Universidad de La Laguna

Facultad de Ciencias
Sección de Física

TRABAJO DE FIN DE GRADO

Bioacústica física:

La parametrización de las ecuaciones para simular el canto de los pájaros.

Pablo Rodríguez Beltrán

Supervisor:

Prof. Fernando Luis Rosa González

Índice general

1. Introducción **1**

1.1. Objetivos 2

1.2. Estructura expositiva 3

2. Antecedentes y marco teórico **4**

2.1. Anatomía del órgano vocal 4

2.2. Modelo físico 5

2.2.1. Espacio de parámetros 7

2.2.2. Inicio de la oscilación 8

2.3. Síntesis con series de Fourier 10

2.4. Método de detección de eventos 12

3. Metodología desarrollada **14**

3.1. Adquisición de la señal 14

3.2. Detección de eventos 15

3.3. Estimación de la frecuencia fundamental 17

3.4. Síntesis de la señal 18

3.5. La ecuación de van der Pol 19

4. Resultados y discusión **22**

4.1. Evento utilizado 22

4.2. Señal sintetizada 23
4.3. Integración de la ecuación de van der Pol 25
4.4. Búsqueda de la frecuencia fundamental 28

5. Conclusiones y perspectivas de futuro **31**
5.1. Líneas futuras de trabajo 31
5.2. Conclusiones 33

Apéndices: **37**

A. Varios códigos del sistema **37**

B. Varios códigos del sistema **40**

Abstract

Animal vocal communication is a wide field of study with numerous potential applications. In physics, the study of acoustic zoosemiotics tries to explain the animal vocalization with models that approximate acoustic production. Birdsong is a rich example of this vocal system, capable to product simple signals or highly complex sounds. This sound production variety sets birdsongs as a good starting point to comprehend animal sound outputs and the physic simulations of them.

This work presents the study of the equations for the production and analysis of birdsongs, starting from the researches of Gabriel Mindlin from Buenos Aires University. The main target will be to simulate the signal emitted by passeriformes birds, parametrizing the differential equations and integrating them, in order to establish the validity of the proposed models.

The analysis of the acoustic signal is a useful feature to study the acoustical behaviour of animal species and, therefore, automating this work, from the very begging of the sound recording, is a top request for the conservation of animal populations.

Furthermore, understanding birdsong production, which is simpler, allows to find the keys of more complex vocalizations from other animals, or even, comprehending the learning process of oral language in humans.

The work begins with the review of the state of the art of this field and the theoretic notions studied, necessary to carry out the following chapters. Several publications of the Gabriel Mindlin research team on the sound production of birds will be studied, for the purpose of selecting a model to simulate the biological vocalization system, in a way as simpler as possible. The chosen model equation is the van der Pol oscillator, that represents the source of the sound as a mass attached to an oscillator, where the pressure, the elastic constant (of the bird tract tissue) and a non linear friction are considered. It is expected to find a solution able to simulate the birdsong signal if the correct parameters are introduced in the equation. The study of this physical model is one of the main

missions of this work, because it can provide some clues to understand the nature of animal vocalizations from a physical point of view.

Otherwise, it is studied a methodology to recognize sound events using energy detection. This is necessary to isolate each vocalization and develop a system able to extract the fundamental frequency, the maximum of the spectral density and its phase. With these parameters it is possible to construct a synthesized signal that reproduce the original data. This synthesized signal offers the advantage to work with the refined data, that allows to study and manipulate the essential information of each event and it is applied to build a parameter space. With this parameter space it is possible follow the state and flux of the vocalization, that may bring forward the possibility of classifying each birdsong event. Also, the parameter space will be used to obtain the optimum coefficients for the van der Pol equations, adjusting the solution of the differential equation to the synthesized signal.

