

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

DEPARTAMENTO DE ELECTRÓNICA
VALPARAÍSO, CHILE



UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

IDENTIFICACIÓN DE MODELOS DE VIBRACIONES PARA CONTROL DE SISTEMAS DE ÓPTICA ADAPTATIVA

Tesis de Grado presentada por

Karen Fernanda González Pérez

como requisito parcial para optar al título de

Ingeniero Civil Electrónico

y al grado de

Magíster en Ciencias de la Ingeniería Electrónica

Supervisor Principal

Dr. Juan Carlos Agüero Vásquez

Co-supervisor

Dr. Rodrigo Carvajal Guerra

Marzo 2018

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

Departamento de Electrónica

**Identificación de modelos de vibraciones para
control de sistemas de óptica adaptativa**

Tesis de Grado presentada por

Karen Fernanda González Pérez

como requisito parcial para optar al título de

Ingeniero Civil Electrónico

y al grado de

Magíster en Ciencias de la Ingeniería Electrónica

Supervisor Principal

Dr. Juan Carlos Agüero Vásquez.

Co-supervisor

Dr. Rodrigo Carvajal Guerra.

Valparaíso, Marzo 2018.

Tesis:

IDENTIFICACIÓN DE MODELOS DE VIBRACIONES PARA CONTROL DE SISTEMAS DE ÓPTICA ADAPTATIVA

Autor:

Karen Fernanda González Pérez

TRABAJO DE TESIS, presentado en cumplimiento parcial de los requisitos para el título de Ingeniero Civil Electrónico y el grado de Magíster en Ciencias de la Ingeniería Electrónica de la Universidad Técnica Federico Santa María, Valparaíso, Chile.

Supervisor Principal,

Dr. Juan Carlos Agüero Vásquez.

Co-Supervisor,

Dr. Rodrigo Carvajal Guerra.

Revisor Interno,

Dr. Francisco Vargas Parra.

Revisor Externo,

Dr. Pedro Escárate Monetta.

Valparaíso, Marzo 2018.

Identificación de modelos de vibraciones para control de sistemas de óptica adaptativa

Tesis de Grado presentada por
Karen Fernanda González Pérez
como requisito parcial para optar al título de
Ingeniero Civil Electrónico
y al grado de
Magíster en Ciencias de la Ingeniería Electrónica

Supervisor Principal
Dr. Juan Carlos Agüero Vásquez

Co-supervisor
Dr. Rodrigo Carvajal Guerra

Revisor Interno
Dr. Francisco Vargas Parra

Revisor Externo
Dr. Pedro Escárate Mónetta.

RESUMEN

En todos los principales observatorios astronómicos terrestres, la óptica adaptativa (AO) se ha convertido en una técnica intrínseca para acercar las observaciones científicas al límite de difracción de los instrumentos astronómicos. Esto se debe a que AO permite la compensación de las aberraciones ópticas causadas por la turbulencia atmosférica, así como las vibraciones de la estructura del telescopio inducidas por elementos dentro de la instrumentación del sistema (como ventiladores y bombas de enfriamiento), el viento y los movimientos del telescopio. Ya que las vibraciones afectan fuertemente el rendimiento de los sistemas AO y dificultan el logro de imágenes de buena calidad, es necesario obtener un modelo de estas vibraciones para posteriormente desarrollar técnicas de control simples, pero efectivas, que se puedan implementar en tiempo real. Es por ello que en esta tesis se propone caracterizar estas vibraciones modelándolas como una combinación lineal de osciladores alimentados cada uno de ellos por un ruido e identificando dichos osciladores de tiempo continuo utilizando un muestreo regular. Se representa el modelo del oscilador como un modelo autoregresivo en tiempo continuo para luego obtener su modelo equivalente en tiempo discreto, en función de los parámetros del

modelo en tiempo continuo. Posteriormente, se identifica el modelo utilizando el método de Máxima Verosimilitud ocupando algoritmos de optimización tanto local como global.

Cuando se utiliza un algoritmo de optimización local, se requiere una buena estimación inicial para los parámetros del sistema para luego realizar la correspondiente optimización, la cual en esta tesis consiste en implementar el algoritmo de cuasi Newton. Por otra parte, cuando se utiliza un algoritmo de optimización global, se analiza el modelo equivalente de datos muestreados para dos casos: i) muestreo instantáneo y ii) muestreo integrado.

Se analiza en detalle ambos tipos de optimización, ilustrando el comportamiento de la función log-verosimilitud a través de ejemplos numéricos que muestran que presenta varios máximos locales.

Palabras Claves – Óptica Adaptativa, Vibraciones, Identificación de Osciladores, modelo CARMA, Método de Máxima Verosimilitud.

ABSTRACT

In all major ground-based astronomical observatories, adaptive optics (AO) has become an intrinsic technique to bring scientific observations closer to the diffraction limit of the astronomical instruments. This is because AO enables the compensation of the optical aberrations caused by atmospheric turbulence, as well as the vibrations of the structure of the telescope induced by elements within the system instrumentation (such as fans and cooling pumps), wind and movements of the telescope. Since vibrations strongly affect the performance of the AO systems and hinder the achievement of good quality images, it is necessary to obtain a model of these vibrations to later develop simple but effective control techniques that can be implemented in real time. It is for this reason that in this thesis it is proposed to characterize these vibrations by modeling them as a linear combination of oscillators each one driven fed by a noise and identifying the continuous-time oscillators using regular sampling. The model of the oscillator is represented as continuous-time autoregressive model, obtaining its discrete-time equivalent model, in terms of the parameters of the model in continuous-time oscillator. Then, the model is identified using the method of Maximum Likelihood using local and global optimization algorithms.

When a local optimization algorithm is used, a good initial estimation is required for the parameters of the system. Then one performs the corresponding optimization, which in this case is implemented using the algorithm of quasi Newton. On the other hand, when a global optimization algorithm is used, the equivalent model of sampled data is analyzed for two cases: i) instantaneous sampling and ii) integrated sampling.

Both types of optimization are analyzed in detail, illustrating the behavior of the log-likelihood function through numerical examples that show that it presents several local maxima.

Keywords – Maximum Likelihood Methods, Continuous Time System Estimation, Oscillators identification, CARMA model, Maximum Likelihood, Adaptive Optics, Vibrations

CONTENTS

ÍNDICE DE FIGURAS	VII
1. INTRODUCCIÓN	2
1.1. Motivación	2
1.2. Principales Contribuciones de la Tesis	4
1.3. Publicaciones por el Autor	5
1.4. Notación	6
1.5. Organización del documento	6
2. ÓPTICA ADAPTATIVA	7
2.1. Introducción	7
2.2. Estado del Arte	9
2.3. Conclusiones	12
3. CONCEPTOS BÁSICOS DE SISTEMAS ESTOCÁSTICOS MUESTRADOS	13
3.1. Introducción	13
3.2. Modelos muestreados	13
3.3. Sistema estocástico	14
3.4. Ejemplos numéricos	18
3.5. Factorización Espectral	20
3.5.1. Modelo al muestrear el sistema oscilatorio instantáneamente	22
3.5.2. Modelo al muestrear el sistema oscilatorio con AFF	23
3.6. Conclusiones	24
4. MODELO Y ESTIMACIÓN DE MÁXIMA VEROSIMILITUD DE OSCILADORES EN TIEMPO-CONTINUO	25
4.1. Introducción	25

4.2. Modelo del sistema	25
4.3. Modelo en espacio de estado equivalente	26
4.4. Modelo equivalente en tiempo discreto	27
4.5. Método de Máxima Verosimilitud	32
4.6. Conclusiones	33
5. OPTIMIZACIÓN DE MÁXIMA VEROSIMILITUD	34
5.1. Introducción	34
5.2. Optimización Local	34
5.3. Optimización Global	42
5.4. Conclusiones	45
6. CONCLUSIONES	46
6.1. Conclusiones	46
6.2. Trabajo Futuro	47
BIBLIOGRAPHY	49
A. APÉNDICE A	55
A.1. Cálculo de las matrices $e^{A\Delta}$ y Q_d del modelo en tiempo-discreto.	55
A.2. Cálculo de derivadas necesarias para obtener el gradiente de la función de máxima verosimilitud.	55
A.3. Procedimiento en Matlab	56
B. APÉNDICE B	58
B.1. Lemas	58
B.2. Algoritmo de Putzer Modificado	61

Índice de figuras

2.1. Imágen del cúmulo globular de la constelación Hércules capturada por el observatorio Gemini Norte: (Abajo) sin el sistema de AO y (Arriba) con sistema de AO.	8
2.2. Diagrama esquemático de un sistema de Óptica Adaptativa.	8
2.3. (a): Instrumentos astronómicos del telescopio de Gemini South. (b): Acelerómetro PCB Piezotronics triaxial (amarillo) instalado en GPI.	9
2.4. Densidad espectral de potencia del acelerómetro triaxial ubicado en el sector a)Ref. Ring (b) Front Frame de GPI.	10
2.5. Diagrama de bloque del lazo cerrado del sistema de AO.	11
5.1. Gráfico de la Función de Verosimilitud de vibraciones generadas con frecuencia en $f_1 = 300[Hz]$ y amplitud $\beta_1 = 1$, utilizando (a) $N = 500$ y (b) $N = 1000$ muestras de la señal de salida y_k	37
5.2. Densidad espectral de potencia para el Ejemplo 1 con $N = 1000$	37
5.3. Estimación de $\theta = [\alpha \ \beta]^T$ para 100 realizaciones de Monte Carlo con (a) $N = 500$ y (b) $N = 1000$ muestras de la señal de salida y_k	38
5.4. Gráfico de la Función de Verosimilitud de vibraciones generadas con frecuencia en $f_1 = 60[Hz]$ y amplitud $\beta_1 = 1$, utilizando (a) $N = 500$ y (b) $N = 1000$ muestras de la señal de salida y_k	39
5.5. Estimación de $\theta = [\alpha \ \beta]^T$ para 100 realizaciones de Monte Carlo con (a) $N = 500$ y (b) $N = 1000$ muestras de la señal de salida y_k	39
5.6. Gráfico de la Función de Verosimilitud de las vibraciones generadas con frecuencias $f_1 = 380[Hz]$ y $f_2 = 400[Hz]$, utilizando (a) $\beta_0 = [1 \ 1]^T$ y $N = 500$, (b) $\beta_0 = [1 \ 1]^T$ y $N = 1000$, (c) $\beta_0 = [3 \ 1]^T$ y $N = 500$, y (d) $\beta_0 = [3 \ 1]^T$ y $N = 1000$	40
5.7. Estimación de $\theta = [\alpha \ \beta]^T$ para 100 realizaciones de Monte Carlo con $N = 1000$ y $\beta_0 = [3 \ 1]^T$	41

5.8. Magnitud del diagrama de bode para el Ejemplo 2 con $N = 1000$ y $\beta_0 = [3 \ 1]^T$	41
5.9. Gráfico de la Función de Verosimilitud de vibraciones generadas con frecuencias en $f_1 = 300$ y $f_2 = 400[Hz]$, amplitud $\beta_1 = 1$ y $\beta_2 = 1$, utilizando (a) $N = 500$ y (b) $N = 1000$ muestras de la señal de salida y_k	42
5.10. Función Log-Verosimilitud para el sistema (4.14) utilizando (a) muestreo instantáneo y (b) muestreo integrado con $\Delta = 1 \text{ ms}$	44
5.11. Función Log-Verosimilitud para el sistema (4.14) utilizando (a) muestreo instantáneo y (b) muestreo integrado con $\Delta = 100 \text{ ms}$	45

INTRODUCCIÓN

1.1. Motivación

La identificación de vibraciones ha sido un tema recurrente en diferentes áreas de investigación, tales como: Óptica Adaptativa (AO) [1–6], Robótica móvil [7], y Procesamiento de Señales y Comunicaciones [8,9].

Los sistemas de AO se emplean en los telescopios modernos para obtener imágenes cercanas al límite de difracción de los instrumentos científicos. Estos sistemas son muy sensibles a las vibraciones inducidas por elementos dentro de la instrumentación del sistema, como ventiladores y bombas de enfriamiento, por lo que caracterizar y mitigar los efectos de las vibraciones es un tema de creciente importancia en muchos observatorios [2, 6, 8]. Es por ello que la motivación principal de esta tesis es modelar y desarrollar técnicas de identificación de las vibraciones inducidas en los modernos sistemas de AO de los telescopios.

En este trabajo de investigación se propone modelar el sistema de interés como una combinación lineal de osciladores alimentados cada uno de ellos por un ruido. Por ende, la propuesta involucra la estimación de los parámetros de cada uno de dichos ruidos. La identificación de sistemas osciladores ha sido un tema de interés recurrente en áreas de investigación, como i) Estadística [9–11] y ii) Física Matemática [12]. Los osciladores de tiempo continuo son muy relevantes para una amplia gama de áreas de investigación ya que están presentes no solo en la naturaleza sino también en algunas aplicaciones tecnológicas (veáse, por ejemplo [3, 12] y sus referencias). Los sistemas osciladores se representan típicamente utilizando ecuaciones diferenciales en tiempo continuo. Además, es bien sabido que cuando se muestrea un sistema dinámico aparece, en general, ceros de muestreo [13, 14]. Por lo tanto, los algoritmos para identificar sistemas de tiempo continuo son altamente relevantes.

Por otro lado, la identificación de sistemas de tiempo continuo se realiza normalmente utilizando modelos aproximados [13, 15, 16]. De hecho, para los sistemas no lineales no existe una representación exacta del modelo equivalente de datos muestreados, y para los

sistemas lineales, el mapeo de los parámetros de tiempo continuo a los ceros de muestreo es, en general, complejo.

El problema de obtener un modelo aproximado utilizando datos muestreados ha atraído la atención de varios grupos de investigación (véase, por ejemplo [17–21]).

Un enfoque típico para identificar un sistema de tiempo continuo es utilizar una aproximación como el método de Euler [13, 16, 22] para obtener un modelo (aproximado) de tiempo discreto. Por ejemplo, en [23] se considera el problema de estimar los parámetros en un modelo autorregresivo en tiempo continuo a partir de muestras de tiempo discreto, reemplazando las derivadas por aproximaciones delta, formando una regresión lineal y utilizando el método de mínimos cuadrados. Además, es bien sabido que el rendimiento de los métodos de identificación obtenidos mediante este enfoque proporciona estimaciones sesgadas [24].

Por otra parte, en [25] se obtiene una evaluación exacta del espectro de potencia en el dominio discreto usando B-splines exponenciales y se presenta un enfoque de estimación que se basa en la función de verosimilitud de Whittle [26].

En el reciente artículo [27] se presenta un algoritmo para identificar un sistema autorregresivo de media móvil de tiempo-continuo (CARMA) que utiliza datos muestreados de forma irregular en el marco de máxima verosimilitud (a través de la maximización directa de la función de verosimilitud y mediante el algoritmo de Esperanza-Maximización (EM) [28]). Los modelos CARMA se modelan como una función de transferencia con polos y ceros alimentada por un ruido [27]. En [27] la función de Máxima Verosimilitud (ML) es optimizada utilizando métodos de Newton [29]. Este enfoque se puede aplicar directamente para identificar osciladores de tiempo continuo. Sin embargo, el éxito de las técnicas en [25, 27] depende de la inicialización de la estimación de los parámetros del sistema.

En este trabajo el procedimiento para conseguir la estimación de los parámetros del modelo de osciladores es el siguiente: (i) Describir la correlación de los distintos ruidos que alimentan al sistema, (ii) expresar el sistema CARMA como un modelo de espacio de estado, (iii) obtener su correspondiente modelo exacto (para estadísticas de segundo orden) equivalente en tiempo discreto considerando el tiempo de muestreo finito Δ , (iv) identificar las vibraciones utilizando la técnica de estimación ML ampliamente utilizada en la literatura [30] y [31]; Para ello se desarrollan dos estrategias de algoritmo de optimización una local y otra global. En el caso de que la estimación de los parámetros del modelo de osciladores se realiza mediante un algoritmo de optimización local, se calcula la función de log-verosimilitud mediante la descomposición del error de predicción a través del filtro de Kalman, para luego identificar los parámetros del modelo localmente utilizando técnicas de cuasi-Newton. Por otra parte, si la estimación se realiza mediante

un algoritmo de optimización global, se calcula la función log-verosimilitud directamente desde el modelo ARMA del sistema de interés. Además, al implementar el algoritmo de optimización global, se analiza la función de verosimilitud para dos casos: i) muestreo instantáneo y ii) muestreo integrador.

El objetivo principal de la presente tesis es obtener una primera aproximación de las dificultades para identificar los parámetros de un oscilador de tiempo continuo. En los ejemplos numéricos, se observa que la función de verosimilitud muestra varios máximos locales, lo que ayuda a comprender las dificultades de aplicar los algoritmos [25, 27] a este problema.

1.2. Principales Contribuciones de la Tesis

Las principales contribuciones de la tesis son las siguientes:

1. Se proponen modelar las vibraciones de sistemas de AO como una combinación lineal de osciladores alimentados cada uno de ellos por un ruido. Por claridad de la exposición el modelo se considera sin condiciones iniciales ni tampoco amortiguamiento.
2. A pesar de que en los sistemas lineales el mapeo de los parámetros de tiempo continuo a los ceros de muestreo es muy complejo, aquí se obtiene el modelo exacto en tiempo discreto (para estadísticas de segundo orden) considerando el tiempo de muestreo finito Δ .
3. Se propone dos métodos de identificación para el modelo de vibraciones utilizando Máxima Verosimilitud: Se maximiza la función de verosimilitud a través de i) método de optimización local y ii) global. En el caso de optimización local se calcula la función de log-verosimilitud utilizando la descomposición del error mediante el filtro de Kalman y se estiman los parámetros utilizando el algoritmo de cuasi-Newton. En cambio, en la optimización global se calcula la función de log-verosimilitud utilizando el modelo ARMA obtenido previamente.
4. Implementación del algoritmo ML para ambos casos de optimización, lo que permite comprender las dificultades para identificar los parámetros de un oscilador de tiempo continuo ya que la función log-verosimilitud presenta varios máximos locales.
5. A través de los ejemplos numéricos presentados se comprueba la validez del análisis de los ceros de muestreo asintóticos ($\Delta \rightarrow 0$) en [13] en el caso de que el sistema estocástico de tiempo continuo es un oscilador con polos en el eje imaginario.

1.3. Publicaciones por el Autor

1. Artículo en Revista

K. González, P. Escárate, R. Carvajal and J.C. Agüero. On Maximum Likelihood Estimation of Continuous Time Oscillators. Submitted in 2017 (november) IEEE Latin America Transactions.- **Abstract:** In this paper, we address the identification of a continuous-time oscillator. We represented the oscillators model as using continuous-time autogressive moving average model. We obtain its equivalent discrete-time model, in terms of the continuous time model parameters. Thus, we identify the model using the Maximum Likelihood method. Our main motivation is modelling the telescope structure vibrations induced by wind and the telescope movements in modern adaptive optics systems.

