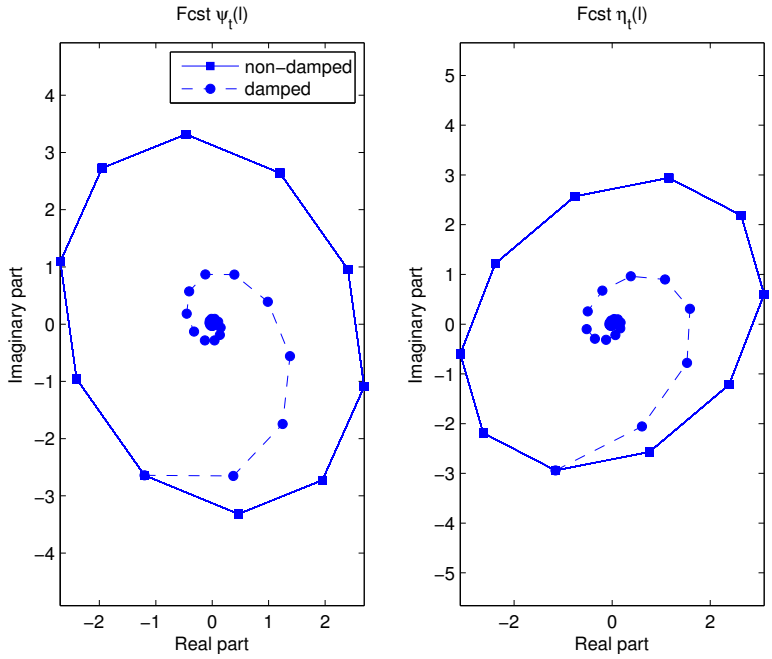


Figure 6.6: Plots of the forecast ellipses of the simulated $\mathbb{C}$VAR(1) time series, showing the damped and non-damped ellipses from each of the the entries of the forecast vector through forecast-time l

Identifiability of the general model To assess the usability of the model we want to investigate if the parameters are uniquely identifiable. We use the simulated data to fit the model parameters using the conditional likelihood from Equation (6.3.12), in order to observe the behaviour of the likelihood function and analyse the identifiability of the parameters. Below in Figure 6.7 we show the surface plot of the log-likelihood in 3-dimensions for estimated parameters ρ and δ.

The main observation is that the surface plot has a clear maximum and is locally convex at the true values of the parameters. Since we used simulated data we know the true values of the parameters and see that they were correctly estimated. We can conclude that the likelihood is well behaved and that the parameters are uniquely identifiable in the parameter space. Figure 6.8 shows the log-likelihood function for a single parameter at the true value of the other one. As we can see the likelihood function has a unique maximum for both parameters.

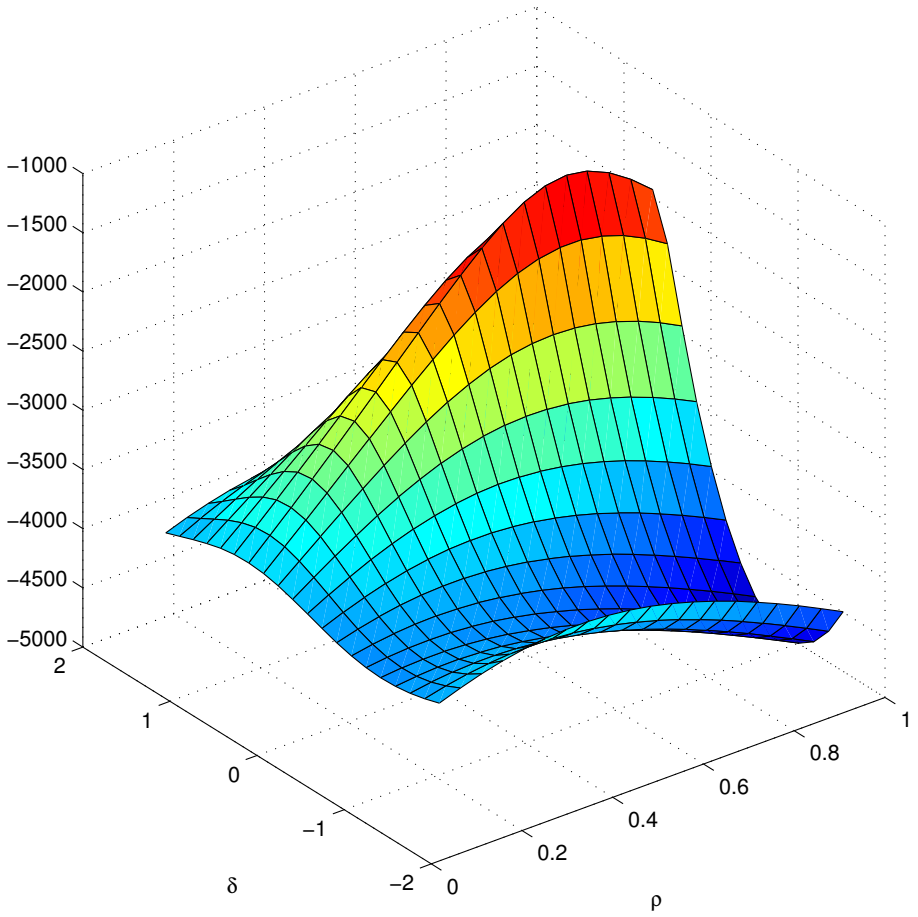


Figure 6.7: Surface plot of the likelihood function of the simulated ℂVAR(1) time series for parameters ρ and δ

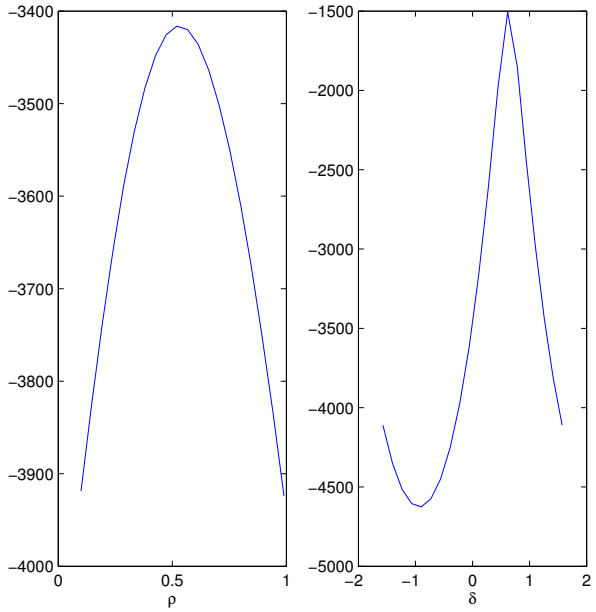


Figure 6.8: Section of above surface plot of the likelihood function for true parameters $\delta = \pi/5$ (left) and $\rho = 0.8$ (right)

Model estimation example We show also the estimation of the parameters using maximum likelihood method using the same simulated data. We the conditional log-likelihood defined in Equation (6.3.12) to estimate the parameters from data generated in this example. Since we use the simulated data we know the true values of the parameters are $\rho = 0.8$, $\delta = \pi/5$ and $\lambda = 0.6283$. We use the log-likelihood function to iteratively estimate the parameters, with starting values that are in the space of possible values for each parameter. The estimation method gave us the following estimates: $\hat{\rho} = 0.7946$, $\hat{\delta} = 0.5333$ and $\hat{\lambda} = 0.6364$, which are very close to the true values.

6.4 Extensions of the parametric time series models

In continuation, we explore further extensions and variations of the parametric time series model developed in the previous chapter. As the basis we will use the bivariate VAR(1) model in (6.3.1) to which we will add a *stretching factor*. We saw that this model features the desired qualities, so we are interested in how can it be further extended and what do such changes mean for the model and its behaviour. We will

look at both real- and complex-valued models. In this research we just mention these models as a possible extension, but we will not develop them further, mainly because they are mathematically a bit more complex and not easy to work with.