The development of the work requires a proper sound file to apply and validate the following methods. To analyse this file (or any other sound file), a sound processing system (SPS) has been developed, in first place, to detect the vocalizations as sound events with energy over the noise threshold. Thus, the SPS isolate the events and search automatically the maximum of the spectral density at each time of the event, which provides the fundamental frequency. However, the method used to find the fundamental frequency shows some unwanted jumps in the instantaneous frequency. Using the parameters, SPS can synthesize a sound signal that model the original. SPS can be used to look for the optimum parameters for a given event in order to integrate the van der Pol equation and find the best-fitting solution.

After the develop and execution of the SPS, the results are displayed and explained. The synthesized signal fits approximately the original data in amplitude but there are some differences in the phase. These differences appear because the raw data has noise, coupled oscillations and very fast changes in frequency. The synthesized signal rounds

up all these irregularities, and develops a signal with smoother oscillations, so it can be used for next approximations.

The parameters obtained from the data follow the breathing cycles of the bird and the results are similar to those described in the literature. Finally, the results of the van der Pol integration are shown and its solutions fit the signal, but there are some slight differences in amplitude and phase. This must come from the lack of numerical precision in the approximations or from the method of integration. Actually, the van der Pol model itself is limited to reproduce the birdsong. For example, the equation can not generate the second harmonic frequency, so there is a loss of energy in the simulation.

The work ends listing the future potential studies and the conclusions. The SPS could be extended in order to implement in it the recording process and to analyse the sound automatically. It is possible to improve the event detector, adding a noise detector and creating a tracking system to find the fundamental frequencies that avoids unwanted jumps. Also, the relations between the parameters of the van der Pol equation require a deeper study.

One of the biggest potentials of this work is the possibility to classify automatically birdsongs and species with the data obtained from each sound event, because it will be a great help for the monitoring of animal species and their conservation.

In conclusion, the SPS works properly but could be improved in several ways: the detection of sound events and its analysis works successfully, the parameters obtained are valid to follow the flux of birdsong cycles and the synthesized signal rounds well the original. As for the simulation of the birdsong using the van der Pol equation, as suggested in the studies of Gabriel Mindlin, the solutions present a valid first approximation to the problem.

Resumen

La comunicación sonora animal es un campo de estudio amplio y con numerosas aplicaciones potenciales. Desde la perspectiva de la física, el estudio de la zoosemiótica acústica trata de explicar la vocalización del animal con modelos que aproximen la producción acústica. El canto de un pájaro es un ejemplo de la riqueza del sistema vocal, capaz de generar señales simples o sonidos mucho más complejos. Esta variedad en la producción hace que su estudio sea un buen punto de partida para la compresión general de la producción sonora animal y los modelos físicos que la pueden simular.

Este trabajo presenta el estudio de las ecuaciones para la producción y análisis de cantos de pájaros, partiendo de las investigaciones de Gabriel Mindlin de la Universidad de Buenos Aires. El objetivo principal será emular la señal emitida por paseriformes, parametrizando las ecuaciones diferenciales e integrándolas por medio de métodos numéricos, con el fin de validar experimentalmente los modelos propuestos.

La obtención de parámetros y el análisis de la producción de la señal son características de gran utilidad para realizar el seguimiento de especies según su comportamiento acústico y, por tanto, la automatización de estas tareas a partir de la grabación de estos sonidos resulta de máximo interés para su aplicación en la conservación de estas poblaciones. Además, entender el problema más sencillo del canto de las aves permite buscar las claves para comprender sistemas de producción más complejos de otros animales, o incluso el proceso de aprendizaje de comunicación de los humanos.

El trabajo comienza con la exposición de los conceptos teóricos estudiados que van

a ser utilizados durante los siguientes apartados. Se analizan diversas publicaciones del equipo de investigación de Mindlin sobre la producción sonora de pájaros, y se escoge el modelo que propone la ecuación de van der Pol para simular el sistema biológico de vocalización, ya que se espera que su solución devuelva una señal como la emitida por un ave, si se introducen los parámetros adecuados. Por otra parte, se estudia la metodología de detección de eventos de sonido para aislar cada vocalización y se desarrolla un sistema de síntesis de señales a partir de su frecuencia fundamental, su densidad espectral máxima y su desfase. Además, se estudia el espacio de parámetros de la señal, que sirve para calcular los coeficientes de la ecuación de van der Pol, que aportan información del estado y flujo del canto y adelantan la posibilidad de clasificar cada vocalización.