2. Artículo en Conferencias

- a) J.C. Agüero, K. González and R. Carvajal. EM-based identification of ARX systems having quantized output data. In 2017 20th IFAC World Congress.- **Abstract:** In this paper we develop a novel algorithm to identify an autoregressive with exogenous signal system utilizing quantized output data. We use the Expectation-Maximization algorithm to obtain the Maximum Likelihood estimate.
- b) K. González, M. Coronel, R. Carvajal, P. Escárate and J.C. Agüero. Maximum Likelihood Identification of a Continuous-Time Oscillator Utilizing Sampled Data. Submitted to 18th IFAC Symposium on System Identification, (SYSID 2018).- **Abstract:** In this paper we analyze the likelihood function corresponding to a continuous-time oscillator utilizing regular sampling. We analyze the equivalent sampled-data model for two cases i) instantaneous sampling and ii) integrated sampling. We illustrate the behavior of the log-likelihood function via numerical examples showing that it presents several local maxima.
- c) P. Escárate, K. González, M. Coronel, R. Carvajal and J.C. Agüero. Vibration Model identification using the Maximum Likelihood method. To appear in SPIE Astronomical Telescopes + Instrumentation (2018).- **Abstract:** The vibration effects acting in the science light path reduce the performance of the adaptive optics systems (AO). In order to mitigate the vibration effects and to improve the performance of the AO, an adequate model for the vibration is necessary. A continuous-time oscillator modelled as a continuous-time autoregressive moving average model will be identified in order to estimate the main frequency vibrations peaks. The use of Maximum Likelihood method to

identify the system of interest (considering the effect of initial conditions) will be discussed.

1.4. Notación

La notación que se utilizará en la presente tesis se define a continuación: A^T denota la matriz transpuesta de A , $\det A$ denota el determinante de la matriz A , $\text{adj}(A)$ denota a la matriz adjunta de A , e^A representa a la matriz exponencial de la matriz A . $0_{a \times b}$ y I_a denota la matriz cero de dimensiones $a \times b$ y la matriz identidad de dimensiones $a \times a$ respectivamente. Se utiliza el siguiente vector columna para representar una subsecuencia de la señal x_t , $x_{k:j} = [x_k^T \ x_{k+1}^T \ \dots \ x_j^T]^T$. θ_0 denota el parámetro verdadero, $\hat{\theta}$ denota una estimación de θ_0 . $\hat{\theta}^m$ denota la m -ésima iteración. $\mathbb{E}\{\cdot\}$ denota el operador esperanza. $\mathbb{E}\{x|y\}$ denota el valor esperado de la variable aleatoria x dado la variable aleatoria y . $\delta(\cdot)$ representa a la función Delta de Dirac.

1.5. Organización del documento

Se proporciona a continuación una breve descripción de los temas tratados en cada capítulo de la tesis :

Capítulo 2: En este capítulo se presentan los conceptos fundamentales de la Óptica Adaptativa (AO) para luego plantear el problema teórico a resolver en la presente tesis.

Capítulo 3: En este capítulo se presentan los conceptos básicos respecto a los sistemas estocásticos muestreados uniformemente. Se obtiene el modelo exacto equivalente en tiempo discreto cuando se utiliza muestreo instantáneo e integrado.

Capítulo 4: Se obtiene el modelo exacto en tiempo discreto para el sistema de vibraciones utilizando muestreo instantáneo e integrado. Además, se presenta el método de Máxima Verosimilitud a utilizar para estimar los parámetros de interés.

Capítulo 5: Se desarrollan dos algoritmos de optimización para estimar los parámetros del sistema de interés, uno global y otro local. Se analizan ambos algoritmos a través de varios ejemplos numéricos.

Capítulo 6: En este capítulo se presentan las conclusiones de la tesis y las opciones de trabajo a futuro que permiten darle continuidad a esta investigación.

ÓPTICA ADAPTATIVA

2.1. Introducción

La Óptica adaptativa (AO) es una técnica óptica utilizada para mejorar el rendimiento de los sistemas ópticos al reducir el efecto de las distorsiones del frente de onda entrante, deformando un espejo para compensar la distorsión. La AO es ampliamente utilizada en los modernos telescopios astronómicos [32], en microscopía [33] y en sistemas de imágenes de la retina [34] con el fin de reducir las aberraciones ópticas.

Los sistemas de AO aplicados al área de la astronomía realizan una corrección en tiempo real de los efectos de distorsión producidos por la turbulencia atmosférica, permitiendo que el sistema óptico corregido observe detalles más finos de objetos astronómicos de lo que sería posible observar sin instrumentación astronómica que cuenten con sistema de AO como se puede apreciar en la Fig. 2.1. En dicha figura se comparan dos imágenes del cúmulo globular de la constelación Hércules. La ampliación de arriba fue tomada sin el uso del sistema de AO y la ampliación de abajo, con el sistema de AO.

El sistema de AO requiere una estrella de referencia bastante brillante que esté muy cerca del objeto en estudio. Esta estrella de referencia se utiliza para medir la borrosidad causada por la atmósfera local de modo que el espejo deformable pueda corregirla. Estrellas suficientemente brillantes no están disponibles en todas las partes del cielo, lo que limita en gran medida la utilidad de la AO de estrellas guía naturales. Para resolver esto se pueden utilizar láseres para generar estrellas artificiales a 90[km] de altura.

Los sistemas de AO consta de tres elementos fundamentales como se puede observar en la Fig. 2.2: i) Sensor de frente de onda (WFS), ii) Espejo Deformable (DM) y iii) el sistema de control. En el camino óptico del telescopio se inserta un espejo deformable controlado por un conjunto de actuadores. Posteriormente se ubica el sensor de frente de onda (mide la distorsión inducida por la atmósfera mediante la evaluación de la luz de una estrella natural o artificial) con el objetivo de monitorear la imagen y así cerrar el lazo permitiendo que el frente de onda inicial (distorsionada por la turbulencia atmosférica) sea corregido.

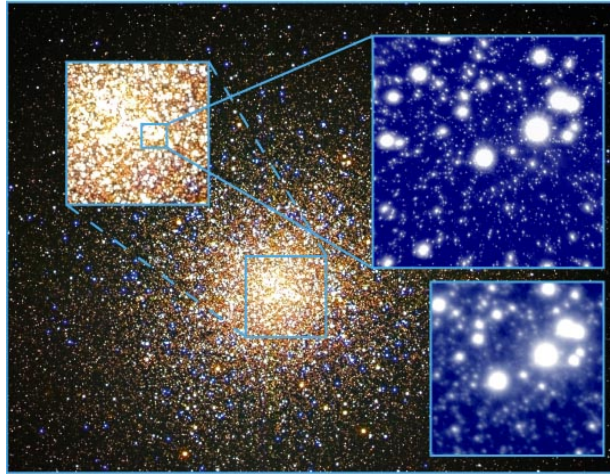


Fig. 2.1: Imagen del cúmulo globular de la constelación Hércules capturada por el observatorio Gemini Norte: (Abajo) sin el sistema de AO y (Arriba) con sistema de AO.

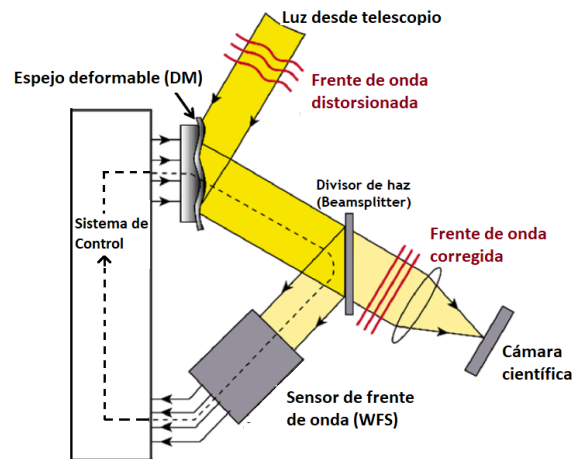


Fig. 2.2: Diagrama esquemático de un sistema de Óptica Adaptativa.

Por otra parte, los telescopios muy o extremadamente grandes que contienen el sistema de AO se ven afectados por picos de vibraciones mecánicas inducidas por el viento, movimientos del telescopio y en mayor medida por elementos dentro de la instrumentación como bombas de enfriamiento y motores.

En la instrumentación astronómica del observatorio Gemini South se han instalados acelerómetros para monitorear las vibraciones mecánicas con el fin de identificar las fuentes de perturbación y posteriormente caracterizarlas. En la Fig. 2.3 se presenta la

estructura de la instrumentación astronómica de Gemini South y de la vista esquemática de un acelerómetro posicionado en el instrumento Gemini Planet Imager (GPI).

En colaboración con el equipo de ingenieros del observatorio Gemini South se realizaron mediciones de aceleración. En la Fig. 2.4 se gráfica la densidad espectral de potencia (PSD) de las mediciones capturadas en dos sectores del instrumento GPI, donde se han marcado los peaks de magnitudes más altas. En ambos gráficos se presentan peaks de 60 [Hz] y sus armónicas (120, 180 y 240 [Hz]), provenientes de dos bombas de enfriamiento de GPI que funcionan a 60 [Hz]. También, están presente peaks de 100[Hz] y sus armónicas debido a las bombas de enfriamiento de Flamingos-2 (instrumento instalado cerca de GPI).

A través de la descomposición Polinomial de Zernike es posible caracterizar los diferentes tipos de aberraciones ópticas [35,36]. Existen aberraciones de bajo orden como lo son: tip, tilt, defocus y astigmatismo. Mientras que de alto orden existen varios tipos de aberraciones como por ejemplo: esférica y coma. Las perturbaciones mecánicas debidas a la vibración conducen principalmente a aberraciones de pistón, tip y tilt, afectando directamente a la nitidez de las imágenes capturadas. Es por ello la importancia de obtener un modelo de los datos muestreados y así posteriormente mitigarlas utilizando técnicas de control [3,4,6,37,38].

2.2. Estado del Arte

Para el modelado de los sistemas de AO se utiliza un esquema de lazo AO estándar [4,6,38], teniendo en cuenta la presencia de perturbaciones debidas a componentes de turbulencia y vibración. Es fundamental obtener una correcta identificación de todas las fuentes de ruido y perturbación, así como los parámetros que definen la dinámica del

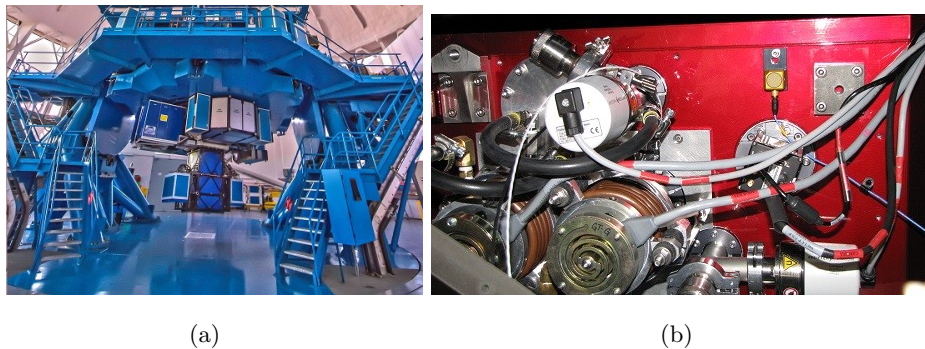
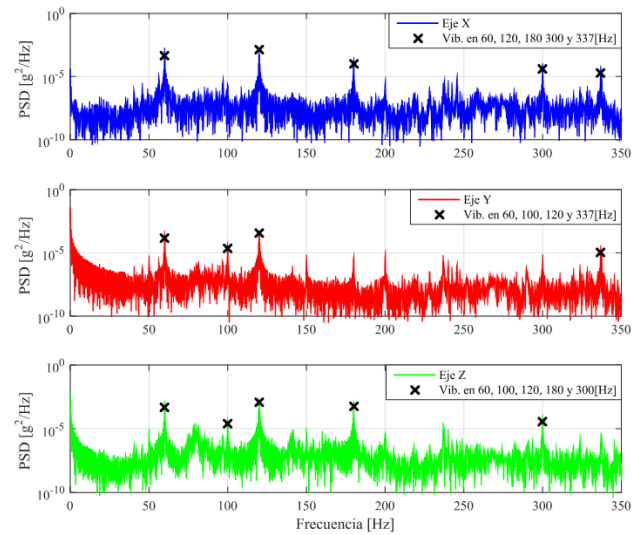
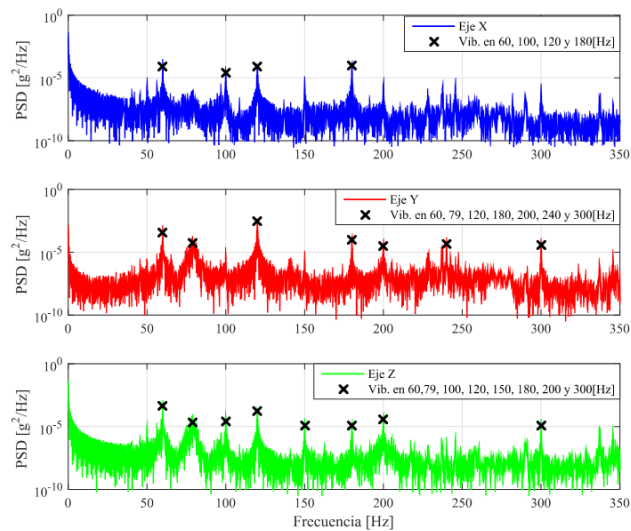


Fig. 2.3: (a): Instrumentos astronómicos del telescopio de Gemini South. (b): Acelerómetro PCB Piezotronics triaxial (amarillo) instalado en GPI.

espejo deformable y el sensor de frente de onda, lo que permite una correcta mitigación de las vibraciones y perturbaciones.



(a)



(b)

Fig. 2.4: Densidad espectral de potencia del acelerómetro triaxial ubicado en el sector a)Ref. Ring (b) Front Frame de GPI.

La estructura estándar para un sistema de óptica adaptativa en lazo cerrado consiste en un sensor de frente de onda (WFS), un espejo deformable (DM) y un controlador (K) [4, 6, 38], el cual se muestra en la Fig. 2.5. En dicha figura, φ , φ^{cor} y φ^{res} representan las amplitudes para la fase de perturbación total, fase de corrección y fase residual, respectivamente.

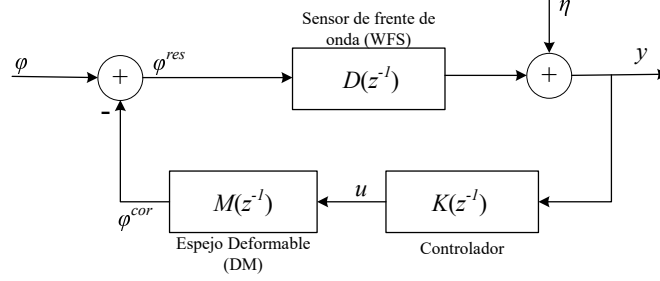


Fig. 2.5: Diagrama de bloque del lazo cerrado del sistema de AO.

Observación 1. Asumiendo que la fase incidente de perturbación total φ puede ser representada usando la descomposición Polinomial de Zernike [7, 8, 39], se puede diseñar una regla de control diferente para cada modo de frente de onda independiente, siguiendo la estructura presentada en la Fig. 2.5. Desde este punto de vista, y sin pérdida de generalidad, se ha considerado solamente los modos tip, tilt y defocus de la descomposición Polinomial de Zernike del frente de onda.

La perturbación total φ se define como la suma de las contribuciones de la señal de turbulencia φ^{tur} y las n señales de vibración. Es decir:

$$\varphi_t = \varphi_t^{tur} + \varphi_t^{vib1} + \varphi_t^{vib2} + \dots + \varphi_t^{vibn}. \quad (2.1)$$

El proceso de medición del WFS y el proceso de corrección del DM se pueden modelar como [3, 4, 40]:

$$\begin{aligned} D(z^{-1}) &= d_0 z^{-1} \\ M(z^{-1}) &= m_0 z^{-1} \end{aligned} \quad (2.2)$$

Así, la señal de salida y está dada por:

$$y_t = d_0 \varphi_{t-1}^{res} + \eta_t, \quad (2.3)$$

donde el ruido $\eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$ es Gaussiano blanco aditivo y

$$\varphi_t^{cor} = m_0 u_{t-1}. \quad (2.4)$$

Un método ampliamente utilizado en la literatura para describir la correlación temporal de las fases de turbulencia y vibración son modelos autoregresivos de segundo orden (AR2) [7, 8, 39].

2.3. Conclusiones

En este capítulo se explicó de manera general la Óptica Adaptativa y el funcionamiento de los sistemas de AO aplicados a la astronomía. Se dio a conocer que los sistemas de AO son sensibles a las vibraciones mecánicas inducidas por elementos al interior de la instrumentación como las bombas de enfriamiento en el caso del instrumento GPI. Debido a ello, es fundamental conseguir un modelo que represente a estas vibraciones con el fin de mitigarlas mediante técnicas de control.

CONCEPTOS BÁSICOS DE SISTEMAS ESTOCÁSTICOS MUESTREADOS

3.1. Introducción

En este capítulo se estudia el muestreo de un sistema estocástico lineal en tiempo continuo. En primer lugar, se comenta de manera general los modelos muestreados y los dispositivos utilizados en el proceso de muestreo. Luego se presentan las ecuaciones generalizadas en tiempo discreto de sistemas estocásticos lineales muestreados instantáneamente y muestreados utilizando previamente un filtro anti-aliasing. Se presenta un ejemplo numérico para cada tipo de muestreo. Posteriormente, se obtienen los modelos de datos muestreados de sistemas oscilatorios indirectamente utilizando factorización espectral. Finalizando con la presentación de las conclusiones.

3.2. Modelos muestreados

Los modelos de sistemas dinámicos en tiempo continuo surgen en general de la aplicación de leyes físicas tales como la conservación de la masa, el momento y la energía. Estos modelos se pueden representar como ecuaciones diferenciales lineales o no lineales, donde los parámetros involucrados generalmente se pueden interpretar en términos de propiedades físicas del sistema. Sin embargo, en la práctica este tipo de modelos no son apropiados para interactuar con dispositivos digitales (por ejemplo actuadores y sensores), debido a que solo se puede acceder a los datos de la señal en instantes de tiempo discreto.

Al interconectar un sistema físico de tiempo continuo a una computadora digital, se necesitan dispositivos de interfaz en la entrada y salida, los cuales determinan las propiedades del sistema muestreado:

- Retenedor: el retenedor se utiliza para convertir la secuencia de tiempo discreto u_k de la computadora en una señal de entrada de tiempo continuo adecuada para la aplicación al sistema. El retenedor de orden cero (ZOH)

$$u(t) = u_k \text{ para } k\Delta \leq t < (k+1)\Delta$$

es la más comúnmente utilizada, donde Δ corresponde al tiempo de muestreo.

- Sistema físico: usualmente se describe mediante un conjunto de ecuaciones diferenciales lineales o no lineales en tiempo continuo.
- Filtro Anti-aliasing (AAF): este dispositivo acondiciona la señal de salida de tiempo continuo antes de tomar las muestras.
- Muestreador: este dispositivo crea una secuencia de tiempo discreto y_k mediante muestreo instantáneo.