6.4.1 $\mathbb{C}$VAR(1) with stretching factor

First, we take the time series model from the previous chapter in (6.3.1) and add a stretching factor α to it. This is done by multiplying the model coefficient by the stretching matrix $\boldsymbol{V}_\alpha$. The stretching introduces an additional asymmetry between the components of the bivariate time series vector. In other words, the axes of the coordinate system are unevenly stretched, if $\alpha \neq 1$. Whereas if $\alpha = 1$ then we have the same model as before in (6.3.1), which means that our original model is nested in this more general class.

The $\mathbb{C}$VAR(1) with stretching factor will have the parametric form as follows

$$\boldsymbol{\psi}_t = \rho \boldsymbol{M}_{\alpha,\delta,\lambda} \boldsymbol{\psi}_{t-1} + \boldsymbol{\epsilon}_t = \rho \underbrace{\tfrac{1}{\sqrt{\alpha}} \boldsymbol{K}_\delta \boldsymbol{R}_\lambda \boldsymbol{K}_\delta^{\mathrm{H}} \boldsymbol{V}_\alpha}_{\boldsymbol{M}_{\alpha,\delta,\lambda}} \boldsymbol{\psi}_{t-1} + \boldsymbol{\epsilon}_t; \tag{6.4.1}$$

$$\begin{aligned}
\begin{bmatrix} \psi_t \\ \eta_t \end{bmatrix} &= \frac{\rho}{\sqrt{\alpha}} \begin{bmatrix} e^{i\delta} & 0 \\ 0 & e^{-i\delta} \end{bmatrix} \begin{bmatrix} \cos(\lambda) & \sin(\lambda) \\ -\sin(\lambda) & \cos(\lambda) \end{bmatrix} \begin{bmatrix} e^{-i\delta} & 0 \\ 0 & e^{i\delta} \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \eta_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t^{(1)} \\ \epsilon_t^{(2)} \end{bmatrix} \\
&= \frac{\rho}{\sqrt{\alpha}} \begin{bmatrix} \alpha\cos(\lambda) & \sin(\lambda)e^{i2\delta} \\ -\alpha\sin(\lambda)e^{-i2\delta} & \cos(\lambda) \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \eta_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t^{(1)} \\ \epsilon_t^{(2)} \end{bmatrix};
\end{aligned} \tag{6.4.2}$$

where the additional parameter is the stretching factor $\alpha \in \mathbb{R}$ and its stretching matrix $\boldsymbol{V}_\alpha$. This represents an affine transformation of stretching of the axes (if $\alpha \neq 1$ then the stretching is non-uniform). Additionally we also divide by factor $\sqrt{\alpha}$ in order to make the determinant of $\boldsymbol{M}_{\alpha,\delta,\lambda}$ equal to one.

The matrix $\boldsymbol{V}_\alpha$ causes a non-uniform (anisotropic) scaling transformation, if $\alpha \neq 1$. Since the other scaling factor is 1, this represents a directional scaling or stretching. This means that the time series vector at $t-1$, $\boldsymbol{\psi}_{t-1}$, is stretched in one direction and the shape of the object is changed. More general discussion on affine transformation is in Chapter 2.1.2.

Stationarity For a general $\mathbb{C}$VAR(1) model the stationarity condition is satisfied, if the roots of $\det\{\Phi(B)\} = 0$ are greater than one in absolute value [51], where $\Phi(B)$ is the characteristic polynomial. For this model this translates to a condition

that will not only depend on the damping factor ρ, but also the scaling factor α and frequency λ. The characteristic polynomial for this process is $\Phi(B) = \mathbf{I}_2 - \rho \boldsymbol{M}_{\alpha,\delta,\lambda} B = \mathbf{I}_2 - \frac{\rho}{\sqrt{\alpha}} \boldsymbol{K}_\delta \boldsymbol{R}_\lambda \boldsymbol{K}_\delta^{\mathrm{H}} \boldsymbol{V}_\alpha B$. Finding the roots of the determinant of the characteristic polynomial gives us the condition for stationarity for this models as

$$\left| \frac{\frac{\alpha+1}{\sqrt{\alpha}} \cos\lambda \pm \sqrt{\frac{(\alpha+1)^2}{\alpha} \cos^2\lambda - 4}}{2\rho} \right| > 1. \tag{6.4.3}$$

Autocovariance function Using the Yule-Walker equations (as introduced in Equation (6.3.14)) we express the ACV matrix of this model as

$$\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(\tau) = \begin{cases} \rho^2 \boldsymbol{M}_{\alpha,\delta,\lambda} \boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0) \boldsymbol{M}_{\alpha,\delta,\lambda}^{\mathrm{H}} + \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} & \text{for } \tau = 0, \\ \left(\frac{\rho}{\sqrt{\alpha}}\right)^{\tau} \boldsymbol{K}_\delta \left[\boldsymbol{R}_\lambda \boldsymbol{V}_\alpha\right]^{\tau} \boldsymbol{K}_\delta^{\mathrm{H}} \boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0) & \text{for } \tau > 0, \\ \left(\frac{\rho}{\sqrt{\alpha}}\right)^{|\tau|} \boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0) \boldsymbol{K}_\delta \left[\boldsymbol{V}_\alpha \boldsymbol{R}_\lambda^{\mathrm{T}}\right]^{|\tau|} \boldsymbol{K}_\delta^{\mathrm{H}} & \text{for } \tau < 0; \end{cases} \tag{6.4.4}$$

and the C-ACV matrix as

$$\tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}(\tau) = \begin{cases} \rho^2 \boldsymbol{M}_{\alpha,\delta,\lambda} \tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}(0) \boldsymbol{M}_{\alpha,\delta,\lambda}^{\mathrm{T}} + \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} & \text{for } \tau = 0, \\ \left(\frac{\rho}{\sqrt{\alpha}}\right)^{\tau} \boldsymbol{K}_\delta \left[\boldsymbol{R}_\lambda \boldsymbol{V}_\alpha\right]^{\tau} \boldsymbol{K}_\delta^{\mathrm{H}} \tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}(0) & \text{for } \tau > 0, \\ \left(\frac{\rho}{\sqrt{\alpha}}\right)^{|\tau|} \tilde{\boldsymbol{\Gamma}}_{\boldsymbol{\psi}}(0) \boldsymbol{K}_\delta^{\mathrm{H}} \left[\boldsymbol{V}_\alpha \boldsymbol{R}_\lambda^{\mathrm{T}}\right]^{|\tau|} \boldsymbol{K}_\delta & \text{for } \tau < 0. \end{cases} \tag{6.4.5}$$

Remark 6.4.1. *We analyse the expression for the ACV as function of τ and try to see its behaviour. We look only at the case when $\tau \geq 0$. Unfortunately, the expression does not have a clear closed form solution in τ*

$$\begin{aligned} \boldsymbol{\Gamma}_{\boldsymbol{\psi}}(\tau) &= \rho^\tau \boldsymbol{M}_{\alpha,\delta,\lambda}^{\tau} \boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0) \\ &= \left[\frac{\rho}{\sqrt{\alpha}} \boldsymbol{K}_\delta \boldsymbol{R}_\lambda \boldsymbol{K}_\delta^{\mathrm{H}} \boldsymbol{V}_\alpha\right]^{\tau} \boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0) \\ &= \left(\frac{\rho}{\sqrt{\alpha}}\right)^{\tau} \boldsymbol{K}_\delta \left[\boldsymbol{R}_\lambda \boldsymbol{V}_\alpha\right]^{\tau} \boldsymbol{K}_\delta^{\mathrm{H}} \boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0) \\ &= \left(\frac{\rho}{\sqrt{\alpha}} \begin{bmatrix} \alpha\cos(\lambda) & \sin(\lambda) e^{i2\delta} \\ -\alpha\sin(\lambda) e^{-i2\delta} & \cos(\lambda) \end{bmatrix}\right)^{\tau} \boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0). \end{aligned} \tag{6.4.6}$$

From the above Remark 6.4.1 we can notice that for every τ the covariance matrix $\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0)$, which does not depend on lag τ, is multiplied by an affine transformation matrix $\rho \mathbf{M}_{\alpha,\delta,\lambda}$. For every τ the covariance matrix is scaled (using the scaling factor

α), rotated (using the frequency λ) and damped (by the damping factor ρ). From this we can conclude that the trajectory of the ACVF matrix as a function of τ will follow an ellipse, although we are not able to show that in closed form as previously.