Para desarrollar el trabajo se construye un sistema capaz de detectar eventos de sonido, de esta forma se aíslan aquellos datos con señal de canto, para poder extraer información de ellos y analizarlos. De cada evento se obtienen sus parámetros esenciales y se realiza una búsqueda de la frecuencia fundamental. Después, con los parámetros extraídos se sintetiza una señal que reproduce los datos originales y permite un estudio numérico más eficaz de cada evento. Además, se utiliza la señal sintetizada y sus parámetros como valores a incluir en la ecuación de van der Pol, estudiando las relaciones que guardan entre sí y los resultados de la integración numérica. De esta forma es posible extraer conclusiones sobre la simulación propuesta por G. Mindlin y sobre las características de canto del pájaro.

Capítulo 1

Introducción

Abstract

Birdsongs contain some of the most complex and beautiful sounds in nature, and physics allow to understand how the song is generated and what parameters are involved in the process. Obtaining these resources may improve the classification and conservation of bird species, or even could set references to expand the work to the sound production of other animals. The project has a few goals to achieve: creating a system able to detect sound events, search the fundamental frequency and other parameters, synthesize the signal with these parameters and study the solutions of the equations that should reproduce the sound.

ooo

La naturaleza ofrece una amplia gama de sonidos que pueden ser estudiados, pero de todos ellos son de particular interés y belleza aquellos producidos por los miembros del reino animal. El análisis y clasificación de la vocalización de los animales es un elemento fundamental para ayudar a su conservación, censo y seguimiento. Además, comprender el conjunto de mecanismos que permiten la comunicación sonora permitiría entender las relaciones que mantienen entre ellos y con su entorno, así como los métodos que utilizan para aprender a emitir dichos sonidos. Por otra parte, el estudio de la zoosemiótica de aquellos animales cuya producción sonora sea más sencilla que la del ser humano, puede ayudar a establecer pautas para esclarecer los medios de aprendizaje y el funcionamiento del habla humana por su similitud [2, 7].

Dado que la comunicación estudiada es sonora, será posible cuantificar, medir y analizar las oscilaciones de las ondas siguiendo los principios físicos que sean necesarios. De estas ondas se pretende obtener diversos parámetros que, una vez conocidos, permitirán sintetizar el sonido fuente, reproduciendo las características originales del canto del animal.

Por su menor complejidad en comparación a otros animales, se estudiarán los datos de aves paseriformes (pájaros cantores), buscando ejemplares sencillos con el objetivo de trabajar con los parámetros de su sonido y comprobar los modelos que se han escogido para explicar su canto.

El entorno de trabajo de este proyecto está estrechamente relacionado con el convenio existente entre la Fundación Loro Parque y la Universidad de La Laguna, que viene realizando el estudio sistemático, tanto del comportamiento vocal como de los mecanismos de producción de sonido de especies marinas, y de cómo el entorno influye en la comunicación de estas especies. Además, la universidad realiza el diseño y fabricación del *hardware* y el *software* necesarios para la adquisición, localización, identificación y clasificación automática de tales sonidos.

Es precisamente la búsqueda de métodos de clasificación más eficientes lo que ha motivado que se inicie un estudio de las ecuaciones diferenciales que controlan el proceso de producción sonora en diversos animales, que en este trabajo se centrará en pájaros cantores.

1.1. Objetivos

El objetivo final es elaborar un método automático de detección, análisis y síntesis del canto del ave, que extraiga información sobre los parámetros de cada evento vocal.