3.3. Sistema estocástico

En esta sección se considera el muestreo de sistemas lineales estocásticos descritos como:

$$y(t) = H(q)\dot{w}(t) \quad (3.1)$$

donde $\dot{w}(t)$ es un proceso de entrada de ruido blanco de tiempo continuo. Se asume que la señal que alimenta al sistema $\dot{w}(t)$ es débilmente estacionaria. Por ende, al aplicar cualquier filtro estable la salida será asintóticamente débilmente estacionaria y si además el ruido \dot{w} tiene distribución Gaussiana entonces la señal de salida será estrictamente estacionaria [41]. Para el sistema (3.1) se puede obtener un modelo de datos muestreados de manera que su secuencia de salida tenga las mismas propiedades de segundo orden que la salida de tiempo continuo en los instantes de muestreo. Un ruido de tiempo continuo, $w(t)$, es un proceso de Wiener si cumple con [14, 16]

- $w(t)$ tiene media cero, es decir, $\mathbb{E}\{w(t)\} = 0$ para todo t .
- Los incrementos de $w(t)$ son independientes, es decir, $\mathbb{E}\{(w(t_1) - w(t_2))(w(s_1) - w(s_2))\} = 0$ para todo $t_1 > t_2 > s_1 > s_2 \geq 0$.
- Para cada s y t , $s \leq t$, los incrementos $w(t) - w(s)$, tiene una distribución Gaussiana de media cero y varianza $\mathbb{E}\{(w(t) - w(s))^2\} = \sigma_w^2 |t - s|$.

Por lo tanto, la derivada de un proceso de Wiener, $\dot{w}(t)$ tiene distribución Gaussiana.

Lema 1. Dado el siguiente sistema estocástico en tiempo continuo expresado en espacio-estado:

$$\dot{x}(t) = Ax(t) + \kappa\dot{w}(t) \quad (3.2)$$

donde $\dot{w}(t)$ es ruido blanco Gaussiano escalar en tiempo continuo con media cero y varianza $\mathbb{E}\{\dot{w}(t)\dot{w}(r)\} = \sigma^2\delta(t-r)$. El sistema equivalente en tiempo discreto equivalente está dado por:

$$x(t_{k+1}) = A_d(t_k) + n(t_{k+1}), \quad (3.3)$$

donde $\Delta = t_{k+1} - t_k$ corresponde al tiempo de muestreo, $A_d = e^{A\Delta}$, $n(t_{k+1}) = \int_0^\Delta e^{A\eta}\kappa\dot{w}(t_{k+1} - \eta)d\eta$ es ruido blanco de media cero y varianza Q_d .

Demostración. Considerando la matriz exponencial $e^A = I + A + \frac{1}{2!}A^2 + \dots + \frac{1}{n!}A^n$ con $n \in \mathbb{R}$. La matriz exponencial satisface lo siguiente:

- $\frac{de^{At}}{dt} = Ae^{At} = e^{At}A$
- $\lim_{\Delta \rightarrow 0} \frac{e^{A\Delta} - I}{\Delta} = A$

Multiplicando la ecuación (3.2) por el término e^{-At} se obtiene que:

$$\begin{aligned} e^{-At}\dot{x}(t) &= e^{-At}Ax(t) + e^{-At}\kappa\dot{w}(t) \\ \frac{d}{dt}\{e^{-At}x(t)\} &= e^{-At}\kappa\dot{w}(t) \end{aligned}$$

Integrando la última expresión con respecto a t entre los instantes t_k y t_{k+1} se obtiene:

$$e^{-At_{k+1}}x(t_{k+1}) - e^{-At_k}x(t_k) = \int_{t_k}^{t_{k+1}} e^{-At}\kappa\dot{w}(t)dt. \quad (3.4)$$

Despejando $x(t_{k+1})$, se obtiene el sistema (3.2) equivalente en tiempo-discreto:

$$x(t_{k+1}) = e^{A\Delta}x(t_k) + n(t_k), \quad (3.5)$$

donde

$$n(t_k) = \int_0^\Delta e^{A\eta}\kappa\dot{w}(t_{k+1} - \eta)d\eta, \quad (3.6)$$

es ruido blanco de media cero y varianza

$$\begin{aligned} \mathbb{E}\{n(t_k)n(t_k)^T\} &= \int_0^\Delta \int_0^\Delta e^{A\eta}\kappa\mathbb{E}\{\dot{w}(t_{k+1} - \eta)\dot{w}(t_{k+1} - \tau)\}\kappa^T e^{A^T\tau}d\eta d\tau \\ &= \int_0^\Delta e^{A\eta}\kappa\kappa e^{A^T\eta}d\eta = Q_d. \end{aligned} \quad (3.7)$$

Para el cálculo anterior se considera la siguiente igualdad

$$\int_0^\Delta \mathcal{F}(\eta, \tau) \delta(\eta - \tau) d\tau = \mathcal{F}(\eta, \eta)$$

En cuanto a las matrices $e^{A\Delta}$ y Q_d , éstas se pueden calcular eficientemente a partir de una matriz exponencial extendida (para más detalles ver Apéndice 4). \square

Si la salida del sistema de tiempo continuo (3.2) satisface

$$y(t) = Cx(t), \quad (3.8)$$

entonces, para un **muestreo instantáneo** se obtiene que la salida en tiempo discreto equivalente está dado por:

$$y(t_k) = Cx(t_k). \quad (3.9)$$

Equivalentemente, la representación de entrada-salida está dada por:

$$y(t_k) = C(zI - A_d)^{-1}n(t_{k+1}) \quad (3.10)$$

donde z es el operador de adelanto. Por otra parte, es habitual en la práctica utilizar un AAF antes de tomar muestras. Un AAF típicamente utilizado para análisis es el siguiente **promediador AAF** [13]:

$$F(s) = \frac{1 - e^{s\Delta}}{s\Delta} \quad (3.11)$$

Este tipo de muestreo también se llama **muestreo integrado** ya que las muestras están dadas por:

$$y(t_{k+1}) = \frac{1}{\Delta} \int_{t_k}^{t_{k+1}} \dot{z}(t) dt = \frac{z(t_{k+1}) - z(t_k)}{\Delta}. \quad (3.12)$$

Lema 2. Se considera el siguiente sistema en tiempo continuo

$$\begin{aligned} \dot{x}(t) &= Ax(t) + \kappa \dot{w}(t) \\ \dot{z}(t) &= Cx(t) + \dot{v}(t), \end{aligned} \quad (3.13)$$

donde $[\kappa \dot{w}(t) \ \dot{v}(t)]^T$ es ruido blanco de tiempo continuo con covarianza dada por:

$$\mathbb{E} \left\{ \begin{bmatrix} \kappa \dot{w}(t) \\ \dot{v}(t) \end{bmatrix} \begin{bmatrix} \kappa \dot{w}(\tau) \\ \dot{v}(\tau) \end{bmatrix}^T \right\} = \begin{bmatrix} Q_c & S_c \\ S_c & R_c \end{bmatrix} \delta(t - \tau). \quad (3.14)$$

Si las muestras son tomadas con un promediador AAF, entonces el correspondiente sistema en tiempo discreto está dado por:

$$\begin{bmatrix} x(t_{k+1}) \\ y(t_{k+1})\Delta \end{bmatrix} = \begin{bmatrix} A_d \\ C_d \end{bmatrix} x(t_k) + n(t_{k+1}), \quad (3.15)$$

donde $A_d = e^{A\Delta}$, $C = CA^{-1}(e^{A\Delta} - I)$ en caso de que la matriz A es invertible. Además, $n(t_{k+1})$ es ruido blanco de media cero con matriz de covarianza dada por:

$$\mathbb{E} \{n(t_{k+1})n(t_{k+1})^T\} = \int_0^\Delta e^{A_x\eta} \begin{bmatrix} Q_c & S_c \\ S_c & R_c \end{bmatrix} e^{A_x^T\eta} d\eta = \begin{bmatrix} Q_d & S_d \\ S_d & R_d \end{bmatrix} \quad (3.16)$$

donde la matriz extendida A_x está dada por

$$A_x = \begin{bmatrix} A & 0 \\ C & 0 \end{bmatrix} \quad (3.17)$$

Demostración. Definiendo un sistema con un estado extendido $\mathcal{X}(t) = [x(t) \ z(t)]^T$, y un ruido extendido $\dot{n}(t) = [\kappa\dot{w}(t) \ \dot{v}(t)]^T$ se tiene que:

$$\dot{\mathcal{X}}(t) = \begin{bmatrix} A & 0 \\ C & 0 \end{bmatrix} \mathcal{X}(t) + \dot{n}(t)$$

Por lo tanto, el sistema de tiempo discreto al utilizar el Lema 1 está dado por:

$$\mathcal{X}(t_{k+1}) = e^{\begin{bmatrix} A & 0 \\ C & 0 \end{bmatrix}\Delta} \mathcal{X}(t_k) + n(t_{k+1}), \quad (3.18)$$

donde

$$n(t_{k+1}) = \int_0^\Delta e^{\begin{bmatrix} A & 0 \\ C & 0 \end{bmatrix}\eta} \dot{n}(t_{k+1} - \eta) d\eta. \quad (3.19)$$

Usando [42] y considerando que A es invertible, se tiene que:

$$e^{\begin{bmatrix} A & 0 \\ C & 0 \end{bmatrix}\Delta} = \begin{bmatrix} e^{A\Delta} & 0 \\ CA^{-1}(e^{A\Delta} - I) & I \end{bmatrix}. \quad (3.20)$$

El calculo de la matriz exponencial $e^{A_x\Delta}$ se obtiene utilizando el algoritmo de Putzer modificado (ver Apéndice B.1). Expandiendo la ecuación (3.18)

$$\begin{bmatrix} x(t_{k+1}) \\ z(t_{k+1}) \end{bmatrix} = \begin{bmatrix} e^{A\Delta} & 0 \\ CA^{-1}(e^{A\Delta} - I) & I \end{bmatrix} \begin{bmatrix} x(t_k) \\ z(t_k) \end{bmatrix} + n(t_{k+1}),$$

se desprenden las siguiente ecuaciones:

$$\begin{aligned} x(t_{k+1}) &= A_d x(t_k) + w_d(t_{k+1}) \\ y(t_{k+1})\Delta &= z(t_{k+1}) - z(t_k) = C_d x(t_k) + v_d(t_{k+1}) \end{aligned}$$

donde $A_d = e^{A\Delta}$, $C_d = CA^{-1}\{e^{A\Delta} - I\}$ y $n(t_{k+1}) = [w_d(t_{k+1}) \ v_d(t_{k+1})]^T$ es ruido blanco Gaussiano de media cero y con matriz de covarianza dada por:

$$\begin{aligned} \mathbb{E} \{n(t_{k+1})n(t_{k+1})^T\} &= \mathbb{E} \left\{ \int_0^\Delta e^{\begin{bmatrix} A & 0 \\ C & 0 \end{bmatrix} \eta} \dot{n}(t_{k+1} - \eta) d\eta \int_0^\Delta \dot{n}(t_{k+1} - \tau)^T e^{\begin{bmatrix} A & 0 \\ C & 0 \end{bmatrix}^T \tau} \right\} \\ &= \int_0^\Delta e^{A_x \eta} \begin{bmatrix} Q_c & S_c \\ S_c & R_c \end{bmatrix} e^{A_x^T \eta} d\eta = \begin{bmatrix} Q_d & S_d \\ S_d & R_d \end{bmatrix}. \end{aligned}$$

Por lo tanto, el sistema en tiempo discreto queda expresado como:

$$\begin{bmatrix} x(t_{k+1}) \\ y(t_{k+1})\Delta \end{bmatrix} = \begin{bmatrix} A_d \\ C_d \end{bmatrix} x(t_k) + n(t_{k+1}),$$

□

desde donde se puede obtener la representación entrada-salida como:

$$y(t_{k+1})\Delta = C_d(zI - A_d)^{-1}w_d(t_{k+1}) + v_d(t_{k+1}) \quad (3.21)$$

3.4. Ejemplos numéricos

Ejemplo 1. Se considera el siguiente sistema CAR

$$y(t) = \frac{1}{s + \alpha} \dot{w}(t), \quad (3.22)$$

donde $s = \frac{d}{dt}$ corresponde al operador de Heavside y $\mathbb{E}\{\dot{w}(t)\dot{w}(\tau)\} = \sigma^2\delta(t - \tau)$. Este sistema puede ser representado en su forma espacio de estado como:

$$\begin{aligned} \dot{x}(t) &= -\alpha x(t) + \dot{w}(t) \\ y(t) &= x(t) \end{aligned} \quad (3.23)$$

Utilizando el Lema 1 se obtiene que el sistema en tiempo discreto equivalente al muestrearlo instantáneamente está dado por:

$$\begin{aligned} x(t_{k+1}) &= A_d x(t_k) + n(t_{k+1}) \\ y(t_k) &= C x(t_k), \end{aligned}$$

donde $A_d = e^{-\alpha\Delta}$, $C = 1$ y $n(t_{k+1}) = \int_0^\Delta e^{-\alpha\eta} \dot{w}(t_{k+1} - \eta) d\eta$ es ruido blanco de media cero y varianza dada por

$$Q_d = \mathbb{E}\{n(t_{k+1})^2\} = \sigma^2 \int_0^\Delta e^{-2\alpha\eta} d\eta = \sigma^2 \frac{1 - e^{-2\alpha\Delta}}{2\alpha} \quad (3.24)$$

Dada la representación entrada-salida (3.10), el sistema en tiempo discreto puede ser re-escrito como el siguiente sistema auto-regresivo de primer orden (AR(1)) en tiempo discreto:

$$y(t_k) = \frac{n(t_k)}{1 - e^{-\alpha\Delta}q^{-1}} = \frac{n(t_k)}{1 - aq^{-1}} \quad (3.25)$$

Ejemplo 2. Se considera el siguiente sistema

$$\dot{z}(t) = \frac{s}{s + \alpha} \dot{w}(t) \quad (3.26)$$

el cual se muestrea utilizando el promediador AFF y cuyas muestras están dadas por

$$y(t_{k+1}) = \frac{1}{\Delta} \int_{t_k}^{t_{k+1}} \dot{z}(t) dt.$$

Este sistema puede ser representado en su forma espacio de estado como:

$$\begin{bmatrix} \dot{x}(t) \\ \dot{z}(t) \end{bmatrix} = \begin{bmatrix} -\alpha \\ 1 \end{bmatrix} x(t) + \dot{n}(t), \quad (3.27)$$

donde $\dot{n}(t) = [-\alpha\dot{w}(t) \ \dot{z}(t)]$ es ruido blanco de tiempo continuo con covarianza dada por

$$\mathbb{E} \left\{ \begin{bmatrix} -\alpha\dot{w}(t) \\ \dot{w}(t) \end{bmatrix} \begin{bmatrix} -\alpha\dot{w}(\tau) \\ \dot{w}(\tau) \end{bmatrix}^T \right\} = \begin{bmatrix} Q_c & S_c \\ S_c & R_c \end{bmatrix} \delta(t - \tau) = \sigma^2 \begin{bmatrix} \alpha^2 & -\alpha \\ -\alpha & 1 \end{bmatrix} \delta(t - \tau).$$

Utilizando el Lema 2, se tiene que el sistema puede ser representado en la siguiente forma espacio de estado:

$$\begin{bmatrix} x(t_{k+1}) \\ y(t_{k+1})\Delta \end{bmatrix} = \begin{bmatrix} A_d \\ C_d \end{bmatrix} x(t_k) + n(t_{k+1}) \quad (3.28)$$

donde $A_d = e^{-\alpha\Delta}$, $C_d = \alpha^{-1}(1 - e^{-\alpha\Delta})$ y $n(t_{k+1})$ es ruido blanco de media cero con matriz de covarianza dada por:

$$\mathbb{E} \{ n(t_{k+1})n(t_{k+1})^T \} = \int_0^\Delta e^{A_x\eta} \begin{bmatrix} Q_c & S_c \\ S_c & R_c \end{bmatrix} e^{A_x^T\eta} d\eta = \begin{bmatrix} Q_d & S_d \\ S_d & R_d \end{bmatrix} \quad (3.29)$$

con

$$A_x = \begin{bmatrix} -\alpha & 0 \\ 1 & 0 \end{bmatrix} \quad (3.30)$$

Al calcular la matriz exponencial de A_x utilizando [42] se obtiene lo siguiente:

$$e^{A_x\Delta} = \begin{bmatrix} e^{-\alpha\Delta} & 0 \\ \int_0^\Delta e^{-\alpha\tau} d\tau & 1 \end{bmatrix} = \begin{bmatrix} e^{-\alpha\Delta} & 0 \\ \frac{1 - e^{-\alpha\Delta}}{\alpha} & 1 \end{bmatrix} \quad (3.31)$$

Por lo tanto, al calcular la matriz de covarianza de $n(t_{k+1})$ se obtiene lo siguiente:

$$\begin{aligned}
\mathbb{E} \{n(t_{k+1})n(t_{k+1})^T\} &= \sigma^2 \int_0^\Delta e^{A_x \eta} \begin{bmatrix} \alpha^2 & -\alpha \\ -\alpha & 1 \end{bmatrix} e^{A_x^T \eta} d\eta \\
&= \sigma^2 \int_0^\Delta \begin{bmatrix} \alpha^2 e^{-2\alpha\eta} & -\alpha e^{-2\alpha\eta} \\ -\alpha e^{-2\alpha\eta} & e^{-2\alpha\eta} \end{bmatrix} d\eta \\
&= \sigma^2 \frac{1 - e^{-2\alpha\Delta}}{2\alpha} \begin{bmatrix} -\alpha \\ 1 \end{bmatrix} \begin{bmatrix} -\alpha \\ 1 \end{bmatrix}^T
\end{aligned} \tag{3.32}$$

Finalmente, se re-escibe la ecuación 3.28 como:

$$\begin{bmatrix} x(t_{k+1}) \\ y(t_{k+1})\Delta \end{bmatrix} = \begin{bmatrix} A_d \\ C_d \end{bmatrix} x(t_k) + \begin{bmatrix} -\alpha \\ 1 \end{bmatrix} e(t_{k+1}), \tag{3.33}$$

donde $e(t_{k+1})$ es ruido blanco de media cero y varianza dada por:

$$\mathbb{E}\{e(t_{k+1})^2\} = \sigma^2 \frac{1 - e^{-2\alpha\Delta}}{2\alpha} \tag{3.34}$$

Dada la representación entrada-salida (3.21) obtenida a partir del Lema 2, el sistema en tiempo discreto puede ser re-escrito como el siguiente sistema ARMA en tiempo discreto:

$$y(t_k)\Delta = \frac{1 - q^{-1}}{1 - e^{-\alpha\Delta}q^{-1}} e(t_k) \Leftrightarrow \frac{\Delta}{1 - q^{-1}} y(t_k) = \frac{1}{1 - e^{-\alpha\Delta}q^{-1}} e(t_k) \tag{3.35}$$

3.5. Factorización Espectral

Es bien sabido que a partir de la estimación del espectro de entrada-salida se puede obtener directamente el modelo de datos muestreados que representa el sistema de interés en tiempo discreto. Para obtener los modelos correspondiente al sistema estocástico de interés muestreado instantáneamente (3.10) e integrado (3.21) se debe introducir previamente lo siguiente:

Teorema 2. Considerando el siguiente sistema lineal:

$$y(t_k) = H(q)w(t_k), \tag{3.36}$$

donde $w(t_k)$ es ruido blanco de media cero y covarianza Σ , $H(q)$ es un filtro lineal y q es el operador de adelanto ($qw(t_k) = w(t_{k+1})$), el espectro de la señal $y(t_k)$ está dada por [30]:

$$\Phi_y(z) = H(z)\Sigma H(z^{-1})^T \tag{3.37}$$

donde el espectro del ruido blanco es igual a su matriz de covarianza.