Remark 6.4.2. *We now perform a similar analysis for the forecast function to observe its behaviour. In this case the l-step ahead forecast of the model in* (6.4.1) *is*

$$\begin{aligned}
\widehat{\boldsymbol{\psi}}_t(l) = \begin{bmatrix}\widehat{\psi}_t(l)\\ \widehat{\eta}_t(l)\end{bmatrix} &= \mathbb{E}[\boldsymbol{\psi}_{t+l} \mid \boldsymbol{\psi}_t, \boldsymbol{\psi}_{t-1}, \ldots] \\
&= \rho \boldsymbol{M}_{\alpha,\delta,\lambda} \widehat{\boldsymbol{\psi}}_t(l-1) = \ldots = \rho^l \boldsymbol{M}^l_{\alpha,\delta,\lambda} \boldsymbol{\psi}_t \\
&= \Big[\frac{\rho}{\sqrt{\alpha}} \boldsymbol{K}_\delta \boldsymbol{R}_\lambda \boldsymbol{K}_\delta^H \boldsymbol{V}_\alpha\Big]^l \boldsymbol{\psi}_t \\
&= \Big(\frac{\rho}{\sqrt{\alpha}}\Big)^l \boldsymbol{K}_\delta \Big[\boldsymbol{R}_\lambda \boldsymbol{V}_\alpha\Big]^l \boldsymbol{K}_\delta^H \boldsymbol{\psi}_t \\
&= \left(\frac{\rho}{\sqrt{\alpha}} \begin{bmatrix} \alpha\cos(\lambda) & \sin(\lambda)e^{i2\delta} \\ -\alpha\sin(\lambda)e^{-i2\delta} & \cos(\lambda)\end{bmatrix}\right)^l \begin{bmatrix}\psi_t \\ \eta_t\end{bmatrix}
\end{aligned} \tag{6.4.7}$$

which, similar to the analysis of the ACVF before, does not have a nice closed form solution in forecast-time l.

From the above remark we see that the forecast vector $\widehat{\boldsymbol{\psi}}_t(l)$ will be at every l-step forecast scaled, phase-shifted, rotated, phase-shifted backwards and damped. Consequently the forecast vector will trace a trajectory of a damped ellipse as a function of the l-step ahead forecast. As we have already seen above in the case of the autocovariance ellipse we cannot show this in closed form.

Simulated example Also in this case, we simulate an example from the model with stretching factor in Equation (6.4.1). To be abel to compare the plots with the simulated example of the general model in Chapter 6.3.6, we use the same parameters as above and add the stretching factor, $\alpha = 2$.

Figure 6.9 shows the plots of the autocovariance ellipses of each of the four entries of the ACVF of the simulated model with stretching factor, and Figure 6.10 shows the plots of the forecast ellipses of each of the two entries of the forecast vector of the simulated model with stretching factor.

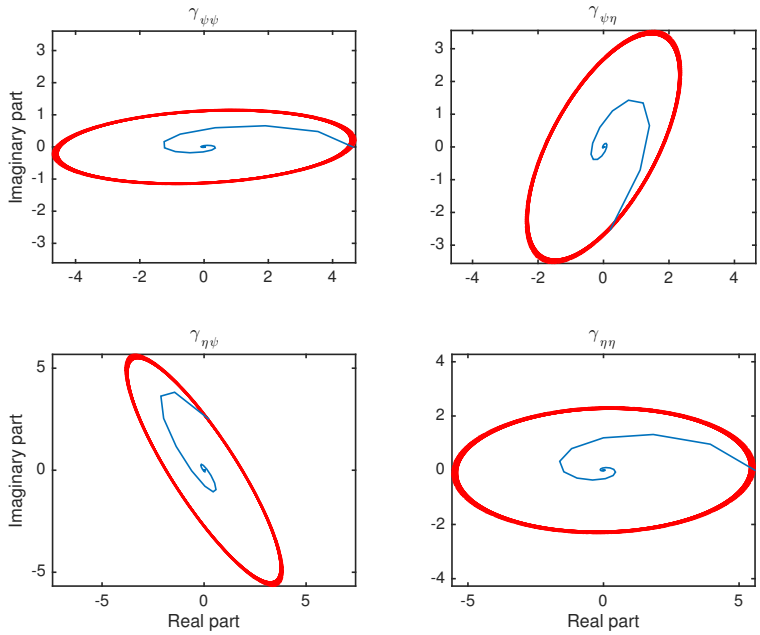


Figure 6.9: Plots of the autocovariance ellipse of the simulated $\mathbb{C}$VAR(1) time series model with stretching factor in Equation (6.4.1), showing the damped and non-damped ellipses from each of the entries of the ACV matrix through lags τ

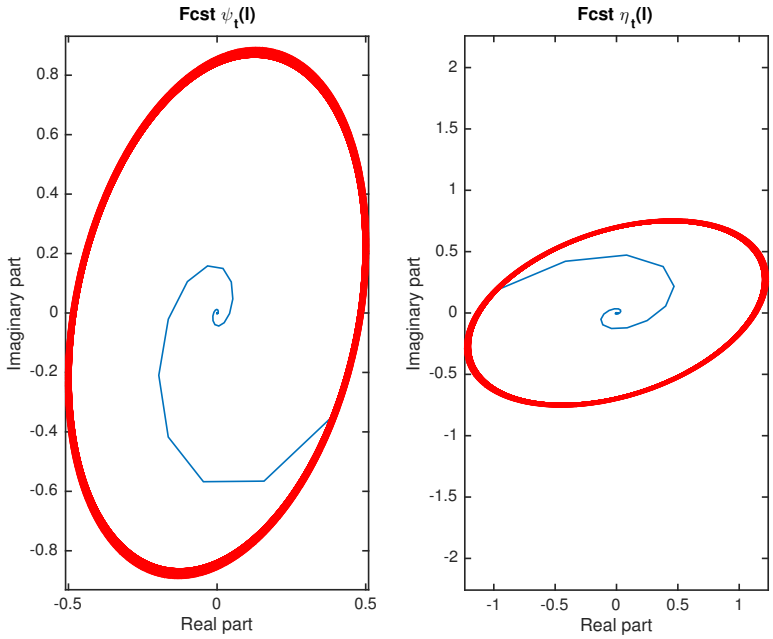


Figure 6.10: Plots of the forecast ellipses of the simulated $\mathbb{C}$VAR(1) time series model with stretching factor in Equation (6.4.1), showing the damped and non-damped ellipses from each of the the entries of the forecast vector through forecast-time l

We compare the simulated trajectories of the autocovariance and forecast ellipses between the general class model and the model with the stretching factor. In both cases the trajectories trace a mapped ellipse in the complex plane. The main difference for the autocovariance ellipses of the model with the stretching factor is that each of the four entries of the ACV matrix has different parameters of the ellipses. Both size and orientation are different, compared to the model without the stretching factor where the ellipses were same in size. We can observe similar differences for the forecast ellipse. In the case of the model with the stretching factor the orientation and size of the forecast ellipses are different, whereas in the case of the model without the stretching factor only the orientation was different.