Los estudios consultados en este trabajo han planteado como describir de forma analítica la producción sonora del canto de las aves. Cabe destacar los estudios del

investigador Gabriel Mindlin, que propone ajustar la señal producida por el pájaro con un oscilador de van de Pol, tal como se explicará en el Capítulo 2. Uno de los propósitos de este trabajo es comprobar la veracidad del método propuesto por Mindlin, para ello se estimarán, a partir de datos experimentales reales, los valores de algunos parámetros, para introducir dichos valores en las ecuaciones y recuperar una aproximación de la señal original. De este modo se estudiarán los parámetros involucrados y se comprobará si la forma matemática propuesta consigue simular el canto. En resumen los objetivos serán:

- La creación de un sistema que que detecte eventos de canto y los analice.
- La síntesis de una señal a partir de los parámetros extraídos del evento.
- El estudio del espacio de parámetros.
- Comprobar la validez del modelo de Mindlin para simular el canto.

1.2. Estructura expositiva

Se comienza con la exposición de los fundamentos teóricos preliminares y el estado del arte en el Capítulo 2, para asentar una serie de conceptos que servirán de apoyo y guía para el desarrollo del trabajo. En el Capítulo 3 se presenta el desarrollo de la metodología, donde se estudia la señal y se construirá un sistema para su tratamiento. Después, se muestran y discuten los resultados obtenidos en el Capítulo 4. Finalmente, en el Capítulo 5 se exponen las conclusiones, se proponen líneas futuras de investigación, ampliaciones posibles del trabajo y aplicaciones potenciales.

Capítulo 2

Antecedentes y marco teórico

Abstract

Before the work gets started it is necessary to summarize the state of the art and the birdsong models that will be used in the following chapters. First at all, it is important to describe the anatomy of the source and how it is related with the equations proposed to reproduce the vocalization. The research of this physic model is one of the main quests of this work, therefore studying the equations may provide some keys to understand the following results. Also, there is a explanation of the synthesis of a signal that will reproduce the original data with a few parameters and the description of a sound recognition system, necessary to detect the events with vocalization.

ooo

2.1. Anatomía del órgano vocal

Para poder entender con más claridad las ideas que se expresarán posteriormente y el vocabulario utilizado es necesario definir ciertos conceptos y explicar brevemente el sistema biológico de vocalización de aves paseriformes sencillas [4, 7, 11], como el que se muestra en la figura 2.1.

Las aves no poseen laringe ni cuerdas vocales, en su lugar tienen una siringe en la que el sonido se produce por vibraciones del tejido de sus paredes forzado por el aire que fluye desde los bronquios. La siringe se ubica donde la traquea se divide en dos para conectar

con los bronquios. En las dos partes de la siringe se encuentran un par de labios, que producirán los sonidos cuando el aire los atraviese. El modelo interpreta los labios como si se trataran de pequeñas masas conectadas a un resorte cuya oscilación es asociada con el sonido emitido. Además, se denominará "sílaba" a cada vocalización, cuya duración suele ser entre 0,025 y 0,25 segundos.

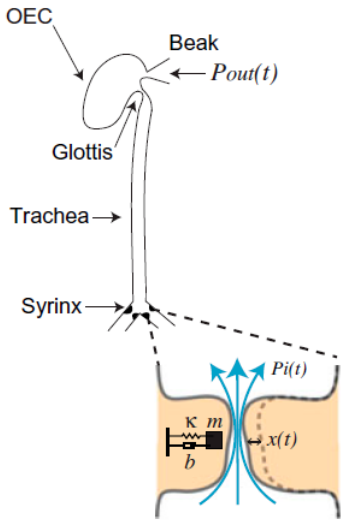

Figura 2.1: Esquema del tracto vocal, los labios de la siringe se toman como osciladores con elasticidad *(k)* y amortiguamiento *(b)* [7].

2.2. Modelo físico

Es posible describir el sistema biológico anterior con ciertos modelos matemáticos. El modelo que será utilizado como referencia fundamental en los capítulos posteriores es el propuesto por el profesor Gabriel B. Mindlin y colaboradores en el libro *The physics of birdsong* [6], teniendo en cuenta algunas otras publicaciones [1, 4, 7]. Dicho modelo emula el sistema de canto del ave con un oscilador unidimensional, ya que interpreta los labios del ave como una masa con un resorte, capaz de oscilar para comprimir el aire y producir ondas sonoras. El oscilador propuesto como una posible aproximación al

sistema físico es un oscilador de van der Pol. La expresión de la ecuación 2.1 muestra la variación diferencial del estado del sistema.