□

En este trabajo el sistema de interés es oscilatorio, por ende la salida y del sistema no es estacionario. Por lo tanto y no tiene espectro, porque el espectro en sistemas estocásticos de tiempo discreto se define como

$$\Phi_y(z) = \sum_{\tau=-\infty}^{\infty} r(\tau)z^{-\tau} \quad (3.38)$$

donde $r_y(\tau)$ es la auto-correlación de la señal $y(t_k)$ la cual puede ser estrictamente estacionaria o débilmente estacionaria

$$r_y(\tau) = \mathbb{E} \{y(t + \tau)y(t)\} \quad (3.39)$$

Siendo así, la densidad espectral de potencia es el espectro evaluado en el círculo unitario, es decir $\Phi_y(e^{jw})$. Por lo tanto, la densidad espectral de potencia se relaciona a la auto-correlación $r_y(\tau)$ a través de la transformada (y su inversa) de Fourier de tiempo discreto. Debido a que el sistema es oscilatorio (cuyo transiente no desaparece), entonces el sistema no es estacionario, por ende no se puede obtener el espectro de la salida directamente. En las siguientes subsecciones se presenta una manera indirecta de obtener el modelo de un sistema oscilatorio a partir de la factorización espectral, pero antes se presenta el siguiente Lema que será de utilidad.

Lema 3. Sea $R_1(z)$ y $R_2(z)$ los polinomios $R_1(z) = z^2 + bz + 1$, $R_2(z) = z^4 + bz^3 + cz^2 + bz + 1$ donde b y $c \in \mathbb{R}$. Estos polinomios se pueden factorizar de la siguiente manera:

$$R_1(z) = (z - z_1)(z^{-1} - z_1) \left\{ -\frac{z}{z_1} \right\} \quad (3.40)$$

$$R_2(z) = (z - z_2)(z - z_3)(z^{-1} - z_2)(z^{-1} - z_3) \left\{ \frac{z^2}{z_2 z_3} \right\} \quad (3.41)$$

donde $|z_i| < 1$ con $i = \{1, 2, 3\}$.

Demostración. Se puede ver directamente que si z_* es una raíz de $R_i(z)$ en (3.40) y (3.41), entonces $\frac{1}{z_*}$ también es una raíz de $R_i(z)$ ($i = \{1, 2\}$). Entonces, el resultado se obtiene re-ordenando los términos. □

Note que en el caso que z_1 y z_2 sean complejos conjugados ($z_2 = \bar{z}_1$), entonces se tendrá que el polinomio R_2 se factoriza de la siguiente manera:

$$R_2(z) = (z - z_1)(z - \bar{z}_1)(z^{-1} - z_1)(z^{-1} - \bar{z}_1) \left\{ \frac{z^2}{|z_1|^2} \right\} \quad (3.42)$$

La manera de obtener el modelo de datos muestreados indirectamente es descomponiendo la salida $y(t_k)$ de tal manera de calcular el espectro a la parte de la señal que sea

estacionaria. En primer lugar, se re-escriben las ecuaciones de entrada-salida (3.10) y (3.21) muestreadas de manera instantánea y con AFF respectivamente. Para ello se utiliza la igualdad de la inversa de la matriz A la cual está dada por [43]:

$$A^{-1} = \frac{\text{adj}(A)}{\det(A)} \quad (3.43)$$

Dado las ecuaciones (3.10) y (3.21) (que corresponden a las expresiones de entrada-salida en tiempo discreto de un sistema lineal estocástico cuando se toman muestras instantáneamente y cuando se usa un prefiltro integrante, respectivamente.) se obtiene a continuación los modelos que representa el sistema cuando es oscilatorio:

3.5.1. Modelo al muestrear el sistema oscilatorio instantáneamente

Utilizando (3.43) la representación de entrada-salida (3.10) al muestrear el sistema oscilatorio instantáneamente se puede re-escribir como:

$$\begin{aligned} y(t_k) &= \frac{C \text{adj}(zI - A_d)n(t_{k+1})}{\det(zI - A_d)} \\ \det(zI - A_d)y(t_k) &= \underbrace{C \text{adj}(zI - A_d)n(t_{k+1})}_{\xi(t_k)} \end{aligned} \quad (3.44)$$

donde $\xi(t_k)$ es asintóticamente débilmente estacionario. Por ende, se puede calcular su espectro y en consecuencia el modelo de datos muestreados correspondiente. El término $\det(zI - A_d) = z^2 - 2 \cos[\alpha\Delta]z + 1$ corresponde a la parte oscilatoria. Definiendo $P(z) = C \text{adj}(zI - A_d)$ y utilizando el Teorema 2 se obtiene que el espectro de potencia de $\xi(t_k)$ está dado por:

$$\Phi_\xi(z) = P(z)Q_dP(z^{-1})^T \quad (3.45)$$

donde Q_d es la varianza del ruido $n(t_{k+1})$. Es posible demostrar que el espectro de potencia de $\xi(t_k)$ está dada por [44]:

$$\Phi_\xi(z) = \frac{\eta a R(z)}{z} \quad (3.46)$$

donde $\eta \in \mathbb{R}^+$, $R(z) = z^2 + \frac{b}{a}z + 1$. Usando el Lema 3 se obtiene que:

$$\Phi_\xi(z) = \frac{\eta a}{(-z_1)} (z - z_1) (z^{-1} - z_1) \quad (3.47)$$

donde z_1 es la raíz de $R(z)$ ubicada dentro del círculo unitario. Note que $z_1 < 0$ para $\alpha\Delta$ suficientemente pequeño. Por lo tanto, el modelo ARMA de tiempo discreto para muestreo instantáneo se puede escribir como:

$$y(t_k) = H(z)e(t_k), \quad (3.48)$$

donde

$$H(z) = \frac{z(z - z_1)}{(z^2 - 2z \cos[\alpha\Delta] + 1)} \quad (3.49)$$

y $e(t_k)$ es ruido blanco Gaussiano de media cero con varianza

$$\text{var}\{e(t_k)\} = \frac{\eta a}{|z_1|}. \quad (3.50)$$

3.5.2. Modelo al muestrear el sistema oscilatorio con AFF

Al igual que en el muestreo instantáneo, se re-escribe la representación de entrada-salida (3.21) utilizando (3.43) como:

$$\begin{aligned} y(t_k + 1)\Delta &= \frac{C_d \text{adj}(zI - e^{A\Delta})}{\det(zI - A_d)} \omega_d(t_{k+1}) + v_d(t_{k+1}) \\ \det(zI - A_d)y(t_{k+1}) &= \underbrace{C_d \text{adj}(zI - e^{A\Delta}) \omega_d(t_{k+1}) + \det(zI - A_d)v_d(t_{k+1})}_{\xi(t_{k+1})}, \end{aligned} \quad (3.51)$$

donde $\det(zI - A_d) = z^2 - 2\cos[\alpha\Delta]z + 1$ corresponde a la parte oscilatoria y $\xi(t_{k+1})$ sí tiene espectro. Es posible demostrar que el espectro de potencia de $\xi(t_{k+1})$ está dada por [44]:

$$\Phi_\xi = \eta R(z), \quad (3.52)$$

donde $\eta \in \mathbb{R}^+$ y el polinomio $R(z)$ tiene la siguiente estructura

$$R(z) = a \left\{ z^4 + \frac{b}{a}z^3 + \frac{c}{a}z^2 + \frac{b}{a}z + 1 \right\}. \quad (3.53)$$

Usando el Lema 3, se obtiene que $R(z) = a(z - z_1)(z - z_2)(z - z_1^{-1})(z - z_2^{-1})$ y

$$\Phi_\xi(z) = \frac{\eta a}{(-z_1)(-z_2)} N(z)N(z^{-1}) \quad (3.54)$$

donde $N(z) = (z - z_1)(z - z_2)$. Entonces, el modelo ARMA en tiempo discreto equivalente está dado por:

$$y(t_{k+1}) = H(z)e(t_k), \quad (3.55)$$

donde

$$H(z) = \frac{(z - z_1)(z - z_2)}{(z^2 - 2z \cos[\alpha\Delta] + 1)} \quad (3.56)$$

y $e(t_k)$ es ruido blanco Gaussiano de media cero y varianza:

$$\text{var}\{e(t_k)\} = \frac{\eta a}{\Delta^2 |z_1| |z_2|}. \quad (3.57)$$

Se utilizó la magnitud de las raíces z_1 y z_2 en el caso de que éstas sean complejas conjugadas.

3.6. Conclusiones

En este capítulo se han presentado los desarrollos necesarios para entender el proceso de muestreo de un sistema estocástico lineal utilizando tanto muestreo instantáneo como integrado. Además, se ha presentado los modelos de datos muestreados de sistemas oscilatorios obtenidos indirectamente mediante el espectro de potencia de una señal auxiliar denominada $\xi(t_k)$.

MODELO Y ESTIMACIÓN DE MÁXIMA VEROSIMILITUD DE OSCILADORES EN TIEMPO-CONTINUO

4.1. Introducción

En este capítulo se obtiene el modelo equivalente exacto en tiempo discreto de los sistemas osciladores cuando se tienen mediciones de posición, velocidad y aceleración de vibraciones. Inicialmente se plantea el modelo en tiempo continuo de las vibraciones. Luego se obtiene el modelo equivalente en espacio de estado. Posteriormente, se obtiene el modelo exacto equivalente en tiempo discreto para cada una de las posibles mediciones muestreando instantáneamente o utilizando un AAF previo al muestreo. Además, se plantea la identificación de los parámetros a modelar a través del método de Máxima Verosimilitud (ML). Finalizando con la presentación de las conclusiones.

4.2. Modelo del sistema

La ecuación diferencial que modela un oscilador está dada por:

$$\frac{d^2}{dt^2}\xi(t) + \alpha^2\xi(t) = \dot{w}(t), \quad (4.1)$$

donde $\xi(t)$ es la salida del oscilador y $\dot{w}(t)$ corresponde al ruido que lo alimenta. Por simplicidad se utiliza el operador de Heaviside $\frac{d}{dt} = s$, con lo que se obtiene

$$s^2\xi(t) + \alpha^2\xi(t) = \dot{w}(t) \Rightarrow \xi(t) = \frac{1}{s^2 + \alpha^2}\dot{w}(t). \quad (4.2)$$

El sistema de vibraciones es modelado como una combinación lineal de osciladores alimentados cada uno de ellos por un ruido. Por lo tanto, el sistema de interés queda

expresado como:

$$y(t) = \sum_{i=1}^n \frac{1}{s^2 + \alpha_i^2} \dot{w}_i(t), \quad (4.3)$$

donde $y(t)$ es la salida del sistema, α_i^2 denota la frecuencia al cuadrado de la vibración expresada en $[\text{rad}/s]^2$. $w_i(t)$ es un proceso de Wiener, y así $\dot{w}_i(t)$ es ruido Gaussiano de media cero y varianza $\mathbb{E}\{\dot{w}_i(t)\dot{w}_i(\tau)\} = \sigma_i^2 \delta(t - \tau)$. Los ruidos $\dot{w}_i(t)$ de cada vibración podrían estar correlacionados entre sí, pero para efectos de simplificar la presentación se consideran las correlaciones cruzadas iguales a cero. Es decir:

$$\mathbb{E} \left\{ \begin{bmatrix} \dot{w}_1(t) \\ \vdots \\ \dot{w}_n(t) \end{bmatrix} \begin{bmatrix} \dot{w}_1(\tau) \\ \vdots \\ \dot{w}_n(\tau) \end{bmatrix}^T \right\} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} \delta(t - \tau). \quad (4.4)$$

Si se considera el espectro de potencia del ruido de tiempo continuo $\dot{w}_i(t)$ como $\sigma_i^2 = 1 \ \forall i$, la magnitud de cada pico del espectro de potencia de la señal $y(t)$ en (4.3) se puede obtener definiendo β_i como la ganancia del oscilador. Es decir:

$$y(t) = \sum_{i=1}^n \frac{\beta_i}{s^2 + \alpha_i^2} \dot{w}_i(t) \quad (4.5)$$

con $\beta_i > 0$ y matriz de covarianza de los ruidos de tiempo continuo $\dot{w}_i(t)$:

$$\mathbb{E} \left\{ \begin{bmatrix} \dot{w}_1(t) \\ \vdots \\ \dot{w}_n(t) \end{bmatrix} \begin{bmatrix} \dot{w}_1(\tau) \\ \vdots \\ \dot{w}_n(\tau) \end{bmatrix}^T \right\} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \delta(t - \tau). \quad (4.6)$$

Note que por abuso de notación se está utilizando la misma letra w para denotar el ruido con varianza unitaria.

4.3. Modelo en espacio de estado equivalente

A modo de obtener un modelo en espacio de estado equivalente del modelo en (4.5), se parametrizan las matrices de espacio de estado en función de las frecuencias de oscilación y de su magnitud. Por lo tanto, el modelo equivalente en espacio de estado está dado por:

$$\dot{x}(t) = Ax(t) + \kappa \dot{w}(t) \quad (4.7)$$

$$y(t) = Cx(t), \quad (4.8)$$

donde la matriz de estado $A \in \mathbb{R}^{2n \times 2n}$, la matriz de entrada $\kappa \in \mathbb{R}^{2n \times 1}$ y la matriz de salida $C \in \mathbb{R}^{1 \times 2n}$ tienen la siguiente estructura:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ -\alpha_1^2 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & -\alpha_2^2 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\alpha_n^2 & 0 \end{bmatrix}, \quad (4.9)$$

$$\kappa = \begin{bmatrix} 0 & \beta_1 & 0 & \beta_2 & \cdots & 0 & \beta_n \end{bmatrix}^T, \quad (4.10)$$

$$C = \begin{bmatrix} 1 & 0 & 1 & 0 & \cdots & 1 & 0 \end{bmatrix}, \quad (4.11)$$

donde C se asume independiente de los parámetros $\vec{\alpha}$ y $\vec{\beta}$. Finalmente, el vector de parámetros a estimar está dado por

$$\theta = \begin{bmatrix} \vec{\alpha}^T & \vec{\beta}^T \end{bmatrix}^T \quad (4.12)$$

donde $\vec{\alpha} = [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n]^T$ y $\vec{\beta} = [\beta_1 \ \beta_2 \ \cdots \ \beta_n]^T$. Cabe mencionar que en esta tesis, por claridad de exposición se consideran las condiciones iniciales nulas, cuya extensión aparece en el artículo de conferencia [45]. ,

4.4. Modelo equivalente en tiempo discreto

El estado $x(t)$ en (4.7) es una señal bien definida y su valor entre muestras discretas $t_k, k = \{0, \dots, N\}$, se puede encontrar fácilmente [14]. Utilizando el Lema 1 y suponiendo un intervalo de muestreo $\Delta = t_{k+1} - t_k$ constante y positivo, se obtiene que el modelo de vibraciones en tiempo discreto equivalente cuando se muestra de forma instantánea está dado por:

$$\begin{aligned} x(t_{k+1}) &= A_d x(t_k) + n(t_{k+1}) \\ y(t_k) &= C x(t_k) \end{aligned} \quad (4.13)$$

donde $A_d = e^{A\Delta}$ y $n(t_{k+1}) = \int_0^\Delta e^{A\eta} \kappa \dot{w}(t_{k+1} - \eta) d\eta$ es ruido blanco de media cero y varianza Q_d .

Ejemplo 1:

En este ejemplo, se consideran tres osciladores en tiempo continuo modelado como un sistema CARMA, cada uno de ellos corresponden a datos de medición de posición, velocidad y aceleración:

(i) Medición de posición

$$\dot{z}(t) = \frac{1}{s^2 + \alpha^2} \dot{w}(t) \quad (4.14)$$

(ii) Medición de velocidad

$$\dot{z}(t) = \frac{s}{s^2 + \alpha^2} \dot{w}(t) \quad (4.15)$$

(iii) Medición de aceleración

$$\dot{z}(t) = \frac{s^2}{s^2 + \alpha^2} \dot{w}(t) \quad (4.16)$$

donde $s = \frac{d}{dt}$ y $\dot{w}(t)$ es ruido blanco Gaussiano en tiempo continuo de media cero y varianza $\mathbb{E}\{\dot{w}(t)\dot{w}^T(\tau)\} = \sigma^2\delta(t - \tau)$. Cada uno de los sistemas (4.14)-(4.16) se puede representar en su forma espacio de estado como:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + \kappa\dot{w}(t) \\ \dot{z}(t) &= Cx(t) \end{aligned}$$

donde las matrices de estado A , de entrada κ y salida C están dadas por:

(i) Medición de posición

$$A = \begin{bmatrix} 0 & 1 \\ -\alpha^2 & 0 \end{bmatrix} \quad \kappa = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad (4.17)$$

(ii) Medición de velocidad

$$A = \begin{bmatrix} 0 & 1 \\ -\alpha^2 & 0 \end{bmatrix} \quad \kappa = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad C = \begin{bmatrix} 0 & 1 \end{bmatrix} \quad (4.18)$$

(iii) Medición de aceleración

$$A = \begin{bmatrix} 0 & 1 \\ -\alpha^2 & 0 \end{bmatrix} \quad \kappa = \begin{bmatrix} 0 \\ -\alpha^2 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad (4.19)$$

Note que para la medición de aceleración no se puede hacer muestreo instantáneo debido a que el proceso $\dot{w}(t)$ en (4.16) tiene paso directo. Por ende, se requiere un filtro previo a muestrear. Típicamente se utiliza el muestreo integrado [13], motivo por el cual se considera en este trabajo.

El modelo de vibraciones equivalente en tiempo discreto, cuando se muestrea de forma instantánea, para las mediciones de posición y velocidad está dado por (4.13) donde el cálculo de las matrices A_d y Q_d se obtienen en los Lemas 4 y 5 del Apéndice B.1 respectivamente.

A través del muestreo integrado, el modelo equivalente en tiempo discreto al utilizar el Lema 2 queda dado por:

$$\begin{bmatrix} x(t_{k+1}) \\ y(t_{k+1})\Delta \end{bmatrix} = \begin{bmatrix} A_d \\ C_d \end{bmatrix} x(t_k) + n(t_{k+1}), \quad (4.20)$$

donde $A_d = e^{A\Delta}$, $C_d = CA^{-1}(e^{A\Delta} - I)$ y el ruido blanco $n(t_{k+1})$ es de media cero con matriz de covarianza dada por:

$$\mathbb{E} \{n(t_{k+1})n(t_{k+1})^T\} = \int_0^\Delta e^{A_x\eta} \begin{bmatrix} Q_c & S_c \\ S_c & R_c \end{bmatrix} e^{A_x^T\eta} d\eta = \begin{bmatrix} Q_d & S_d \\ S_d & R_d \end{bmatrix}$$

donde la matriz extendida A_x está dada por

$$A_x = \begin{bmatrix} A & 0 \\ C & 0 \end{bmatrix}$$

El cálculo de las matrices exponenciales $e^{A\Delta}$ y $e^{A_x\Delta}$ se encuentran en el Lema 4 del Apéndice B.1. Respecto a la matriz de covarianza del ruido $n(t_{k+1})$, su cálculo se desarrolla en el Lema 6 del Apéndice B.1. La matriz de salida $C_d = CA^{-1}\{e^{A\Delta} - I\}$ para cada una de las mediciones está dada por:

- Medición de posición

$$C_d = \begin{bmatrix} \frac{\sin[\alpha\Delta]}{\alpha} & \frac{1-\cos[\alpha\Delta]}{\alpha^2} \end{bmatrix} \quad (4.21)$$

- Medición de velocidad

$$C_d = \begin{bmatrix} \cos[\alpha\Delta] - 1 & \frac{\sin[\alpha\Delta]}{\alpha} \end{bmatrix} \quad (4.22)$$

- Medición de aceleración

$$C_d = \begin{bmatrix} \frac{\sin[\alpha\Delta]}{\alpha} & \frac{1-\cos[\alpha\Delta]}{\alpha^2} \end{bmatrix} \quad (4.23)$$

El modelo de tiempo discreto equivalente para sistemas lineales de segundo orden se presentó en [13, Ejemplo 3.1] en el caso asintótico ($\Delta \rightarrow 0$) y no asintótico. La suposición subyacente era que los polos del sistema de tiempo continuo eran reales. Esto significa que el modelo de tiempo discreto equivalente no asintótico para el oscilador en (4.14) no está disponible. En el siguiente teorema, se presenta este resultado.