6.4.2 Bivariate real-valued VAR(1) with stretching factor

Now we look at a similar parametric model as above with the main difference that it will be real-valued VAR(1). This is an extension to the model described in (6.1.1) and at the same time a constrained complex-valued model defined in Equation (6.4.1). We need to restrict the $\mathbb{C}$VAR(1) model parameter $\rho \boldsymbol{M}_{\alpha,\delta,\lambda}$ to be real-valued, which

we do by setting the phase shift factor to be $\delta = 0$ and making sure the noise term $\boldsymbol{\epsilon}_t$ is real-valued as well. By applying these restrictions we obtain the following model

$$\begin{aligned} \boldsymbol{\psi}_t &= \rho \boldsymbol{M}_{\alpha,\lambda} \boldsymbol{\psi}_{t-1} + \boldsymbol{\epsilon}_t = \rho \underbrace{\frac{1}{\sqrt{\alpha}} \boldsymbol{R}_\lambda \boldsymbol{V}_\alpha}_{\boldsymbol{M}_{\alpha,\lambda}} \boldsymbol{\psi}_{t-1} + \boldsymbol{\epsilon}_t, \\ \begin{bmatrix} \psi_t \\ \eta_t \end{bmatrix} &= \frac{\rho}{\sqrt{\alpha}} \begin{bmatrix} \alpha\cos(\lambda) & \sin(\lambda) \\ -\alpha\sin(\lambda) & \cos(\lambda) \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \eta_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t^{(1)} \\ \epsilon_t^{(2)} \end{bmatrix}, \end{aligned} \tag{6.4.8}$$

where $\{\boldsymbol{\epsilon}_t\} \in \mathbb{R}^2$. The stationarity of the model will also depend on the ρ, α and λ, where the condition will be the same as for the previous model in Equation (6.4.3).

The model parameter $\rho \boldsymbol{M}_{\alpha,\lambda} = \frac{\rho}{\sqrt{\alpha}} \boldsymbol{R}_\lambda \boldsymbol{V}_\alpha$ can be viewed as a linear transformation (a restricted case of an affine transformation). Due to the transformation of the matrix $\boldsymbol{V}_\alpha$ the vector $\boldsymbol{\psi}_{t-1}$ is stretched in one direction, so the shape of the object is changed. Based on the transformation sequence composed of non-uniform scaling (stretching) and rotation of the vector $\boldsymbol{\psi}_{t-1}$, we can conclude that the transformation causes the vector to trace a trajectory of an ellipse.

Complex composition To simplify the analysis of this process we construct a univariate complex-valued process $\{\xi_t\}$ from the entries of the bivariate real VAR(1) as $\xi_t = \psi_t + i\eta_t$. We also need to write the augmented vector as $\boldsymbol{\xi}_t = [\xi_t, \xi_t^*]^{\mathrm{T}} \in \mathbb{C}^2$, which can also be obtained as $\boldsymbol{\xi}_t = \mathbb{T}_2 \boldsymbol{\psi}_t$. The complex-valued error term is constructed as $\zeta_t = \epsilon_t^{(1)} + i\epsilon_t^{(2)}$, with its augmented vector $\boldsymbol{\zeta}_t = [\zeta_t, \zeta_t^*]^{\mathrm{T}} \in \mathbb{C}^2$. Using this notation and the model in (6.4.8) it follows that the complex-valued time series model $\{\xi_t\}$ can be expressed with its real-valued components in the following way

$$\begin{aligned} \xi_t = \psi_t + i\eta_t &= \rho\frac{\alpha}{\sqrt{\alpha}} e^{-i\lambda}\psi_{t-1} + i\rho\frac{1}{\sqrt{\alpha}} e^{-i\lambda}\eta_{t-1} + \zeta_t \\ &= \frac{\rho}{\sqrt{\alpha}} e^{-i\lambda}(\alpha\psi_{t-1} + i\eta_{t-1}) + \zeta_t \\ &= \frac{\rho}{2\sqrt{\alpha}} e^{-i\lambda}[\alpha(\xi_{t-1} + \xi_{t-1}^*) + (\xi_{t-1} - \xi_{t-1}^*)] + \zeta_t \\ &= \frac{\rho}{2\sqrt{\alpha}} e^{-i\lambda}[(\alpha+1)\xi_{t-1} + (\alpha-1)\xi_{t-1}^*] + \zeta_t, \end{aligned} \tag{6.4.9}$$

which in case if $\alpha = 1$, simplifies to $\xi_t = \rho e^{-i\lambda}(\psi_{t-1} + i\eta_{t-1}) + \zeta_t = \rho e^{-i\lambda}\xi_{t-1} + \zeta_t$.

Autocovariance function The ACV matrix of the model can be obtained by using the Yule-Walker equations (as introduced in Equation (6.3.14)) in the following way

$$\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(\tau) = \begin{cases} \frac{\rho^2}{\alpha}\boldsymbol{R}_\lambda \boldsymbol{V}_\alpha \boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0)\boldsymbol{V}_\alpha \boldsymbol{R}_\lambda^{\mathrm{T}} + \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} & \text{for } \tau = 0, \\ \left(\frac{\rho}{\sqrt{\alpha}}\right)^{\tau}\Big[\boldsymbol{R}_\lambda \boldsymbol{V}_\alpha\Big]^{\tau}\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0) & \text{for } \tau > 0, \\ \left(\frac{\rho}{\sqrt{\alpha}}\right)^{|\tau|}\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0)\Big[\boldsymbol{V}_\alpha \boldsymbol{R}_\lambda^{\mathrm{T}}\Big]^{|\tau|} & \text{for } \tau < 0. \end{cases} \tag{6.4.10}$$

We want to investigate how does the ACVF of the bivariate VAR(1) process evolve as we move through lags τ. Since the model is real-valued its ACV matrix will also be real-valued, so in order to observe the autocovariance ellipse, as per Definition 5.2.3, we look at pairs of its entries in the $x - y$ plane. Other option is to look at the ACVF of the complex composition signal $\{\xi_t\}$. If we look at the ACV matrix as function of lag τ we can notice that for every τ the ACVF is scaled, rotated and damped. Since the operation of scaling is non-uniform, or in other words directional, we can deduct that the shape of the ACVF through τ will be deformed.

The ACV matrix for the augmented vector of the complex composition $\{\boldsymbol{\xi}_t\}$ of our time series $\{\boldsymbol{\psi}_t\}$ can easily be obtained by using the real-to-complex transformation matrix $\mathbb{T}_2$ as

$$\boldsymbol{\Gamma}_{\boldsymbol{\xi}}(\tau) = \mathbb{T}_2 \boldsymbol{\Gamma}_{\boldsymbol{\psi}}(\tau)\mathbb{T}_2^{\mathrm{H}} = \left(\frac{\rho}{\sqrt{\alpha}}\right)^{\tau}\mathbb{T}_2\Big[\boldsymbol{R}_\lambda \boldsymbol{V}_\alpha\Big]^{\tau}\boldsymbol{\Gamma}_{\boldsymbol{\psi}}(0)\mathbb{T}_2^{\mathrm{H}}. \tag{6.4.11}$$

The stretching and rotation will make the ACVF trace out an elliptical trajectory that will be damped, with exponential decay, by the factor $\rho < 1$. Unfortunately, as we have seen in the previous chapter, due to the stretching factor α none of these expressions will have a simplified closed form solution.

In the special case when the VAR(1) is without the stretching ($\alpha = 1$), same as in the model introduced at the beginning of our discussion about parametric models in (6.1.1), the model parameter becomes just a rotation matrix $\boldsymbol{R}_\lambda$. Now the ACVF and the complementary ACVF of the complex composition $\{\xi_t\}$ in lag-domain traces

out a trajectory of a circle in the complex plane, as shown below

$$\begin{aligned}
\gamma_\xi(\tau) &= \gamma_{\psi\psi}(\tau) + \gamma_{\eta\eta}(\tau) + i(\gamma_{\eta\psi}(\tau) - \gamma_{\psi\eta}(\tau)) \\
&= \rho^\tau[\gamma_{\psi\psi}(0) + \gamma_{\eta\eta}(0) + i(\gamma_{\eta\psi}(0) - \gamma_{\psi\eta}(0))]e^{i\tau\lambda} \\
&= \rho^\tau\gamma_\xi(0)e^{i\tau\lambda}; \\
\tilde{\gamma}_\xi(\tau) &= \gamma_{\psi\psi}(\tau) - \gamma_{\eta\eta}(\tau) + i(\gamma_{\eta\psi}(\tau) + \gamma_{\psi\eta}(\tau)) \\
&= \rho^\tau[\gamma_{\psi\psi}(0) - \gamma_{\eta\eta}(0) + i(\gamma_{\eta\psi}(0) + \gamma_{\psi\eta}(0))]e^{i\tau\lambda} \\
&= \rho^\tau\tilde{\gamma}_\xi(0)e^{-i\tau\lambda}.
\end{aligned} \tag{6.4.12}$$

Forecast The l-step ahead forecast for the time series model in (6.4.8) is expressed in the following way

$$\begin{aligned}
\widehat{\boldsymbol{\psi}}_t(l) &= \begin{bmatrix}\widehat{\psi}_t(l) \\ \widehat{\eta}_t(l)\end{bmatrix} = \mathbb{E}[\boldsymbol{\psi}_{t+l} \mid \boldsymbol{\psi}_t, \boldsymbol{\psi}_{t-1}, \boldsymbol{\psi}_{t-2}, \ldots] \\
&= \frac{\rho}{\sqrt{\alpha}}\boldsymbol{R}_\lambda\boldsymbol{V}_\alpha\widehat{\boldsymbol{\psi}}_t(l-1) = \ldots \\
&= \left(\frac{\rho}{\sqrt{\alpha}}\boldsymbol{R}_\lambda\boldsymbol{V}_\alpha\right)^l\boldsymbol{\psi}_t \\
&= \left(\frac{\rho}{\sqrt{\alpha}}\begin{bmatrix}\alpha\cos(\lambda) & \sin(\lambda) \\ -\alpha\sin(\lambda) & \cos(\lambda)\end{bmatrix}\right)^l\begin{bmatrix}\psi_t \\ \eta_t\end{bmatrix},
\end{aligned} \tag{6.4.13}$$

which shows that the forecast vector $\{\widehat{\boldsymbol{\psi}}_t(l)\}$, will, at every l-step forecast, be stretched, rotated and damped. Here we aim at describing the forecast ellipse for this time series model in light of the Definition 5.2.2. Similar to above we can conclude that the forecast vector will also trace a trajectory of a damped ellipse in the $x-y$ plane as a function of the forecast-time l, but unfortunately this cannot be shown in a closed form solution.

In the special case when the VAR(1) is without the stretching ($\alpha = 1$), same as in the model introduced at the beginning of our discussion about parametric models in (6.1.1), the l-step ahead forecast will simplify to

$$\widehat{\boldsymbol{\psi}}_t(l) = \begin{bmatrix}\widehat{\psi}_t(l) \\ \widehat{\eta}_t(l)\end{bmatrix} = \rho^l\begin{bmatrix}\psi_t\cos(l\lambda) + \eta_t\sin(l\lambda) \\ \eta_t\cos(l\lambda) - \psi_t\sin(l\lambda)\end{bmatrix}, \tag{6.4.14}$$

which can be analysed using the complex composition ξ_t as

$$\widehat{\xi}_t(l) = \widehat{\psi}_t + i\widehat{\eta}_t = \rho^l[\psi_t e^{-il\lambda} + i\eta_t e^{-il\lambda}] = \rho^l[\psi_t + i\eta_t]e^{-il\lambda}. \tag{6.4.15}$$

This clearly shows that conditionally on knowing the present and the past, we expect the trajectory of the time series model will follow a circular shape.

Chapter 7

Conclusion and discussion

The contribution of our research is that we have provided a synthesis of the existing elliptical definitions, introduced two new definitions of the ellipse and proposed a bivariate parametric model for the modelling of stochastic oscillations. In this thesis we have first reviewed literature and provided background theory for a holistic understanding of the research topic. Since, our research is centred around complex-valued stochastic processes, we provided background theory about complex-valued random vectors and random processes. We also provided some basic theory about time series analysis as defined for complex-valued time points. All this background theory was intended to facilitate the understanding of complex-valued oscillatory models and elliptical structure in the rest of our work.

After the initial background theory we provided an introduction to elliptical models and oscillations. This introduced the reader to general processes as elliptical structures. The central part of our research were the definitions of the elliptical models. We identified and outlined several definitions from literature. The existing definitions are very intuitive in the case of deterministic signals both in time and frequency domains. As noticed, for stochastic processes the approach of how to define the ellipse is a bit more challenging. Most of these definitions are in the frequency domain, which we think, is less intuitive and less useful due to the inherent randomness of frequency domain definitions. Our aim was to identify a non-random time domain definition of stochastic processes. That is why we proposed two new elliptical definitions, the autocovariance ellipse and the forecast ellipse. Both of these definitions are in time domain and are deterministic representations. Due to these two properties of the representations we think they are a useful contribution to the

research in this field. Moreover, we believe this field of research is partially unclear because theree is no synthesis of the different ellipse definitions in the literature. Thus, our aim was to introduce clarity by providing a summary and illustrate all the definitions at the same place. We believe this is also a very important contribution of our work.

To complete the theoretic research and the definitions of ellipses, we also proposed parametric models for modelling of stochastic oscillations. We proposed a bivariate complex-valued model that is equipped to accommodate complex-valued time points. Its parameters control the shape and form of the oscillation and the spectral density function. This model also illustrated the concepts developed theoretically, and showd how the elliptical definitions work in practice. The development of this parametric model was based on affine transformations. We used the different affine transformations to create the stochastic oscillation and to control its shape. Apart from having proposed and analysed the bivariate complex-valued model, we also looked at the restrictions and extensions of that model that stems from affine transformation theory. We explored restrictions in the direction of real-valued processes, whereas extensions can be made by introducing additional parameters. We based ourselves in the context of affine transformations for the purpose of building the stochastic oscillation models.

We believe, that the two new elliptical definitions we added to the theory in this field of research, are very intuitive and add many possible ways to analysing stochastic oscillations. They can be used in statistical research and in signal processing. There are several applied fields that use the idea of elliptical models to understand and model natural behaviour, such as oceanography, econometrics, seismology, etc. The fact that we have also proposed a parametric model gives additional usefulness to the definitions. This model is bivariate and complex-valued, which is very useful in applications where we have coupled phenomena that come in bivariate series. Such as, for example, in oceanography, where we model the north-south and east-west components of ocean currents or surface winds. The proposed model is also useful because it is relatively easy to non-parametrically estimate the autocovariance function of a time series and observe its behaviour. Like that we can easily observe the elliptical properties of the natural phenomena in light of the autocovariance ellipse.

Nevertheless, we also have to be aware that the proposed definitions and the parametric models have limitations. The limitation of the theoretical definitions

we proposed, is that the autocovariance and/or the forecast function of the model need to be defined. Ideally we should find such a model that allows for a closed form solution of these two functions. The limitation of our research is that we do not provide insight into inference of these models and we do not provide a practical application. At this point of research we have decided to illustrate the examples and the proposed models with simulated data only. We believe that this provides enough insight into how the models work. In the future an applied example would be very illustrative.