Teorema 3. El modelo de datos muestreados equivalente para los osciladores de tiempo continuo descrito en (4.14)-(4.16) utilizando un muestreo regular Δ está dado por el siguiente modelo ARMA:

$$y(t_k) = H(z)e(t_k) \quad (4.24)$$

donde $H(z)$ es un sistema estable de fase mínima y $e(t_k)$ es ruido blanco Gaussiano de media cero con varianza λ . Considerando el Ejemplo 4.4, el filtro $H(z)$ y la varianza λ al utilizar un **muestreo instantáneo** están dados por:

$$H(z) = \frac{z(z - z_1)}{(z^2 - 2z \cos[\alpha\Delta] + 1)}, \quad (4.25)$$

donde z_1 es la raíz dentro del círculo unitario del siguiente polinomio $R_1(z)$:

$$R_1(z) = z^2 + \frac{b}{a}z + 1 \quad (4.26)$$

Los valores de a y b para las mediciones de posición y velocidad de un oscilador están dadas por:

- Medición de posición

$$\begin{aligned} b &= 2\alpha\Delta - \sin[2\alpha\Delta] \\ a &= \sin[\alpha\Delta] - \alpha\Delta \cos[\alpha\Delta] \end{aligned} \quad (4.27)$$

- Medición de velocidad

$$\begin{aligned} b &= 2\alpha\Delta + \sin[2\alpha\Delta] \\ a &= -\sin[\alpha\Delta] - \alpha\Delta \cos[\alpha\Delta] \end{aligned} \quad (4.28)$$

La varianza λ del ruido $e(t_k)$ para las mediciones de posición y velocidad de un oscilador con muestreo instantáneo están dadas por:

- Medición de posición

$$\lambda = \frac{\sigma^2 a}{2\alpha|z_1||z_2|} \quad (4.29)$$

- Medición de velocidad

$$\lambda = \frac{\sigma^2 a}{2\alpha^5|z_1||z_2|} \quad (4.30)$$

Por otra parte, si se utiliza un muestreo integrado $H(z)$ tiene la siguiente estructura:

$$H(z) = \frac{(z - z_1)(z - z_2)}{(z^2 - 2z \cos[\alpha\Delta] + 1)}, \quad (4.31)$$

donde z_1 y z_2 son las raíces dentro del círculo unitario del siguiente polinomio $R_2(z)$:

$$R_2(z) = z^4 + \frac{b}{a}z^3 + \frac{c}{a}z^2 + \frac{b}{a}z + 1 \quad (4.32)$$

Los valores de a , b y c para cada una de las mediciones de vibración están dadas por:

- Medición de posición

$$\begin{aligned}
 a &= 2\alpha\Delta + \alpha\Delta \cos[\alpha\Delta] - 3 \sin[\alpha\Delta] \\
 b &= -2\alpha\Delta - 10\alpha\Delta \cos[\alpha\Delta] + 6 \sin[\alpha\Delta] + 6 \sin[\alpha\Delta] \cos[\alpha\Delta] \\
 c &= 12\alpha\Delta + 2\alpha\Delta \cos[\alpha\Delta] + 4\alpha\Delta \cos[2\alpha\Delta] - 6 \sin[\alpha\Delta] - 12 \sin[\alpha\Delta] \cos[\alpha\Delta]
 \end{aligned} \tag{4.33}$$

- Medición de velocidad

$$\begin{aligned}
 a &= \alpha\Delta \cos[\alpha\Delta] - \sin[\alpha\Delta] \\
 b &= -2\alpha\Delta - 2\alpha\Delta \cos[\alpha\Delta] + 2 \sin[\alpha\Delta] + \sin[2\alpha\Delta] \\
 c &= 4\alpha\Delta + 2\alpha\Delta \cos[\alpha\Delta] - 2 \sin[\alpha\Delta] - 2 \sin[2\alpha\Delta]
 \end{aligned} \tag{4.34}$$

- Medición de aceleración

$$\begin{aligned}
 a &= \alpha\Delta \cos[\alpha\Delta] + \sin[\alpha\Delta] \\
 b &= -2\alpha\Delta - 2\alpha\Delta \cos[\alpha\Delta] - 2 \sin[\alpha\Delta] - \sin[2\alpha\Delta] \\
 c &= 4\alpha\Delta + 2\alpha\Delta \cos[\alpha\Delta] + 2 \sin[\alpha\Delta] + 2 \sin[2\alpha\Delta]
 \end{aligned} \tag{4.35}$$

La varianza λ del ruido $e(t_k)$ para cada una de las mediciones de oscilación están dadas por:

- Medición de posición

$$\lambda = \frac{\sigma^2 a}{2\alpha^5 \Delta^2 |z_1| |z_2|} \tag{4.36}$$

- Medición de velocidad

$$\lambda = \frac{\sigma^2 a}{2\alpha^3 \Delta^2 |z_1| |z_2|} \tag{4.37}$$

- Medición de aceleración

$$\lambda = \frac{\sigma^2 a}{2\alpha \Delta^2 |z_1| |z_2|} \tag{4.38}$$

Demostración. Ver la Sección 3.5.1 y 3.5.2 en caso de muestreo instantáneo e integrado, respectivamente. \square

Observación 4. Note que el resultado asintótico cuando $\Delta \rightarrow 0$ corresponde al cero de muestreo (asintótico) asociado [13].

4.5. Método de Máxima Verosimilitud

Un método ampliamente utilizado para la mayoría de los problemas de estimación es el método de Máxima Verosimilitud (ML) [30] y [31], en el cual el problema de estimación se puede escribir como el siguiente problema de optimización:

$$\hat{\theta}_{ML} = \arg \max_{\theta} L_N(\theta), \quad (4.39)$$

donde $L_N(\theta)$ es la función de verosimilitud que se obtiene a partir del modelo del sistema y las mediciones. A modo de simplificar los cálculos, se define la función log-verosimilitud:

$$l_N(\theta) = \log[L_N(\theta)]. \quad (4.40)$$

Dado que $\log[\cdot]$ es una función monótona, el problema de optimización puede ser reescrito como el siguiente problema de minimización:

$$\hat{\theta}_{ML} = \arg \min_{\theta} -l_N(\theta). \quad (4.41)$$

La función log-verosimilitud se puede calcular mediante:

- La descomposición del error de predicción a través del filtro de Kalman [27, 31, 46].

$$l_N(\theta) = -\frac{1}{2} \sum_{k=1}^N \frac{\varepsilon_k^2}{\Lambda_k} - \frac{1}{2} \sum_{k=1}^N \log[\det\{\Lambda_k\}], \quad (4.42)$$

donde el error de predicción ε_k se define como

$$\varepsilon_k(\theta) = y_k - \mathbb{E}\{y_k | y_{0:k-1}, \theta\}, \quad (4.43)$$

con media cero y varianza

$$\Lambda_k(\theta) = \mathbb{E}\{\varepsilon_k^2 | y_{0:k-1}, \theta\}, \quad (4.44)$$

donde y_0 se asume distribuido normal con media $\hat{y}_{0|-1}(\theta)$ y varianza $\Sigma_{0|-1}(\theta)$.

- La siguiente expresión:

$$\ell(\theta) = -\frac{N}{2} \log[\lambda] - \frac{1}{2} \sum_{t=1}^N \left[\frac{y(t_k)}{H(z, \theta)} \right]^2, \quad (4.45)$$

donde $y(t_k)$ corresponde a un modelo ARMA $y(t_k) = H(z)e(t_k)$, $H(z)$ es un sistema estable de fase mínima y $e(t_k)$ es ruido blanco Gaussiano que tiene media cero y varianza λ .

4.6. Conclusiones

En este capítulo, se ha planteado modelar las vibraciones como una combinación lineal de osciladores alimentados cada uno de ellos por un ruido. En ese sentido, el problema a resolver corresponde a estimar los parámetros de cada uno de dichos ruidos. El sistema de interés se expresó como modelos de espacio de estado en tiempo continuo muestreados de manera uniforme. Luego se obtuvo el modelo equivalente exacto (sin aproximaciones) en tiempo discreto considerando un tiempo muestreo finito Δ y utilizando i) muestreo instantáneo y ii) muestreo integrado. A pesar de que la representación entrada salida en tiempo discreto no contiene espectro por ser el sistema de interés oscilatorio, es posible obtener indirectamente el modelo de datos muestreados que representa al sistema en tiempo discreto. Los modelos exactos en tiempo discreto se obtuvieron para los tres casos de medición de vibración: posición, velocidad y aceleración. Esta última medición solo puede muestrearse utilizando un filtro previo a tomar muestras. Por último, se planteó el problema de estimación como un problema de optimización a resolver con el algoritmo de Máxima Verosimilitud y se presentaron dos formas de calcular la función de log-verosimilitud: i) descomposición del error de predicción a través del filtro de Kalman o ii) utilizando directamente el modelo ARMA exacto en tiempo discreto. Estas dos formas de cálculo se utilizarán en el siguiente capítulo. Además, se presentó que el análisis de los ceros de muestreo asintóticos ($\Delta \rightarrow 0$) en [13] también es válido en el caso en que el sistema de tiempo continuo es un oscilador con polos en el eje imaginario .

OPTIMIZACIÓN DE MÁXIMA VEROSIMILITUD

5.1. Introducción

En este capítulo se identifican los modelos exactos equivalente en tiempo discreto de las vibraciones al medir posición velocidad y aceleración. Se desarrollan dos algoritmos de optimización para la estimación de los parámetros, uno local y otro global. Con el fin de analizar el comportamiento de la función log-verosimilitud se presentan varios ejemplos numéricos que muestran la presencia de varios máximos locales, para finalmente presentar las conclusiones correspondientes.

5.2. Optimización Local

Para desarrollar el algoritmo de optimización local se utiliza la descomposición del error de predicción. De esta manera se calcula la función de log-verosimilitud ocupando (4.42). Los valores de error de prediccción (4.43) y su varianza (4.44) se pueden obtener utilizando el filtro de Kalman estándar, estimando así el estado y la varianza del sistema espacio de estado en (4.7) como:

- Etapa de predicción del Filtro de Kalman

$$K_k = \Sigma_{k|k-1} C^T (C \Sigma_{k|k-1} C^T)^{-1} \quad (5.1)$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (y_k - C \hat{x}_{k|k-1}) \quad (5.2)$$

$$\Sigma_{k|k} = (I - K_k C) \Sigma_{k|k-1} \quad (5.3)$$

- Etapa de corrección del Filtro de Kalman

$$\hat{x}_{k+1|k} = e^{A\Delta} \hat{x}_{k|k} \quad (5.4)$$

$$\Sigma_{k+1|k} = e^{A\Delta} \Sigma_{k|k} e^{A^T \Delta} + Q_d \quad (5.5)$$

donde K_k es la ganancia del filtro de Kalman, $\hat{x}_{k|k}$ y $\Sigma_{k|k}$ son las estimaciones del estado y de la matriz de covarianza del error de estimación en el instante k dadas las mediciones entrada-salida del sistema hasta el instante k . Mientras que $\hat{x}_{k+1|k}$ y $\Sigma_{k+1|k}$ son las estimaciones del estado y de la matriz de covarianza del error de estimación en el instante $k+1$ dadas las mediciones entrada-salida del sistema hasta el instante k . Por lo tanto, el error de predicción y su varianza se pueden representar como:

$$\varepsilon_k = y_k - C\hat{x}_{k|k-1} \quad (5.6)$$

$$\Lambda_k = C\Sigma_{k|k-1}C^T. \quad (5.7)$$

Note que en este caso la función log-verosimilitud (4.40) es no lineal y no convexa, por lo que es necesario resolver el problema de minimización en (4.41) a través de métodos numéricos. Dado lo anterior se utilizará la técnica de Cuasi-Newton [29], donde no se requiere calcular la inversa de la matriz Hessiana de manera exacta, sino que se aproxima numéricamente. De esta manera, la estimación del parámetro se actualiza iterativamente como:

$$\hat{\theta}^{m+1} = \hat{\theta}^m - \mu^m S^m \nabla l(\theta)|_{\theta=\hat{\theta}^m}, \quad (5.8)$$

donde $\nabla l(\theta)$ corresponde al vector gradiente de $l(\theta)$ y S^m es una matriz que aproxima la inversa de la matriz Hessiana $\nabla^2 l(\theta)|_{\theta=\hat{\theta}}$, mientras que μ^m corresponde al paso de la m -ésima iteración [29].

Observación 5. Los métodos de cuasi-newton resultan en más iteraciones que el método de Newton, pero cada iteración con un menor costo computacional. Dentro de los métodos de cuasi newton más populares está BFGS [29] y Barzilai-Borwein [47]. \square

El vector gradiente de se calcula como:

$$\frac{\partial l(\theta)}{\partial \theta_j} = - \sum_{k=1}^N \frac{\varepsilon_k}{\Lambda_k} \frac{\partial \varepsilon_k}{\partial \theta_j} - \frac{1}{2} \sum_{k=1}^N \left(\frac{1}{\Lambda_k} - \frac{\varepsilon_k^2}{\Lambda_k^2} \right) \frac{\partial \Lambda_k}{\partial \theta_j}, \quad j = 1, \dots, n, \quad (5.9)$$

donde θ_j denota a la j -ésima componente del vector de parámetros θ . Los términos $\frac{\partial \varepsilon_k}{\partial \theta_j}$ y $\frac{\partial \Lambda_k}{\partial \theta_j}$ se calculan como

$$\frac{\partial \varepsilon_k}{\partial \theta_j} = -C \frac{\partial \hat{x}_{k|k-1}}{\partial \theta_j} \quad (5.10)$$

$$\frac{\partial \Lambda_k}{\partial \theta_j} = C \frac{\partial \Sigma_{k|k-1}}{\partial \theta_j} C^T \quad (5.11)$$

Note que $\frac{\partial \varepsilon_k}{\partial \theta_j}$ y $\frac{\partial \Lambda_k}{\partial \theta_j}$ dependen de las derivadas del estado y la varianza estimada (provenientes del filtro de Kalman) respecto al vector de parámetros θ . Los detalles de los cálculos correspondientes se encuentran en el Apéndice B.

Por otra parte, en la optimización de sistemas no lineales se encuentran generalmente máximos locales, por cuanto se requiere que la inicialización de la optimización sea cercana al máximo global.

Ejemplo Numérico

Ejemplo 1: En este ejemplo, se considera un oscilador en tiempo continuo modelado como un sistema CARMA de la siguiente manera

$$\dot{z}(t) = \frac{\beta}{s^2 + \alpha^2} \dot{w}(t).$$

El modelo de espacio de estado equivalente presenta las siguientes matrices de espacio de estado:

$$A = \begin{bmatrix} 0 & 1 \\ -\alpha^2 & 0 \end{bmatrix} \quad \kappa = \begin{bmatrix} 0 \\ \beta \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

Se asume que el proceso $\dot{w}(t)$ es de media cero y tiene un espectro de potencia $\sigma^2 = 1$. El vector de parámetros a identificar es $\theta = [\alpha \ \beta]^T$. Los datos simulados se generan con los siguientes parámetros: (i) $\alpha = 600\pi[\text{rad/s}]$, (ii) $\beta = 1$, (iii) período de muestreo $\Delta = 10^{-4}[\text{s}]$. La simulación se realizó utilizando muestreo instantáneo con $N = 500$ y $N = 1000$ muestras de y_k , para así obtener el valor de la función de verosimilitud a maximizar.

En la Fig. 5.1 se muestran las gráficas de las curvas de nivel de la función de verosimilitud, donde se observa claramente la presencia de máximos locales. En dicha figura se puede observar también que el número de máximos locales y los anchos de los lóbulos de la función de máxima verosimilitud varía con el número de datos. En particular, con $N = 500$ el ancho de los lóbulos es mayor y la cantidad de mínimos locales es menor. Dado que estamos identificando osciladores, la inicialización del algoritmo de estimación se realiza a partir de la frecuencia correspondiente al máximo de la densidad espectral de potencia (PSD) de la señal $y_{0:N-1}$, como se muestra en la Fig. 5.2. La PSD se obtiene a partir de la transformada rápida de Fourier de los datos. El algoritmo de optimización fue implementado a través de la función *fminunc* de Matlab, con la opción '*Algorithm', 'quasi-Newton'*. Para más detalles ver el Apéndice A.3.

En la Fig. 5.3(a) se presentan los resultados de la simulación de 100 realizaciones de Monte Carlo, utilizando $N = 500$, tomando como punto de partida para la ganancia del oscilador $\hat{\beta}^0 = 2$. En dicha figura se observa claramente que se obtiene una buena estimación de los parámetros. Esto se debe a que el punto de inicialización se encuentra lo suficientemente cerca de los valores reales de los parámetros. En la Fig. 5.3(b) se presentan los resultados de la simulación de 100 realizaciones de Monte Carlo, utilizando $N = 1000$ y $\hat{\beta}^0 = 2$. Se puede apreciar que los resultados de la estimación son mejores

con $N = 500$, ya que el ancho de la región cuasi-convexa de la función ML alrededor del máximo global es más grande por lo que la inicialización escogida con $N = 500$ fue más adecuada en dicho caso. Por lo tanto, la Fig. 5.3 muestra lo importante que es la inicialización del algoritmo de estimación para encontrar el óptimo global de la función ML (4.42).

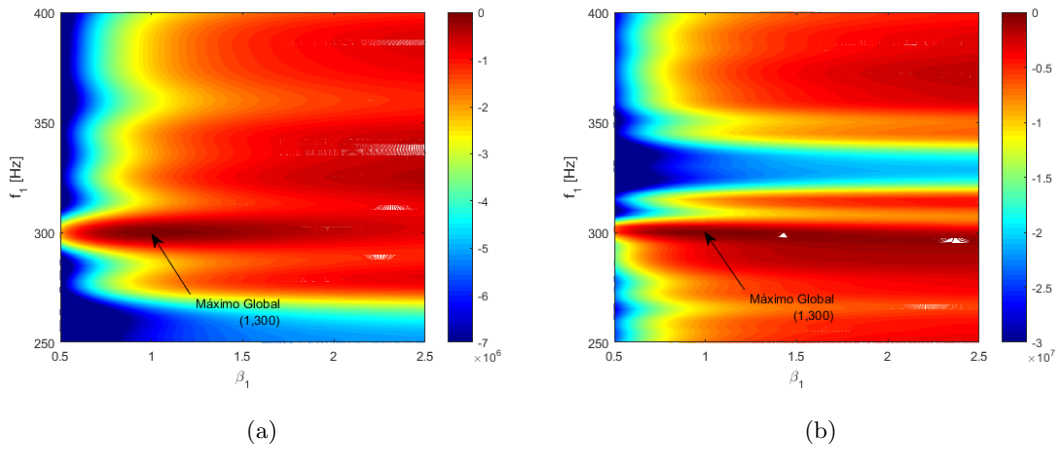


Fig. 5.1: Gráfico de la Función de Verosimilitud de vibraciones generadas con frecuencia en $f_1 = 300[Hz]$ y amplitud $\beta_1 = 1$, utilizando (a) $N = 500$ y (b) $N = 1000$ muestras de la señal de salida y_k .