Last but not least, by proposing the two new theoretical definitions we created space for new research opportunities and further exploration in this direction of time series analysis. The immediate extensions for future work in this area are both in the theoretical and modelling work. On one hand, we would be interested to extend these definitions to non-stationary elliptical models. Since the autocovariance or the forecast functions of a non-stationary process change over time, it would require a local time extension of the above definitions. On the other hand, for the modelling part the immediate extension would be to increase the dimensionality of the models. We started with bivariate models, which are the initial step from univariate analysis. For the future we would be interested to define n- dimensional stochastic oscillation models.

7.1 Discussion on feasibility of the multivariate extension

The increase of dimensionality for these models would mean we would need to break away form the natural bivariate structure introduced above. There are several considerations of a multivariate extension, such as how to construct an intuitive parametric model, how to make sure such model would not suffer from too many parameters to estimate, and how do we understand an ellipse in higher-dimensions.

The main challenge for multivariate extensions is that such models suffer the 'curse of dimensionality'. In other words, the higher-dimensional the structure is the more parameters are needed to estimate. This causes that the identifiability of the parameters needed becomes a serious issue. If such models were to be proposed the identifiability would need to be researched more in detail and possibly a simplified structure proposed.

Apart from the identifiability also the intuitiveness and applicability of higher-dimensional models becomes an issue. To construct higher-dimensional models we

would need to break the simple bivariate structure of the affine transformation matrices. These matrices have an intuitive explanation, as well as they also create the oscillation needed to model cyclical behaviour. A different structure with similar properties would need to be proposed to recreate higher-dimensional structures that would replicate the same effect.

Finally, understanding oscillations and consequently ellipses in higher dimensions is also a challenge. Bivariate structure is very natural and is inherent in our complex-valued models. Higher-dimensional ellipses are not easily pictured in a plot, hence it would be more difficult to see that the structure is following an n-dimensional ellipse.

Based on this multivariate extension might either be very complex or not possible at all, either due to lack of identifiability or the perspective of recreating the desired structure. Usefulness of such extension might also be questionable. Hence, at this point we propose for one to divide the observations in pairwise problems and use the bivariate models to analyse higher dimensional models.

7.2 Discussion on non-stationary extension

Stationary models are desirable in time series analysis as already outlined in the thesis above (see 2.2.2). However, if we cannot assume stationarity it means that the model properties change over time. In order to model this in a parametric model we would need to introduce a local time extension of the models proposed above.

To illustrate the local expansion, we analyse a univariate ACVF that can be easily linked to the above parametric models. Local time expansion of the ACVF can be represented with a series of deviations from a set of constant-amplitude oscillations evolving with some common instantaneous frequency $\omega(\tau)$. Following the approach taken in [31] let us represent the ACVF of the process of interest ψ_t in the neighbourhood of lag τ with deviations from a set of sinusoids with common time-varying frequency $\omega(\tau)$.

Assuming the ACVF is continuous in lag $\tau \in \mathbb{R}$ we are interested in analysing the ACVF in the vicinity offset by h. We start by expressing the ACVF $\gamma(\tau)$ shifted by 'local lag' in the form

$$\gamma(\tau + h) = e^{i\omega(\tau)h}[e^{-i\omega(\tau)h}\gamma(\tau + h)].$$

Next we use Taylor expansion, as defined in [1, p. 880], on the term in the square brackets above with respect to $h = 0$, to obtain the *local modulation expansion*

$$\gamma(\tau + h) = e^{i\omega(\tau)h}\big\{\gamma(\tau) + h\bar{\gamma}_1(\tau;\omega(t)) + \tfrac{1}{2}h^2\bar{\gamma}_2(\tau;\omega(t)) + \zeta_3(\tau, h;\omega(t))\big\} \quad (7.2.1)$$

The $\bar{\gamma}_p(\tau;\omega(t))$ can be called a *deviation factor* and defined as

$$\bar{\gamma}_p(\tau;\omega(t)) = \frac{\partial^p}{\partial h^p}[e^{-i\omega(\tau)h}\gamma(\tau + h)]\Big|_{h=0} \quad (7.2.2)$$

and the remainder term or *inaccuracy* follows the Lagrange form of remainder in Taylor series for some unknown point u $(0 < u < h)$, as

$$\zeta_3(\tau;\omega(t)) = \frac{1}{6}h^3\frac{\partial^3}{\partial h^3}[e^{-i\omega(\tau)h}\gamma(\tau + h)]\Big|_{h=u}. \quad (7.2.3)$$

The first derivative is usually referred to as the velocity and the second derivative as the acceleration. In the above Equation (7.2.1) we can think of the $\bar{\gamma}_1(\tau;\omega(t))$ as the intrinsic ACVF velocity and the $\bar{\gamma}_2(\tau;\omega(t))$ as the intrinsic ACVF acceleration. Whereas the $\zeta_3(\tau, h;\omega(t))$ is the jerk, but since we assume smoothness of the function this remainder should be small and the function should not exhibit too much jerkiness.

This analysis shows that in the non-stationary expansion the qualities related to the models, such as the ACVF, would need to be examined on their behaviour in the neighbourhood of each lag τ. This largely reduces the usefulness of such models and introduces many additional factors to consider during the modelling.

Appendix A

Notation and abbreviations

Notation

$\mathbf{x}$	denotes a real-valued non-random vector in $\mathbb{R}^n$
$\mathbf{u}$	denotes a complex-valued non-random vector in $\mathbb{C}^n$
$\mathbf{X}$	denotes a real-valued random vector in $\mathbb{R}^n$
$\mathbf{U}$	denotes a complex-valued random vector in $\mathbb{C}^n$
$\{\mathbf{u}_t\}$	denotes a complex-valued deterministic signal, such as $\mathbf{u}_t = \mathbf{x}_t + i\mathbf{y}_t$
$\{\mathbf{U}_t\}$	denotes a complex-valued random time series, such as $\mathbf{U}_t = \mathbf{X}_t + i\mathbf{Y}_t$
$\mathbf{V}_t$	denotes the real augmented vector, such as $\mathbf{V}_t = [\mathbf{X}_t^{\mathrm{T}}, \mathbf{Y}_t^{\mathrm{T}}]^{\mathrm{T}}$
$\mathbf{W}_t$	denotes the complex augmented vector composed of the complex vector and its complex conjugate, such as $\mathbf{W}_t = [\mathbf{U}_t^T, \mathbf{U}_t^H]^T$
$\mathbf{I}_n$	denotes the $n \times n$ identity matrix
$\det\{\cdot\}$	determinant of a matrix
$\mathrm{tr}\{\cdot\}$	trace of a matrix
$\mathrm{var}\{\cdot\}$	variance operator
$\mathrm{cov}\{\cdot,\cdot\}$	covariance operator
$\mathrm{mse}\{\cdot\}$	mean square error
τ	denotes lag in the auto-covariance function
B	lag- (or back-) shift operator
$\mathbf{\Sigma}_\epsilon$	covariance matrix of the vector random noise ϵ
$\gamma_U(\tau)$	autocovariance function (ACVF) of univariate process $\{U_t\}$ at lag τ
$\tilde{\gamma}_U(\tau)$	complimentary autocovariance function (C-ACVF) of univariate process $\{U_t\}$ at lag τ