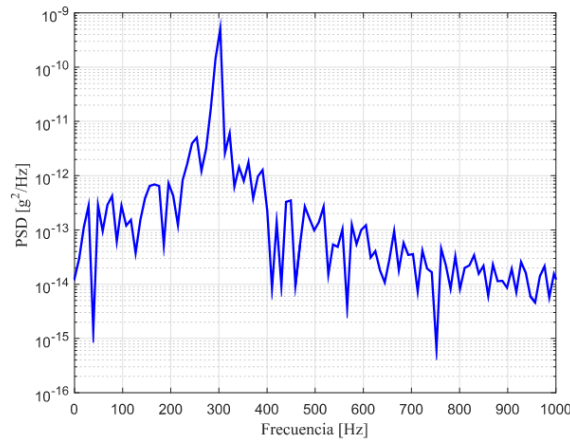


Fig. 5.2: Densidad espectral de potencia para el Ejemplo 1 con $N = 1000$

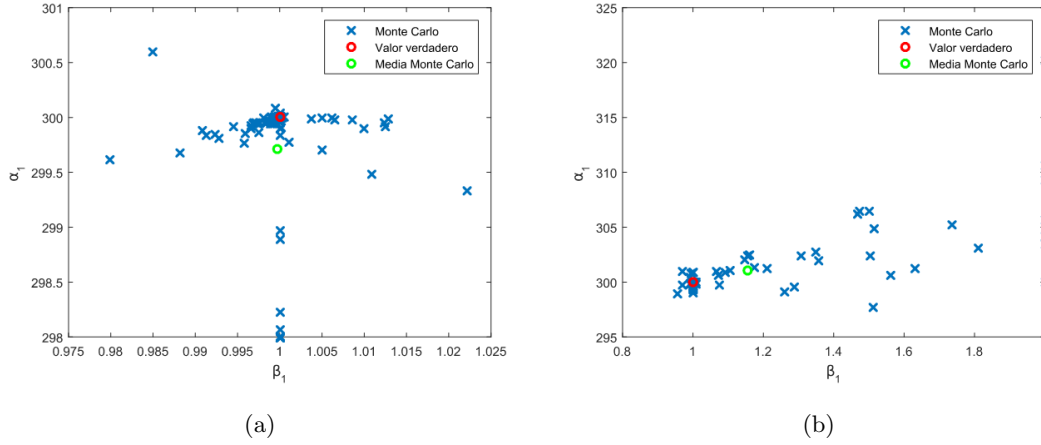


Fig. 5.3: Estimación de $\theta = [\alpha \ \beta]^T$ para 100 realizaciones de Monte Carlo con (a) $N = 500$ y (b) $N = 1000$ muestras de la señal de salida y_k .

Cambiando el parámetro de la frecuencia a $\alpha = 120\pi[\text{rad/s}]$, manteniendo $\beta = 1$ y período de muestreo $\Delta = 10^{-4}[\text{s}]$, se realiza la simulación utilizando muestreo instantáneo con $N = 500$ y $N = 1000$ muestras de y_k , para así obtener el valor de la función de verosimilitud a maximizar. En la Fig. 5.4 se muestran las gráficas de las curvas de nivel de la función de verosimilitud, donde se observa claramente la presencia de varios máximos locales, pero el ancho de los lóbulos de la función de máxima verosimilitud varía con el número de datos. En particular, con $N = 500$ el ancho del lóbulo alrededor del valor del parámetro verdadero es mayor.

En la Fig. 5.5(a) se presentan los resultados de la simulación de 100 realizaciones de Monte Carlo, utilizando $N = 500$, tomando como punto de partida para la ganancia del oscilador $\hat{\beta}^0 = 2$. En dicha figura se observa claramente que se obtiene una buena estimación de los parámetros ya que el punto de inicialización es generalmente cercano a los valores reales de los parámetros.

Ejemplo 2: En este ejemplo se considera un sistema con dos osciladores:

$$y(t) = \frac{\beta_1}{s^2 + \alpha_1^2} \dot{w}_1(t) + \frac{\beta_2}{s^2 + \alpha_2^2} \dot{w}_2(t)$$

El modelo espacio de estado equivalente está dado por las siguientes matrices:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\alpha_1^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\alpha_2^2 & 0 \end{bmatrix}; \quad \kappa = \begin{bmatrix} 0 \\ \beta_1 \\ 0 \\ \beta_2 \end{bmatrix}; \quad C = [1 \ 0 \ 1 \ 0]$$

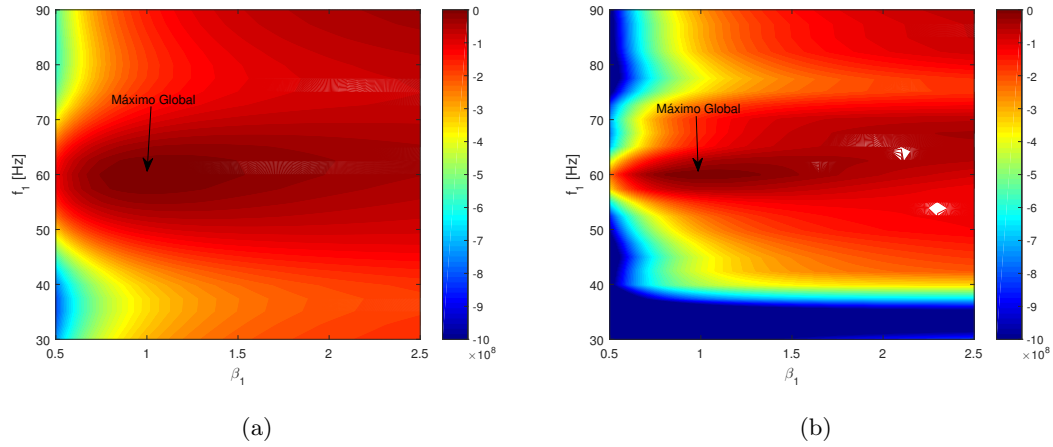


Fig. 5.4: Gráfico de la Función de Verosimilitud de vibraciones generadas con frecuencia en $f_1 = 60[Hz]$ y amplitud $\beta_1 = 1$, utilizando (a) $N = 500$ y (b) $N = 1000$ muestras de la señal de salida y_k .

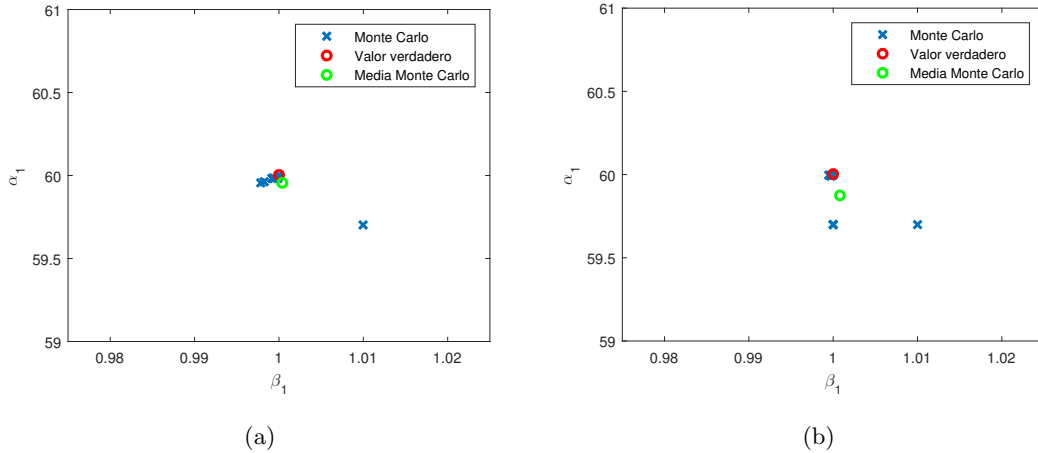


Fig. 5.5: Estimación de $\theta = [\alpha \beta]^T$ para 100 realizaciones de Monte Carlo con (a) $N = 500$ y (b) $N = 1000$ muestras de la señal de salida y_k .

Se asume que los procesos de $\dot{w}_1(t)$ y $\dot{w}_2(t)$ son de media cero y espectro de potencia $\sigma^2 = 1$ y no correlacionados. Con la finalidad de analizar un problema que se puede graficar se simplifica el vector de parámetros a identificar en $\theta = [\alpha_1 \alpha_2]^T$.

Los datos simulados se generan con los siguientes parámetros: (i) $\alpha_1 = 760\pi[rad/s]$, $\alpha_2 = 800\pi[rad/s]$, (ii) periodo de muestreo $\Delta = 10^{-4}[s]$, (iii) caso 1: $\beta_0 = [1 \ 1]^T$ y caso 2:

$\beta_0 = [3 \ 1]^T$. Al igual que en el ejemplo anterior, se analizan la función de verosimilitud considerando $N = 500$ y $N = 1000$. La función de ML se presenta en la Fig. 5.6, donde se observa claramente que los gráficos contienen varios máximos locales. Sin embargo, al igual que en Ejemplo 1, con $N = 500$ hay un menor número de máximos locales y la vecindad cuasi-convexa de la función ML alrededor del máximo global es más grande. Además, en los gráficos (c) y (d), se observa que la vecindad cuasi-convexa de la función ML alrededor del máximo global es ancha entorno a la frecuencia del segundo oscilador y delgada entorno a la frecuencia del primer oscilador debido al aumento de la ganancia de esta última.

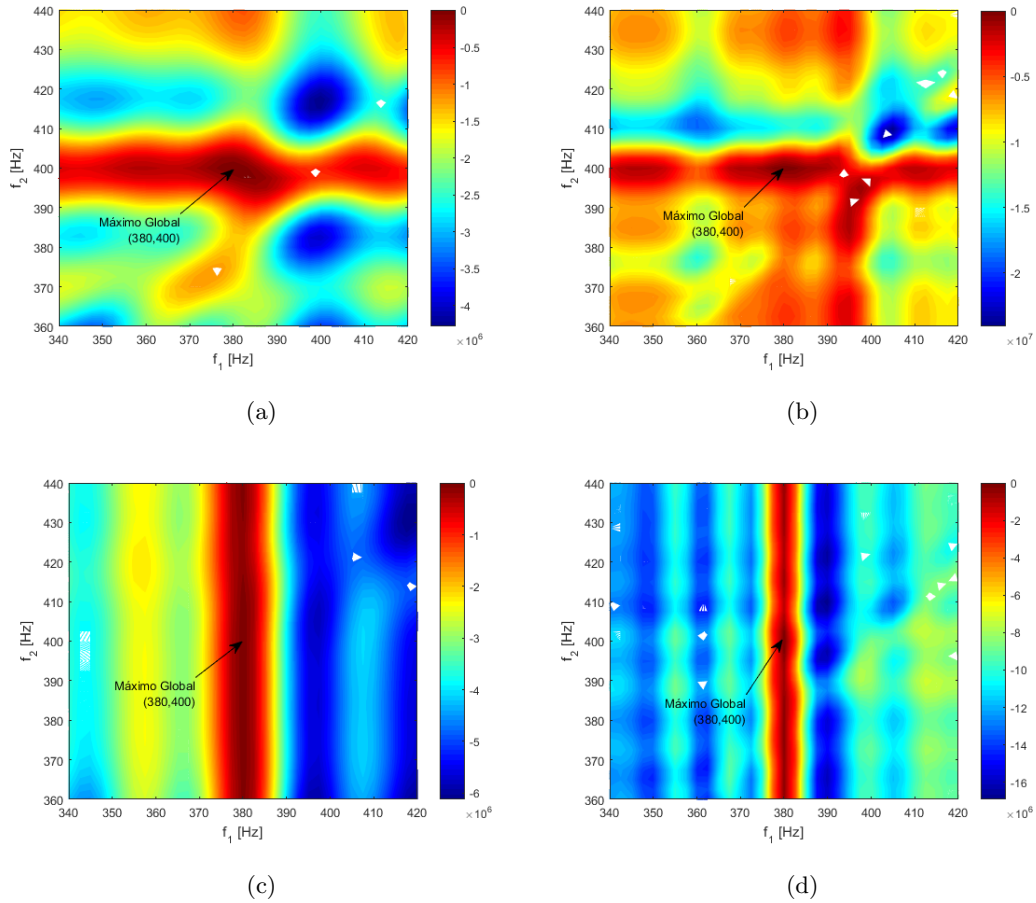


Fig. 5.6: Gráfico de la Función de Verosimilitud de las vibraciones generadas con frecuencias $f_1 = 380[Hz]$ y $f_2 = 400[Hz]$, utilizando (a) $\beta_0 = [1 \ 1]^T$ y $N = 500$, (b) $\beta_0 = [1 \ 1]^T$ y $N = 1000$, (c) $\beta_0 = [3 \ 1]^T$ y $N = 500$, y (d) $\beta_0 = [3 \ 1]^T$ y $N = 1000$.

En la Fig. 5.7 se presentan los resultados de la simulación de 100 realizaciones de Monte Carlo, utilizando $N = 1000$ y $\beta_0 = [3 \ 1]^T$. Se puede apreciar que se obtiene una buena estimación, resultado que también se observa en la Fig. 5.8, donde se presenta la magnitud del diagrama de bode ocupando los parámetros verdaderos y los estimados (media de Monte Carlo). Por lo tanto, a pesar del cambio en la ganancia del primer oscilador se logra conseguir una estimación apropiada, teniendo en cuenta lo importante de establecer el punto de inicialización del algoritmo cerca de los valores reales de los parámetros.

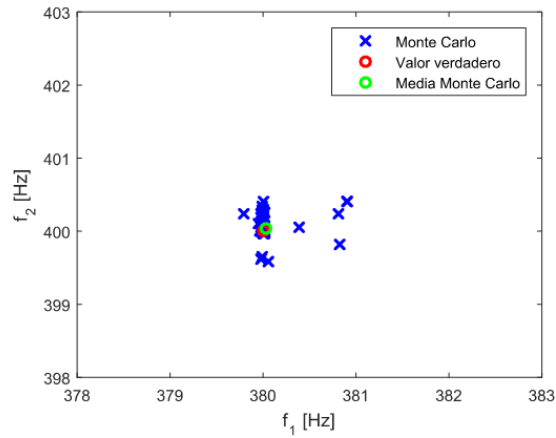


Fig. 5.7: Estimación de $\theta = [\alpha \ \beta]^T$ para 100 realizaciones de Monte Carlo con $N = 1000$ y $\beta_0 = [3 \ 1]^T$.

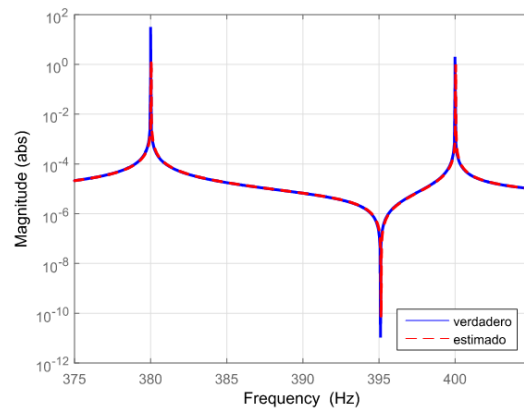


Fig. 5.8: Magnitud del diagrama de bode para el Ejemplo 2 con $N = 1000$ y $\beta_0 = [3 \ 1]^T$.

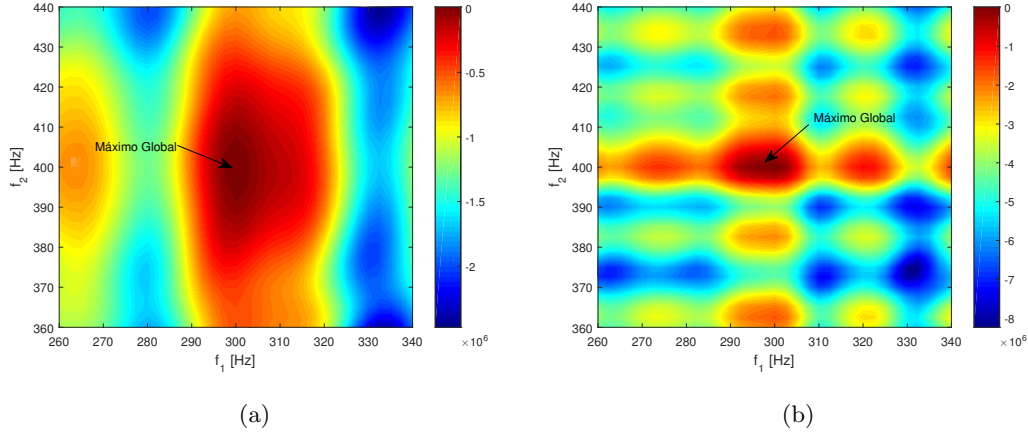


Fig. 5.9: Gráfico de la Función de Verosimilitud de vibraciones generadas con frecuencias en $f_1 = 300$ y $f_2 = 400[Hz]$, amplitud $\beta_1 = 1$ y $\beta_2 = 1$, utilizando (a) $N = 500$ y (b) $N = 1000$ muestras de la señal de salida y_k .

Se analiza la función de verosimilitud cuando las frecuencias de los osciladores están alejadas entre ellas por ejemplo $f_1 = 300$ y $f_2 = 400[Hz]$. Entonces, los datos simulados se generan con los siguientes parámetros: (i) $\alpha_1 = 600\pi[rad/s]$, $\alpha_2 = 800\pi[rad/s]$, (ii) periodo de muestreo $\Delta = 10^{-4}[s]$, (iii) $\beta_0 = [1 \ 1]^T$. La gráfica de la función de ML se presenta en la Fig. 5.9, donde se observa claramente que con $N = 500$ hay un menor número de máximo locales y la vecindad cuasi-convexa de la función ML alrededor del máximo global es más grande. Una vez más queda en evidencia la importancia de una inicialización del algoritmo cercano a los parámetros verdaderos.

5.3. Optimización Global

La función de log-verosimilitud se obtiene utilizando directamente el modelo ARMA exacto en tiempo discreto (4.24) como:

$$\ell(\theta) = -\frac{N}{2} \log[\lambda] - \frac{1}{2} \sum_{t=1}^N \left[\frac{y(t_k)}{H(z, \theta)} \right]^2,$$

donde $\theta = [\alpha \ \sigma^2]^T$, el filtro $H(z, \theta)$ y la varianza λ del ruido son proporcionados por el Teorema 3. Note que la función log-verosimilitud depende de las raíces de los polinomios $R_1(z)$ y $R_2(z)$. Esto implica que la optimización de la función log-verosimilitud debe realizarse numéricamente. Además, note que las raíces del polinomio $R_2(z) = z^4 + \frac{b}{a}z^3 +$

$\frac{c}{a}z^2 + \frac{b}{a}z + 1$ están dados por:

$$\begin{aligned}
z_1 &= -\frac{\bar{b}}{4} - \mathcal{L} - \frac{1}{2}\sqrt{-2 + \frac{\bar{b}^2}{2} - \bar{c} - \mathcal{M}} \\
z_2 &= -\frac{\bar{b}}{4} - \mathcal{L} + \frac{1}{2}\sqrt{-2 + \frac{\bar{b}^2}{2} - \bar{c} - \mathcal{M}} \\
z_3 &= -\frac{\bar{b}}{4} + \mathcal{L} - \frac{1}{2}\sqrt{-2 + \frac{\bar{b}^2}{2} - \bar{c} + \mathcal{M}} \\
z_4 &= -\frac{\bar{b}}{4} + \mathcal{L} + \frac{1}{2}\sqrt{-2 + \frac{\bar{b}^2}{2} - \bar{c} + \mathcal{M}} \\
\mathcal{L} &= \frac{\sqrt{8 + \bar{b}^2 - 4\bar{c}}}{4} \\
\mathcal{M} &= \frac{4\bar{b}\bar{c} - 8\bar{b} - \bar{b}^3}{2\sqrt{8 + \bar{b}^2 - 4\bar{c}}},
\end{aligned} \tag{5.12}$$

donde $\bar{b} = \frac{b}{a}$, y $\bar{c} = \frac{c}{a}$. En este caso, la funcionalidad de los ceros de muestreo (las raíces de $R_2(z)$ dentro del círculo unitario) de $H(z, \theta)$ cambia con los valores de los parámetros del correspondiente oscilador de tiempo continuo y del tiempo de muestreo. De hecho, si $\sigma^2 = 1$, $\Delta = 0,1$ y $\alpha = 50$, entonces las raíces de $R_2(z)$ que están dentro del círculo unitario corresponden a z_1 and z_3 . Sin embargo, para $\sigma^2 = 1$, $\Delta = 0,1$, y $\alpha = 15$ las raíces de $R_2(z)$ que están dentro del círculo unitario corresponden a z_2 y z_4 . Por lo tanto, se debe modificar la función de verosimilitud (y el algoritmo de optimización correspondiente) según el valor del parámetro α .

Ejemplo Numérico

Considerando el sistema oscilador dado en (4.14), con $\alpha = 10$ y $\sigma^2 = 1$. Se analizan dos casos: i) **muestreo rápido**: $\Delta = 1[ms]$ y ii) **muestreo lento**: $\Delta = 100[ms]$.

Muestreo Rápido : La función de verosimilitud correspondiente al muestreo instantáneo y muestreo AAF se presenta en la Fig. 5.10. Se observa que en ambos casos la función de verosimilitud es cóncava en la región de interés, pero la función de verosimilitud no cambia demasiado en el máximo global. Se realiza la optimización evaluando directamente la función de verosimilitud en una grilla fina. Los resultados se muestran en la tabla 5.1 donde se observa que el óptimo global corresponde al modelo "verdadero".

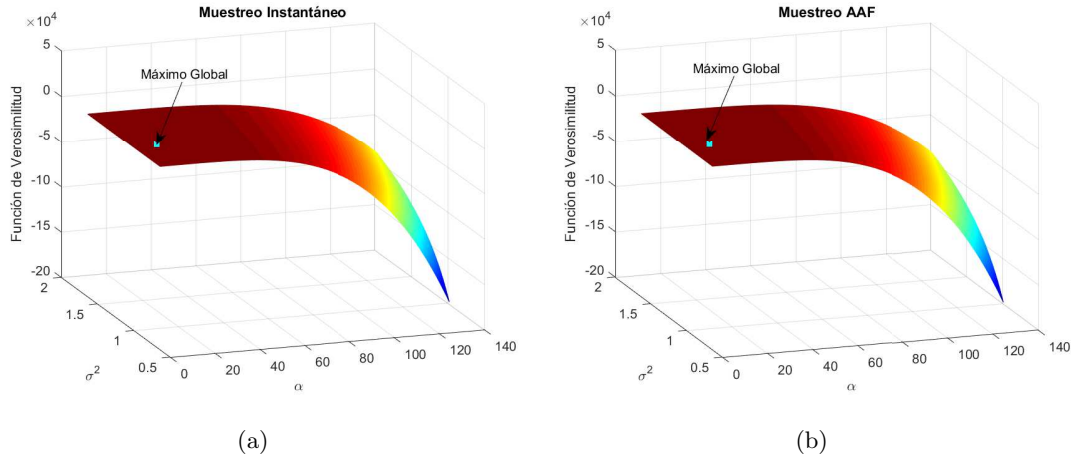


Fig. 5.10: Función Log-Verosimilitud para el sistema (4.14) utilizando (a) muestreo instantáneo y (b) muestreo integrado con $\Delta = 1 \text{ ms}$.

Muestreo Lento: En la Fig. 5.11 se presenta la función log-verosimilitud para (a) muestreo instantáneo y (b) integrado. Se observa claramente que la función de verosimilitud logarítmica muestra varios máximos locales, lo que dificulta el cálculo del máximo global. Se realiza la optimización evaluando directamente la función log-verosimilitud en una grilla fina. Los resultados se presentan en la tabla 5.1 donde se ve que el óptimo global corresponde al modelo "verdadero". También se observa que el AAF modifica la función log-verosimilitud, pero todavía presenta varios máximos locales.

Cuadro 5.1: Estimación de Máxima Verosimilitud.

	$\Delta = 1 \text{ ms}$		$\Delta = 100 \text{ ms}$	
	α	σ^2	α	σ^2
Muestreo Instantáneo	10,06	0,99	10,06	0,93
Muestreo Integrado	10,06	0,99	10,06	0,93

Por lo tanto, los algoritmos de optimización global son necesarios para obtener las estimaciones de máxima verosimilitud para los parámetros del oscilador de tiempo continuo.

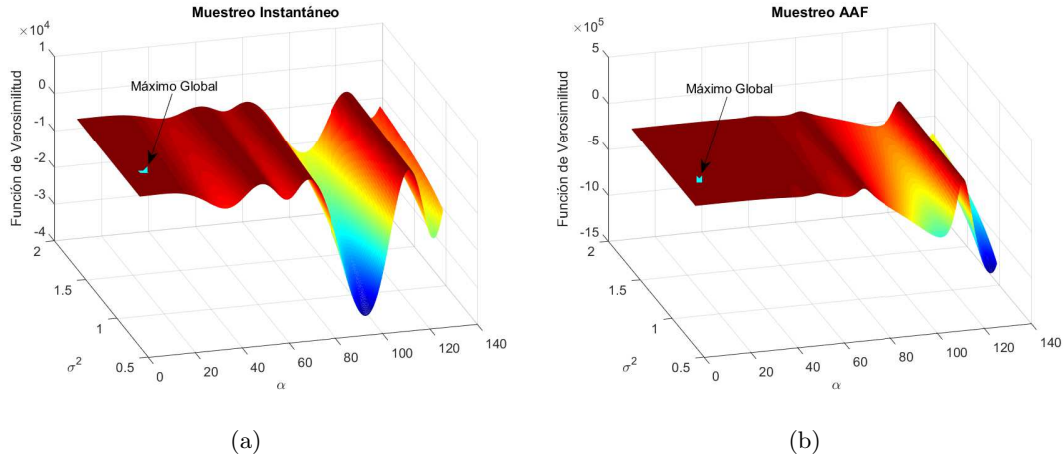


Fig. 5.11: Función Log-Verosimilitud para el sistema (4.14) utilizando (a) muestreo instantáneo y (b) muestreo integrado con $\Delta = 100 \text{ ms}$.

5.4. Conclusiones

En este capítulo se desarrollaron dos algoritmos de optimización para estimar los parámetros del modelo exacto en tiempo discreto del sistema de oscilación. Utilizando el algoritmo de optimización local, la función log-verosimilitud se calculó mediante la descomposición del error de predicción a través del filtro de Kalman, mientras que los parámetros fueron estimados localmente utilizando técnicas de cuasi-Newton. En este tipo de optimización, la estimación de los parámetros depende del número de muestras que se consideren, ya que hay una fuerte variación de máximos locales y ancho de los lóbulos de la función de verosimilitud para distintos números de mediciones. Además, el análisis también mostró que la inicialización del algoritmo de optimización local es fundamental para la estimación de los parámetros. Respecto a la optimización global se calcula la función log-verosimilitud utilizando el modelo ARMA. Por lo tanto, la función log-verosimilitud depende de las raíces de los polinomios $R_1(z)$ y $R_2(z)$, por consiguiente el problema de optimización se debe realizar numéricamente. Los ejemplos numéricos se centraron en el caso en que la tasa de muestreo es lenta y se mostró que la función log-verosimilitud muestra varios máximos locales. Sorprendentemente, la presencia de máximos locales ocurre incluso en el caso donde las muestras son tomadas utilizando un AAF.

CONCLUSIONES

6.1. Conclusiones

En esta tesis se ha definido un problema teórico a resolver a partir de la motivación de obtener un modelo de las vibraciones inducidas en los sistemas de Óptica Adaptativa. El problema se planteó modelando las vibraciones como una combinación lineal de osciladores alimentados cada uno de ellos por un ruido sin considerar amortiguamiento ni tampoco condición inicial. En ese sentido, el problema a resolver corresponde a estimar los parámetros de cada uno de dichos ruidos. El sistema de interés se expresó como modelos de espacio de estado en tiempo continuo muestreados de manera uniforme. Luego se obtuvo el modelo equivalente exacto (sin aproximaciones) en tiempo discreto considerando un tiempo muestreo finito Δ y utilizando i) muestreo instantáneo y ii) muestreo integrado. A través de tener un modelo exacto del sistema oscilante, el análisis de los ceros de muestreo asintóticos ($\Delta \rightarrow 0$) en [13] (análisis realizado suponiendo que los polos del sistema de tiempo continuo son reales) se valida en el caso en que el sistema de tiempo continuo es un oscilador con polos en el eje imaginario.

A pesar de que la representación entrada salida en tiempo discreto no contiene espectro por ser el sistema de interés oscilatorio, se obtuvo indirectamente el modelo de datos muestreados que representa al sistema en tiempo discreto. Los modelos exactos en tiempo discreto se obtuvieron para los tres casos de medición de vibración: posición, velocidad y aceleración. Esta última medición solo puede muestrearse utilizando un filtro previo a tomar muestras.

La identificación se realizó a través del método de Máxima Verosimilitud, empleando un algoritmo de optimización local y otro global. La función de verosimilitud presenta varios máximos locales como se observó en los ejemplos numéricos analizados. Al calcular la función de verosimilitud por descomposición del error de predicción a través del filtro de Kalman los parámetros se estimaron localmente utilizando técnicas de cuasi-Newton. Los resultados aplicando la optimización local indican que la estimación de los parámetros depende del número de muestras que se consideren, ya que hay una fuerte

variación de máximos locales y ancho de los lóbulos de la función de verosimilitud para distintos números de mediciones. A través de las gráficas de la función de verosimilitud se observa que la inicialización del algoritmo es muy importante para la estimación de los parámetros.

Por otra parte, en la optimización global se calcula la función log-verosimilitud utilizando el modelo ARMA directamente. La función de log-verosimilitud al depender de las raíces de los polinomios $R_1(z)$ y $R_2(z)$, el problema de optimización se debe realizar numéricamente. Los ejemplos numéricos muestran que la función log-verosimilitud presenta varios máximos locales. Sorprendentemente, la presencia de máximos locales ocurre incluso en el caso donde las muestras son tomadas utilizando un AAF. Luego de analizar varios ejemplos numéricos relevantes para esta investigación, queda en evidencia las dificultades para identificar los parámetros de un oscilador en tiempo continuo.

6.2. Trabajo Futuro

El problema teórico resuelto en esta tesis tan solo contempló modelar el sistema de oscilación sin condiciones iniciales y sin amortiguamiento, y aplicando únicamente el método de Máxima Verosimilitud para identificar los parámetros del sistema en cuestión. Por ende, se genera una serie de problemas a seguir investigando para identificar en la práctica los modelos de vibración:

- Incluir en el modelo a identificar las condiciones iniciales. Un análisis preliminar de este nuevo modelo se presenta en el artículo de conferencia [45].
- Modelar el sistema de oscilación con amortiguamiento.
- Analizar la función de verosimilitud al ocupar el algoritmo EM enfocado a sistemas con datos cuantizados. Un trabajo preliminar se presenta en el artículo de conferencia [48].
- Resolver el problema de estimación planteando el modelo en el dominio de la frecuencia.

BIBLIOGRAPHY

- [1] M. Glück, J. Pott, and O. Sawodny, “Investigations of an accelerometer-based disturbance feedforward control for vibration suppression in adaptive optics of large telescopes,” *Publications of the Astronomical Society of the Pacific*, vol. 129, no. 976, p. 065001, 2017.
- [2] S. Zúñiga, J. Garcés, L. Close, J. Males, K. Morzinski, P. Escárate, and M. Castro, “Vibrations in magao: analysis of on-sky data and future challenges in vibrations mitigation,” *Imaging and Applied Optics*, 2015.
- [3] P. Escárate, R. Carvajal, L. Close, J. Males, K. Morzinski, and J. C. Agüero, “Minimum variance control for mitigation of vibrations in adaptive optics systems,” *Applied Optics*, vol. 56, no. 19, pp. 5388–5397, Jul 2017.
- [4] C. Petit, J. Conan, C. Kulcsar, H. Raynaud, and T. Fusco, “First laboratory validation of vibration filtering with lqg control law for adaptive optics,” *Opt. Express*, vol. 16, no. 1, pp. 87–97, Jan 2008.
- [5] M. Castro, P. Escárate, J. Garcés, S. Zúñiga, D. Rojas, J. Marchioni, and A. Guesalaga, “Closed-loop control for tip-tilt compensation on systems under vibration,” *Proc.SPIE*, vol. 9909, pp. 9909 – 9909 – 6, 2016.
- [6] G. Sivo, C. Kulcsár, J.-M. Conan, H.-F. Raynaud, Éric Gendron, A. Basden, F. Vidal, T. Morris, S. Meimon, C. Petit, D. Gratadour, O. Martin, Z. Hubert, A. Sevin, D. Perret, F. Chemla, G. Rousset, N. Dipper, G. Talbot, E. Younger, R. Myers, D. Henry, S. Todd, D. Atkinson, C. Dickson, and A. Longmore, “First on-sky scao validation of full lqg control with vibration mitigation on the canary pathfinder,” *Opt. Express*, vol. 22, no. 19, pp. 23 565–23 591, Sep 2014.
- [7] D. Tsetserukou, N. Kawakami, and S. Tachi, “Vibration damping control of robot arm intended for service application in human environment,” *Humanoids 2008 - 8th IEEE-RAS International Conference on Humanoid Robots*, Dec 2008.

-
- [8] T. L. Hayward, M. Rippha, H. Bonnet, C. Cavedoni, R. Galvez, G. Gausachs, and M. Cho, “Characterizing the vibration environments of the gemini telescopes,” *Proceedings of the SPIE*, p. 99065N, 07 2016.
- [9] E. J. Hannan, “The estimation of frequency,” *Journal of Applied Probability*, vol. 10, no. 3, pp. 510–519, 1973.
- [10] J. Roberts and P. Spanos, “Stochastic averaging: An approximate method of solving random vibration problems,” *International Journal of Non-Linear Mechanics*, vol. 21, no. 2, pp. 111 – 134, 1986.
- [11] G. Kubin, C. Lainscsek, and E. Rank, “Identification of nonlinear oscillator models for speech analysis and synthesis,” in *Nonlinear Speech Modeling and Applications: Advanced Lectures and Revised Selected Papers*, G. Chollet, A. Esposito, M. Faundez-Zanuy, and M. Marinaro, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 74–113.
- [12] E. Boujo and N. Noiray, “Robust identification of harmonic oscillator parameters using the adjoint Fokker–Planck equation,” *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 473, no. 2200, 2017.
- [13] B. Wahlberg, “Limit results for sampled systems,” *International Journal of Control*, vol. 48, no. 3, pp. 1267–1283, 1988.
- [14] W. Aström, *Computer Controlled Systems*. Prentice Hall, 1996.
- [15] L. Ljung and A. Wills, “Issues in sampling and estimating continuous-time models with stochastic disturbances,” *Automatica*, vol. 46, no. 5, pp. 925 – 931, 2010.
- [16] G. C. Goodwin, J. C. Agüero, M. E. Cea-Garrido, M. E. Salgado, and J. I. Yuz, “Sampling and sampled-data models: The interface between the continuous world and digital algorithms,” *IEEE Control Systems*, vol. 33, no. 5, pp. 34–53, Oct 2013.
- [17] K. Åström, P. Hagander, and J. Sternby, “Zeros of sampled systems,” *Automatica*, vol. 20, no. 1, pp. 31 – 38, 1984.
- [18] M. J. Blachuta, “On approximate pulse transfer functions,” *IEEE Transactions on Automatic Control*, vol. 44, no. 11, pp. 2062–2067, Nov 1999.
- [19] S. R. Weller, W. Moran, B. Ninness, and A. D. Pollington, “Sampling zeros and the Euler-Frobenius polynomials,” *IEEE Transactions on Automatic Control*, vol. 46, no. 2, pp. 340–343, Feb 2001.

-
- [20] D. Nesic and D. S. Laila, “A note on input-to-state stabilization for nonlinear sampled-data systems,” *IEEE Transactions on Automatic Control*, vol. 47, no. 7, pp. 1153–1158, Jul 2002.
- [21] D. S. Carrasco, G. C. Goodwin, and J. I. Yuz, “Sampled data models for nonlinear stochastic systems: Truncation errors and sampling zero dynamics,” *IEEE Transactions on Automatic Control*, vol. 61, no. 1, pp. 103–115, Jan 2016.
- [22] J. I. Yuz, J. Alfaro, J. C. Agüero, and G. C. Goodwin, “Identification of continuous-time state-space models from non-uniform fast-sampled data,” *IET Control Theory Applications*, vol. 5, no. 7, pp. 842–855, May 2011.
- [23] E. K. Larsson and T. Söderström, “Continuous-time AR parameter estimation by using properties of sampled systems,” *IFAC Proceedings Volumes*, vol. 35, no. 1, pp. 103 – 108, 2002, 5th IFAC World Congress.
- [24] E. K. Larsson, M. Mossberg, and T. Söderström, “Identification of continuous-time ARX models from irregularly sampled data,” *IEEE Transactions on Automatic Control*, vol. 52, no. 3, pp. 417–427, March 2007.
- [25] H. Kirshner, S. Maggio, and M. Unser, “A sampling theory approach for continuous arma identification,” *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4620–4634, Oct 2011.
- [26] P. Whittle, “The analysis of multiple stationary time series,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 15, no. 1, pp. 125–139, 1953.
- [27] F. Chen, J. C. Agüero, M. Gilson, H. Garnier, and T. Liu, “EM-based identification of continuous-time ARMA models from irregularly sampled data,” *Automatica*, vol. 77, no. 3, pp. 293 – 301, 2017.
- [28] G. McLachlan and T. Krishnan, *The EM algorithm and extensions*, 1st ed., ser. Wiley series in probability and statistics. Applied probability and statistics. Wiley, 1997.
- [29] R. Fletcher, *Practical methods of optimization*, 1st ed. J. Wiley, 1980, vol. 1.
- [30] T. Söderström, *Discrete-time Stochastic Systems: Estimation and Control*, 2nd ed., ser. Advanced Textbooks in Control and Signal Processing. Springer-Verlag London, 2002.
- [31] L. Ljung, *System Identification: Theory for the User*, 2nd ed., ser. 2nd Edition. Prentice Hall, 1999.