$\mathbf{\Gamma}_U(\tau)$	autocovariance function (ACVF) matrix of a complex-valued process $\{U_t\}$ at lag τ
$\tilde{\mathbf{\Gamma}}_U(\tau)$	complimentary autocovariance function (C-ACVF) matrix of a complex-valued process $\{U_t\}$ at lag τ
$\mathbf{\Gamma}_W(\tau)$	augmented autocovariance matrix of a complex-valued process $\{U_t\}$ defined by using the augmented vector $\mathbf{W}_t$ at lag τ
$Z_U(f)$	spectral process for complex-valued random process $\{U_t\}$ with orthogonal increments $dZ_U(f)$
f	frequency measured in cycles per unit, where $f = \omega/2\pi$
ω	angular frequency, where $\omega = 2\pi f$
$S_U(f)$	spectral density function (SDF) of complex time series $\{U_t\}$ at frequency f
$\tilde{S}_U(f)$	complimentary spectral density function (C-SDF) of complex time series $\{U_t\}$ at frequency f
$\mathbf{S}_W(f)$	augmented spectral matrix of the complex time series $\{U_t\}$ defined by using the augmented vector $\mathbf{W}$ at frequency f
$\mathfrak{S}_U(f)$	spectral distribution function of time series $\{U_t\}$
$\mathcal{C}$	spectral coherence
$\mathcal{H}\{x_t\}$	discrete Hilbert transform of signal $\{x_t\}$
$x_{+,t}$	analytic signal created from real vector signal $\{x_t\}$
$\mathfrak{H}(\cdot)$	unit step function
$\{\boldsymbol{\epsilon}_t\}$	$n \times 1$ vector of innovation process (error term), usually i.i.d. N($\mathbf{0}$, $\mathbf{\Sigma}_\epsilon$)
$\widehat{X}_t(l)$	l-step ahead forecast of vector time series $\{X_t\}$
$\hat{x}(f)$	discrete Fourier transform of signal $\{x_t\}$
$\boldsymbol{\mu}_t$	vector of trend components of a composite stochastic time series model
$\boldsymbol{\gamma}_t$	vector of seasonal components of a composite stochastic time series model
$\boldsymbol{\psi}_t$	vector of cyclical components of a composite stochastic time series model

Abbreviations

RV	random vector
CW	clockwise
CCW	counter-clockwise
PDF	probability density function
ACV	autocovariance
(C-)ACVF	(complimentary) autocovariance function
ACVS	autocovariance sequence
(C-)SDF	(complimentary) spectral density function
WSS	wide-sense stationary
AR(p)	autoregressive process of order p
AR(p)	complex-valued autoregressive process of order p
MA(q)	moving average process of order q
ARMA(p,q)	autoregressive moving average process of orders p and q
VAR(p)	vector autoregressive process of order p
$\mathbb{C}$AR(p)	complex-valued vector autoregressive process of order p
STSM	structural time series model
i.i.d.	independent and identically distributed

Appendix B

Measures of size and ellipticity

1. The area of ellipse:

$$\begin{aligned} Area = AB\pi &= (A_+ + A_-)|A_+ - A_-|\pi = |A_+^2 - A_-^2|\pi \\ &= |A_x A_y \sin\varphi|\pi \\ &= \mathcal{A}^2 \cos\chi \sin\chi\pi = \tfrac{1}{2}\mathcal{A}^2 \sin(2\chi)\pi \end{aligned} \tag{B.0.1}$$

2. The root-mean-square amplitude:

$$\begin{aligned} \kappa = \sqrt{\frac{A^2 + B^2}{2}} &= \mathcal{A}/\sqrt{2} \\ &= \sqrt{\frac{A_x^2 + A_y^2}{2}} \\ &= \sqrt{A_+^2 + A_-^2} \end{aligned} \tag{B.0.2}$$

3. The ellipse parameter:

$$\begin{aligned} \lambda = \pm\frac{A^2 - B^2}{A^2 + B^2} &= \pm\cos 2\chi \\ &= \pm\frac{2A_+A_-}{A_+^2 + A_-^2} \\ &= \frac{A_x^2 - A_y^2}{A_x^2 + A_y^2}\sec 2\theta \end{aligned} \tag{B.0.3}$$

4. The eccentricity:

$$\begin{aligned}\mathcal{E} = \pm\sqrt{1 - \frac{B^2}{A^2}} &= \pm\frac{\sqrt{\cos 2\chi}}{\cos^2\chi} \\ &= \pm\frac{2\sqrt{A_+ A_-}}{A_+ + A_-} \\ &= \pm\sqrt{\frac{2|(A_x^2 - A_y^2)\sec 2\theta|}{(A_x^2 + A_y^2 + |(A_x^2 - A_y^2)\sec 2\theta|}}\end{aligned} \tag{B.0.4}$$

5. The (signed) *aspect ratio* is

$$\begin{aligned}\mathcal{AR} = \pm\frac{B}{A} &= \tan\chi \\ &= \frac{A_+ - A_-}{A_+ + A_-} \\ &= \pm\sqrt{\frac{A_x^2 + A_y^2 - |(A_x^2 - A_y^2)\sec 2\theta|}{A_x^2 + A_y^2 + |(A_x^2 - A_y^2)\sec 2\theta|}}\end{aligned} \tag{B.0.5}$$

References

[1] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York, 1972.

[2] T. Adali, P.J. Schreier, and L.L. Scharf. Complex-Valued Signal Processing: The Proper Way to Deal With Impropriety. *IEEE Trans. Signal Process.*, 59(11):5101–5125, 2011.

[3] P.O. Amblard, M. Gaeta, and J.L. Lacoume. Statistics for complex variables and signals Part I: variables. *Signal Process.*, 53(1):1–13, 1996.

[4] C.B. Bell and E.P. Smith. Infrence for non-negative autoregressive schemes. *Communications in Statistics - Theory and Methods*, 15(8):2267–2293, 1986.

[5] P. Bloomfield. *Fourier Analysis of Time Series: An Introduction*. Wiley series in probability and statistics: Applied probability and statistics. Wiley, 2004.

[6] M. Born, E. Wolf, and A.B. Bhatia. *Principles of optics: Electromagnetic theory of propagation, interference and diffraction of light*. Cambridge University Press, 1999.

[7] G.E.P. Box, G.M. Jenkins, and G.C. Reinsel. *Time series analysis: forecasting and control*. Forecasting and Control Series. Prentice Hall, 1994.

[8] P.J. Brockwell and R.A. Davis. *Time Series: Theory and Methods*. Springer Series in Statistics. Springer, 2009.

[9] L. Cohen. *Time-frequency analysis*. Prentice Hall, Upper Sadler River, NJ, 1995.

[10] E. Collett. *Field guide to polarization*. SPIE field guides. SPIE Press, 2005.

[11] L. Dewald and P. Lewis. A new laplace second-order autoregressive time-series model–nlar(2). *Information Theory, IEEE Transactions on*, 31(5):645–651, Sep 1985.

[12] J. L. Doob. The elementary gaussian processes. *The Annals of Mathematical Statistics*, 15(3):229–282, 09 1944.

[13] W.J. Emery and R.E. Thomson. *Data Analysis Methods in Physical Oceanography*. Elsevier, New York, USA, 2004.

[14] D.H. Evans, W.N. McDicken, Skidmore R., and J.P. Woodcock. *Doppler Ultrasound: Physics, Instrumentation, and Clinical Applications*. A Wiley medical publication. Wiley, 1989.

[15] D. Gabor. Theory of communication. *Proc. Inst. Elect. Eng.*, 93(26):429–457, 1946.

[16] A. Ghatak. *Optics*. Tata McGraw-Hill, 2005.

[17] G. K. Grunwald, R. J. Hyndman, and L. M. Tedesco. A unified view of linear ar(1) models. Technical report, 1995.

[18] J.D. Hamilton. *Time series analysis*. Princeton University Press, 1994.

[19] A.C. Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, 1990.

[20] A.C. Harvey and S.J. Koopman. Multivariate Structural Time Series Models. In C. Heij et.al., editor, *System Dynamics in Economic and Financial Models*, pages 269–298. John Wiley & Sons Ltd, 1997.

[21] A.C. Harvey and N. Shephard. 10 structural time series models. In *Econometrics*, volume 11 of *Handbook of Statistics*, pages 261–302. Elsevier, 1993.