-
- [32] R. Tyson, *Adaptive Optics Engineering Handbook*, 1st ed., ser. Optical engineering 67. Marcel Dekker, 2000.
- [33] A. J. Wright and S. P. Poland, *Adaptive Optics for Aberration Correction in Optical Microscopy*. Springer Netherlands, 2017, pp. 585–612.
- [34] I. Kozak, “Retinal imaging using adaptive optics technology,” *Saudi Journal of Ophthalmology*, 02 2014.
- [35] J. Schmidt, *Numerical Simulation of Optical Wave Propagation With Examples in MATLAB*, pap/chrt ed., 2010.
- [36] R. Tyson, *Adaptive Optics Engineering Handbook*, 1st ed., ser. Optical engineering 67. M. Dekker, 2000.
- [37] S. Meimon, C. Petit, T. Fusco, and C. Kulcsar, “Tip-tilt disturbance model identification for kalman-based control scheme: application to xao and elt systems,” *J. Opt. Soc. Am. A*, vol. 27, no. 11, pp. A122–A132, Nov 2010.
- [38] C. Kulcsár, H. Raynaud, C. Petit, J. Conan, and P. Lesegno, “Optimal control, observers and integrators in adaptive optics,” *Opt. Express*, vol. 14, no. 17, pp. 7464–7476, Aug 2006.
- [39] J.-L. Lizon, G. Jakob, B. De Marneffe, and A. Preumont, “Different ways of reducing vibrations induced by cryogenic instruments,” 07 2010.
- [40] J. Garcés, S. Z. niga, L. Close, J. Males, K. Morzinski, P. Escárate, M. Castro, J. Marchioni, and D. Rojas, “Vibrations in magao: resonance sources identification and first approaches for modeling and control,” *Proc.SPIE*, vol. 9909, pp. 9909 – 9909 – 17, 2016.
- [41] C. Chatfield, *The analysis of time series: an introduction*, 5th ed. Chapman and Hall/CRC, 1996.
- [42] C. F. Van Loan, “Computing integrals involving the matrix exponential,” *IEEE Transactions on Automatic Control*, vol. 23, no. 3, pp. 395–404, Jun 1978.
- [43] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas: Second Edition*, 2nd ed. Princeton University Press, 2009.
- [44] K. González, P. Escárate, R. Carvajal, and J. Agüero, “On maximum likelihood estimation of continuous time oscillators,” *Submitted in IEEE Latin America Transactions*, 2017.

-
- [45] P. Escárate, K. González, M. Coronel, R. Carvajal, and J. Agüero, “Vibration model identification using the maximum likelihood method,” *To appear in SPIE Astronomical Telescopes + Instrumentation*, 2018.
- [46] J. C. Agüero, W. Tang, J. I. Yuz, R. Delgado, and G. C. Goodwin, “Dual time-frequency domain system identification,” *Automatica*, vol. 48, no. 12, pp. 3031–3041, Dec. 2012.
- [47] J. Barzilai and J. Borwein, “Two-point step size gradient methods,” *IMA Journal of Numerical Analysis*, vol. 8, no. 1, pp. 141–148, 1988.
- [48] J. Agüero, K. González, and R. Carvajal, “Em-based identification of arx systems having quantized output data,” *In 20th IFAC World Congress*, vol. 50, pp. 8367 – 8372, 2017.
- [49] E. Putzer, “Avoiding the jordan canonical form in the discussion of linear systems with constant coefficients,” *The American Mathematical Monthly*, vol. 73, no. 1, pp. 2–7, jan 1966.
- [50] W. A. Adkins and M. G. Davidson, “Putzer’s algorithm for eat via the laplace transform,” *Mathematics Magazine*, vol. 83, no. 4, pp. 267–275, 2010.

APÉNDICE A

A.1. Cálculo de las matrices $e^{A\Delta}$ y Q_d del modelo en tiempo-discreto.

Considere el sistema en variable de estado:

$$\dot{x}(t) = Ax(t) + \dot{w}(t),$$

con $\mathbb{E}\{\dot{w}(t)\dot{w}(\tau)^T\} = Q_c\delta(t - \tau)$. El sistema muestreado con tiempo de muestreo Δ , está dado por:

$$x(t_{k+1}) = e^{A\Delta}x(t_k) + n(t_k),$$

donde $e^{A\Delta}$ y la covarianza de $n(t_k)$ se obtienen con el método presentado en [42] y [15]. El método realiza el cálculo en forma eficiente y compacta en el cálculo de integrales que incluyen la matriz exponencial. A partir de la siguiente matriz exponencial extendida

$$\exp\left\{\begin{bmatrix} -A & Q_c \\ 0 & A^T \end{bmatrix} \Delta\right\} = \begin{bmatrix} P_{11} & P_{12} \\ 0 & P_{22} \end{bmatrix}$$

se puede obtener simultáneamente tanto la matriz $e^{A\Delta}$ como la varianza del ruido $n(t_k)$ del modelo en tiempo discreto (14) y (17).

$$\begin{aligned} e^{A\Delta} &= P_{22}^T \\ Q_d &= \int_0^\Delta e^{A^T\eta} d\eta = P_{22}^T P_{12}. \end{aligned} \tag{A.1}$$

A.2. Cálculo de derivadas necesarias para obtener el gradiente de la función de máxima verosimilitud.

- Derivadas de la etapa de predicción del Filtro de Kalman respecto a los parámetros a estimar

$$\begin{aligned}
\frac{\partial K_k}{\partial \theta_j} &= \frac{\partial \Sigma_{k|k-1}}{\partial \theta_j} C^T (C \Sigma_{k|k-1} C^T)^{-1} - \Sigma_{k|k-1} C^T (C \Sigma_{k|k-1} C^T)^{-2} \left(C \frac{\partial \Sigma_{k|k-1}}{\partial \theta_j} C^T \right) \\
\frac{\partial \hat{x}_{k|k}}{\partial \theta_j} &= \frac{\partial \hat{x}_{k|k-1}}{\partial \theta_j} + \frac{\partial K_k}{\partial \theta_j} (y_k - C \hat{x}_{k|k-1}) - K_k C \frac{\partial \hat{x}_{k|k-1}}{\partial \theta_j} \\
\frac{\partial \Sigma_{k|k}}{\partial \theta_j} &= (I - K_k C) \frac{\partial \Sigma_{k|k-1}}{\partial \theta_j} - \frac{\partial K_k}{\partial \theta_j} C \Sigma_{k|k-1}
\end{aligned} \tag{A.2}$$

- Derivadas de la etapa de corrección del Filtro de Kalman respecto a los parámetros a estimar

$$\begin{aligned}
\frac{\partial \hat{x}_{k+1|k}}{\partial \alpha_i} &= \frac{\partial e^{A\Delta}}{\partial \alpha_i} \hat{x}_{k|k} + e^{A\Delta} \frac{\partial \hat{x}_{k|k}}{\partial \alpha_i} \\
\frac{\partial \Sigma_{k+1|k}}{\partial \alpha_i} &= \frac{\partial e^{A\Delta}}{\partial \alpha_i} \Sigma_{k|k} e^{A^T \Delta} + e^{A\Delta} \frac{\partial \Sigma_{k|k}}{\partial \alpha_i} e^{A^T \Delta} + e^{A\Delta} \Sigma_{k|k} \frac{\partial e^{A^T \Delta}}{\partial \alpha_i} + \frac{\partial Q_d}{\partial \alpha_i} \\
\frac{\partial \hat{x}_{k+1|k}}{\partial \beta_i} &= e^{A\Delta} \frac{\partial \hat{x}_{k|k}}{\partial \beta_i} \\
\frac{\partial \Sigma_{k+1|k}}{\partial \beta_i} &= e^{A\Delta} \frac{\partial \Sigma_{k|k}}{\partial \beta_i} e^{A^T \Delta} + \frac{\partial Q_d}{\partial \beta_i}
\end{aligned} \tag{A.3}$$

- Derivadas de la matriz de varianza Q respecto a los parámetros a estimar

$$\begin{aligned}
\frac{\partial Q_d}{\partial \alpha_i} &= \int_0^\Delta \left(\frac{\partial e^{A\eta}}{\partial \alpha_i} Q_c e^{A^T \eta} + e^{A\eta} Q_c \frac{\partial e^{A^T \eta}}{\partial \alpha_i} \right) d\eta \\
\frac{\partial Q_d}{\partial \beta_i} &= \int_0^\Delta e^{A\eta} \text{frac} \partial Q_c \partial \beta_i e^{A^T \eta} d\eta
\end{aligned} \tag{A.4}$$

- Derivadas de la matriz $e^{A\Delta}$ respecto al parámetro α_i

$$\frac{\partial e^{A\Delta}}{\partial \alpha_i} = \Delta e^{A\Delta} \frac{\partial A}{\partial \alpha_i} \tag{A.5}$$

A.3. Procedimiento en Matlab

Para el desarrollo de la identificación del vector de parámetros en (13) se ha seguido el siguiente procedimiento en Matlab.

1. Obtener la frecuencia correspondiente al máximo de la PSD a partir del cálculo de la transformada rápida de Fourier de los datos.
2. Definir condiciones iniciales de los parámetros $\hat{\theta}^0$.
3. Se define la función *MLgradiente* el cual entrega el valor de la función de verosimilitud y el vector gradiente, para ello se debe seguir los siguientes pasos:

- 3.1. Obtener las matrices del modelo en espacio de estado (4.7) y (4.8).
 - 3.2. Calcular las matrices $e^{A\Delta}$ y Q_d del modelo en espacio de estado en tiempo discreto (4.13).
 - 3.3. Calcular las matrices del filtro de Kalman $K_k, \hat{x}_{k|k}, \Sigma_{k|k}, \hat{x}_{k+1|k}, \Sigma_{k+1|k}$ (5.1)-(5.5).
 - 3.4. Calcular el error de predicción ε_k y su varianza Λ_k en (5.6) y (5.7).
 - 3.5. Calcular el valor de la función log-verosimilitud $l_N(\theta)$ en (4.42).
 - 3.6. Calcular las derivadas del filtro de Kalman, ε_k y Λ_k respecto al vector de parámetros en (5.10) y (5.11).
 - 3.7. Calcular el vector de gradiente $\nabla l(\theta)_{\theta=\hat{\theta}}$ en (5.9).
4. Optimizar ecuación (4.41), utilizando comando `fminunc`.

Código Matlab:

```
options=optimoptions(@fminunc,'Algorithm','quasi-Newton','GradObj','on');
```

```
 $\hat{\theta} = fminunc(@()MLgradiente, \theta_{inicial}, options);$ 
```

APÉNDICE B

B.1. Lemas

Lema 4. Sea A_1 , A_2 y A_3 las siguientes matrices:

$$A_1 = \begin{bmatrix} 0 & 1 \\ -\alpha^2 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ -\alpha^2 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 1 & 0 \\ -\alpha^2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad (\text{B.1})$$

entonces la matriz exponencial de $A_1\Delta$, $A_2\Delta$ y $A_3\Delta$ son dados por:

$$e^{A_1\Delta} = \begin{bmatrix} \cos[\alpha\Delta] & \frac{1}{\alpha} \sin[\alpha\Delta] \\ -\alpha \sin[\alpha\Delta] & \cos[\alpha\Delta] \end{bmatrix} \quad (\text{B.2})$$

$$e^{A_2\Delta} = \begin{bmatrix} \cos[\alpha\Delta] & \frac{1}{\alpha} \sin[\alpha\Delta] & 0 \\ -\alpha \sin[\alpha\Delta] & \cos[\alpha\Delta] & 0 \\ \frac{1}{\alpha} \sin[\alpha\Delta] & \frac{1}{\alpha^2} (\cos[\alpha\Delta] - 1) & 1 \end{bmatrix} \quad (\text{B.3})$$

$$e^{A_3\Delta} = \begin{bmatrix} \cos[\alpha\Delta] & \frac{1}{\alpha} \sin[\alpha\Delta] & 0 \\ -\alpha \sin[\alpha\Delta] & \cos[\alpha\Delta] & 0 \\ \cos[\alpha\Delta] - 1 & \frac{1}{\alpha} \sin[\alpha\Delta] & 1 \end{bmatrix} \quad (\text{B.4})$$

Demostración. Directamente utilizando el modificado algoritmo de Putzer (para más detalles ver Apéndice B.2). □

A continuación se presenta la varianza para el equivalente ruido de tiempo discreto utilizando muestreo instantáneo e integrado en (4.13) y (3.21) respectivamente.

Lema 5. La varianza de $n(t_{k+1})$ para los sistemas de medición de posición y velocidad muestreados instantáneamente está dado por:

$$Q_d = \sigma^2 \begin{bmatrix} \frac{2\alpha\Delta - \sin[2\alpha\Delta]}{4\alpha^3} & \frac{\sin^2[\alpha\Delta]}{2\alpha^2} \\ \frac{\sin^2[\alpha\Delta]}{2\alpha^2} & \frac{2\alpha\Delta + \sin[2\alpha\Delta]}{4\alpha} \end{bmatrix} \quad (\text{B.5})$$

donde α es la frecuencia del oscilador y Δ corresponde al tiempo de muestreo.

Demostración. Utilizando la definición de varianza y el Lema 4 se obtiene que :

$$\begin{aligned} Q_d &= \mathbb{E}\{n(t_{k+1})n^T(t_{k+1})\} \\ &= \sigma^2 \int_0^\Delta (e^{A\eta\kappa}) (e^{A\eta\kappa})^T d\eta \\ &= \sigma^2 \int_0^\Delta \begin{bmatrix} \frac{1}{\alpha} \sin[\alpha\eta] \\ \cos[\alpha\eta] \end{bmatrix} \begin{bmatrix} \frac{1}{\alpha} \sin[\alpha\eta] & \cos[\alpha\eta] \end{bmatrix} d\eta \\ &= \sigma^2 \int_0^\Delta \begin{bmatrix} \frac{1}{\alpha^2} \sin^2[\alpha\eta] & \frac{1}{\alpha} \sin[\alpha\eta] \cos[\alpha\eta] \\ \frac{1}{\alpha} \cos[\alpha\eta] \sin[\alpha\eta] & \cos^2[\alpha\eta] \end{bmatrix} d\eta \\ &= \sigma^2 \begin{bmatrix} \frac{2\alpha\Delta - \sin[2\alpha\Delta]}{4\alpha^3} & \frac{\sin^2[\alpha\Delta]}{2\alpha^2} \\ \frac{\sin^2[\alpha\Delta]}{2\alpha^2} & \frac{2\alpha\Delta + \sin[2\alpha\Delta]}{4\alpha} \end{bmatrix} \end{aligned} \quad (\text{B.6})$$

□

Lema 6. La varianza $n(t_{k+1})$ para los sistema de medición de posición, velocidad y aceleración utilizando muestreo integrado están dados por:

$$\text{cov} \left\{ \begin{bmatrix} w_d(t_{k+1}) \\ v_d(t_{k+1}) \end{bmatrix} \right\} = Q = \begin{bmatrix} Q_d(\alpha, \sigma^2) & S_d(\alpha, \sigma^2) \\ S_d^T(\alpha, \sigma^2) & R_d(\alpha, \sigma^2) \end{bmatrix} \quad (\text{B.7})$$

donde

- Medición de posición

$$\begin{aligned}
 Q_d &= \sigma^2 \begin{bmatrix} \frac{2\alpha\Delta - \sin[2\alpha\Delta]}{4\alpha^3} & \frac{\sin[\alpha\Delta]^2}{2\alpha^2} \\ \frac{\sin[\alpha\Delta]^2}{2\alpha^2} & \frac{2\alpha\Delta + \sin[2\alpha\Delta]}{4\alpha} \end{bmatrix} \\
 R_d &= \sigma^2 \cdot \frac{6\alpha\Delta - 8\sin[\alpha\Delta] + \sin[2\alpha\Delta]}{4\alpha^5} \\
 S_d &= \sigma^2 \begin{bmatrix} \frac{2\sin[\frac{\alpha\Delta}{2}]^4}{\alpha^4} \\ \frac{4\sin[\alpha\Delta] - 2\alpha\Delta - \sin[2\alpha\Delta]}{4\alpha^3} \end{bmatrix}
 \end{aligned} \tag{B.8}$$

- Medición de velocidad

$$\begin{aligned}
 Q_d &= \sigma^2 \begin{bmatrix} \frac{2\alpha\Delta - \sin[2\alpha\Delta]}{4\alpha^3} & \frac{\sin^2[\alpha\Delta]}{2\alpha^2} \\ \frac{\sin^2[\alpha\Delta]}{2\alpha^2} & \frac{2\alpha\Delta + \sin[2\alpha\Delta]}{4\alpha} \end{bmatrix} \\
 R_d &= \sigma^2 \cdot \frac{2\alpha\Delta - \sin[2\alpha\Delta]}{4\alpha^3} \\
 S_d &= \sigma^2 \begin{bmatrix} \frac{2\alpha\Delta - \sin[2\alpha\Delta]}{4\alpha^3} \\ \frac{\sin^2[\alpha\Delta]}{2\alpha^2} \end{bmatrix}
 \end{aligned} \tag{B.9}$$

- Medición de aceleración

$$\begin{aligned}
 Q_d &= \sigma^2 \begin{bmatrix} \frac{\alpha}{4} (2\alpha\Delta - \sin[2\alpha\Delta]) & \frac{\alpha^2}{2} \sin^2[\alpha\Delta] \\ \frac{\alpha^2}{2} \sin^2[\alpha\Delta] & \frac{\alpha^3}{4} (2\alpha\Delta + \sin[2\alpha\Delta]) \end{bmatrix} \\
 R_d &= \sigma^2 \cdot \frac{2\alpha\Delta + \sin[2\alpha\Delta]}{4\alpha} \\
 S_d &= \sigma^2 \begin{bmatrix} \frac{1}{2} \sin^2[\alpha\Delta] \\ \frac{-\alpha}{4} (2\alpha\Delta + \sin[2\alpha\Delta]) \end{bmatrix}
 \end{aligned} \tag{B.10}$$

Demostración. Directamente ocupando la definición de varianza y el Lema 4 como en la demostración del Lema 5. \square

B.2. Algoritmo de Putzer Modificado

La matriz exponencial se puede calcular usando la siguiente fórmula

$$e^{At} = \sum_{j=0}^{n-1} r_{j+1} P_j \quad (\text{B.11})$$

donde las matrices P_j están dadas por:

$$P_0 = I \quad ; \quad P_j = \prod_{k=1}^j (A - \lambda_k I) \quad j = 1, \dots, n \quad (\text{B.12})$$

y

$$r_j(t) = \mathcal{L}^{-1} \left\{ \frac{1}{(s - \lambda_j) \dots (s - \lambda_1)} \right\} \quad (\text{B.13})$$

donde λ_j son los autovalores de A .

El algoritmo se puede resumir de la siguiente manera;

- i) Determinar los autovalores de A .
- ii) Calcular las matrices P_j .
- iii) Calcular los coeficientes r_j .
- iv) Obtener e^{At} .

Demostración. Ver [49] y [50].

□