[22] R.J. Hodrick and E.C. Prescott. Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1–16, 1997.

[23] G. Johnson. The Complex (and Circular) Argument Continues. *IEEE Signal Process. Mag.*, 13(5):42–44, 1996.

[24] A.G. Jones. On the Difference Between Polarizaton and Coherence. *J. Geophys.*, 45:223–229, 1979.

[25] S.C. Kak. The discrete hilbert transform. *Proceedings of the IEEE*, 58(4):585–586, 1970.

[26] M. Lemoine. A model of the stochastic convergence between business cycles. Documents de Travail de l'OFCE 2005-05, Observatoire Francais des Conjonctures Economiques (OFCE), 2005.

[27] J.M. Lilly. Modulated oscillations in three dimensions. *IEEE Trans. Signal Process.*, 59(12):5930–5943, 2011.

[28] J.M. Lilly. The unity of instantaneous spectral moments and the physical moments of a time-varying ellipse. Submitted to IEEE Trans. Signal Process., 2012.

[29] J.M. Lilly and J. Gascard. Nonlinear processes in geophysics wavelet ridge diagnosis of time-varying elliptical signals with application to an oceanic eddy. *Nonlin. Processes Geophys.*, 13:467–483, 2006.

[30] J.M. Lilly and S.C. Olhede. Bivariate Instantaneous Frequency and Bandwidth. *IEEE Trans. Signal Process.*, 58:591–603, 2010.

[31] J.M. Lilly and S.C. Olhede. Analysis of Modulated Multivariate Oscillations. *IEEE Trans. Signal Process.*, 60(2):600–612, 2012.

[32] I.V. Lindell. *Methods for Electromagnetic Field Analysis*. IEEE Press series on electromagnetic wave theory. IEEE Press, 1996.

[33] H. Lütkepohl. *New introduction to multiple time series analysis*. New York, 2005.

[34] T. Medkour and A.T. Walden. Statistical properties of the estimated degree of polarization. *IEEE Trans. Signal Process.*, 56(1):408–414, 2008.

[35] K.S. Miller. *Complex stochastic processes: an introduction to theory and application*. Addison-Wesley Pub. Co., Advanced Book Program, 1974.

[36] C.N.K. Mooers. A technique for the cross spectrum analysis of pairs of complex-valued time series, with emphasis on properties of polarized components and rotational invariants. *Deep-Sea Research*, 20(12):1129–1141, 1973.

[37] D.F. Morrison. Expectations and covariances of serial and cross-correlation coefficients in a complex stationary time series. *Biometrika*, 50(1/2):213–216, 1963.

[38] F.D. Neeser and J.L. Massey. Proper complex random processes with applications to information theory. *IEEE Trans. Inf. Theory*, 39(4):1293–1302, 1993.

[39] S. C. Olhede and A. T. Walden. Noise reduction in directional signals illustrated on quadrature doppler ultrasound. *IEEE Trans. Biomed. Eng.*, 50:51–57, 2003.

[40] S.C. Olhede. On Probability Density Functions for Complex Variables. *IEEE Trans. Inf. Theory*, 52(3):1212–1217, 2006.

[41] S.C. Olhede and A.T. Walden. Local directional denoising. *IEEE Trans. Signal Process.*, 53(12):4725–4730, 2005.

[42] D.B. Percival and A.T. Walden. *Spectral analysis for physical applications: multitaper and conventional univariate techniques*. Cambridge University Press, 1993.

[43] B. Picinbono. On Circularity. *IEEE Trans. Signal Process.*, 42(12):3473–3482, 1994.

[44] B. Picinbono. Second-Order Complex Random Vectors and Normal Distributions. *IEEE Trans. Signal Process.*, 44(10):2637–2640, 1996.

[45] B. Picinbono and P. Bondon. Second-Order Statistics of Complex Signals. *IEEE Trans. Signal Process.*, 45(2):411–420, 1997.

[46] B. Picinbono and M. Bouvet. Complex white noises and autoregressive signals. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '84.*, volume 9, pages 596–599, mar 1984.

[47] M.A. Poletti. The homomorphic analytic signal. *IEEE Trans. Signal Process.*, 45(8):1943–1953, 1997.

[48] M.B. Priestley. *Spectral analysis and time series*. Number v. 1-2 in Probability and mathematical statistics. Academic Press, 1982.

[49] B.G. Quinn and E.J. Hannan. *The Estimation and Tracking of Frequency*. Cambridge University Press, Cambridge, UK, 2001.

[50] P.S. Rao and D.H. Johnson. A first-order ar model for non-gaussian time series. In *Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on*, pages 1534–1537 vol.3, Apr 1988.

[51] G.C. Reinsel. *Elements of multivariate time series analysis*. Springer Series in Statistics. Springer, 2003.

[52] P. Rubin-Delanchy and A. T. Walden. Simulation of Improper Complex-Valued Sequences. *IEEE Trans. Signal Process.*, 55:5517–5521, 2007.

[53] P. Rubin-Delanchy and AT. Walden. Kinematics of Complex–Valued Time Series. *IEEE Trans. Signal Process.*, 56:4189–4198, 2008.

[54] G. Rünstler. Modelling phase shifts among stochastic cycles. *Econome. J.*, 7(1):232–248, 2004.

[55] J.C. Samson. Matrix and Stokes vector representations of detectors for polarized waveforms: theory, with some applications to teleseismic waves. *Geophys. J. R. Astr. Soc.*, 51(3):583–603, 1977.

[56] J.C. Samson. Comments on Polarization and Coherence. *J. of Geophys.*, 48:195–198, 1980.

[57] J.C. Samson and J.V. Olson. Some comments on the descriptions of the polarization states of waves. *Geophys. J. R. Astr. Soc.*, 61(1):115–129, 1980.

[58] S. Sanei and J. Chambers. *EEG Signal Processing*. Wiley, New York, USA, 2007.

[59] P.J. Schreier. Polarization Ellipse Analysis of Nonstationary Random Signals. *IEEE Trans. Signal Process.*, 56(9):4330–4339, 2008.

[60] P.J. Schreier. A Unifying Discussion of Correlation Analysis for Complex Random Vectors. *IEEE Trans. Signal Process.*, 56(4):1327–1336, 2008.

[61] P.J. Schreier and L.L. Scharf. Second-Order Analysis of Improper Complex Random Vectors and Processes. *IEEE Trans. Signal Process.*, 51(3):714–725, 2003.

[62] P.J. Schreier and L.L. Scharf. *Statistical signal processing of complex-valued data: the theory of improper and noncircular signals.* Cambridge University Press, 2010.

[63] C. H. Sim. First-order autoregressive logistic processes. *Journal of Applied Probability*, 30(2):pp. 467–470, 1993.

[64] C.G. Small. *The Statistical Theory of Shape.* Springer series in statistics. Springer, 1996.

[65] D.J. Thomson. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, 70(9):1055–1096, 1982.

[66] S. Tou. *Visualization of Fields and Applications in Engineering.* John Wiley & Sons, 2011.

[67] D. Vakman. On the analytic signal, the Teager-Kaiser energy algorithm, and other methods for defining amplitude and frequency. *IEEE Trans. Signal Process.*, 44(4):791–797, 1996.

[68] H. van Haren and C. Millot. Rectilinear and circular inertial motions in the western mediterranean sea. *Deep-Sea Res. I*, 51(11):1441–1455, 2004.

[69] A.T Walden. Rotary components, random ellipses and polarization: a statistical perspective. *Phil Trans R Soc A*, 371(1984), 2013.

[70] A.M. Walker. On the estimation of a harmonic component in a time series with stationary independent residuals. *Biometrika*, 58(1):21–36, 1971.

[71] E. Wolf. Coherence Properties of Partially Polarized Electromagnetic Radiation. *Il Nuovo Cimento (1955-1965)*, 13:1165–1181, 1959.