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -kx + py - \beta y - \gamma x^2 y \end{aligned} \tag{2.1}$$

La solución *x(t)* de dicha ecuación diferencial debe coincidir con la señal de canto producida por el ave. La masa no aparece explícitamente en la expresión porque se considera a los parámetros por unidad de masa.

La ecuación del oscilador de van der Pol aporta términos que permiten trabajar con amortiguamientos y no linealidad:

1. El término p será la presión efectiva, responsable de la ganancia de energía del sistema al comprimir el aire y producir la vibración. Está relacionado con la intensidad del sonido, siendo la presión proporcional al cuadrado de la amplitud [10].

2. El término β se opone a la ganancia de energía y al movimiento, explica las pérdidas del sistema en forma de rozamiento. Se puede agrupar este término con la presión, ya que ambos actúan exclusivamente sobre la velocidad y.

3. El término con γ implica una fuerza disipativa de rozamiento adicional relacionada con la posición y la velocidad. Depende de las paredes del conducto y también se puede describir como un coeficiente de reflexión [10].

4. El término k es la constante de elasticidad de los labios del pájaro, relacionada con la frecuencia de la oscilación como $k = w^2 = (2\pi\nu)^2$.

Llegado a este punto, es necesario plantear si los coeficientes del oscilador son realmente constantes, dado que el canto no es constante ni en frecuencia ni en amplitud, se puede deducir que tanto la k como la p son susceptibles de ser variables en el tiempo.

La forma matemática que siguen *k(t)* y *p(t)*, está descrita en el libro de Mindlin *The physics of birdsong* [6] en el capítulo 5, sección 2.2, donde se propone el funcionamiento

de las ecuaciones 2.2 y 2.3.

$$p(t) = p_0 + p_1 cos(\phi(t)) \tag{2.2}$$

$$k(t) = k_0 + k_1 cos(\phi(t) + \phi_0) \tag{2.3}$$

Donde p_0, k_0, p_1 y k_1 son constantes que acotan y disponen los valores de la presión y la constante elástica. mientras que $\phi(t)$ describe la velocidad de cambio y ϕ_0 da idea de la diferencia de fase. Ambas ecuaciones son vitales para construir el espacio de parámetros, que dará una idea de como se ha emitido el sonido.

Si la constante recuperadora varía con el tiempo *k=k(t)*, también lo hará la frecuencia, por tanto es necesario utilizar la frecuencia instantánea ν_i, como se verá en el apartado 2.3.

2.2.1. Espacio de parámetros

El modelo propuesto permite construir un espacio de parámetros para la constante de elasticidad k en el eje de ordenadas y la presión p en el eje de abscisas [4, 6].

La posición en el espacio de parámetros indicará el estado en que se encuentra el sistema y condicionará su solución. La evolución en ese espacio permite explicar las variaciones de esa sílaba de canto (figura 2.2).

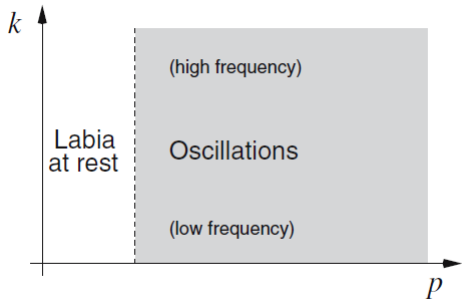

Figura 2.2: Espacio de parámetros. The Physics of Birdsong, figura 5.3 [6].

La velocidad de las variaciones viene descrita por $\phi(t)$, mientras que la diferencia de fase ϕ_0 entre en *k(t)* y *p(t)* dará lugar a los diferentes tipos de sílabas (figura 2.3).

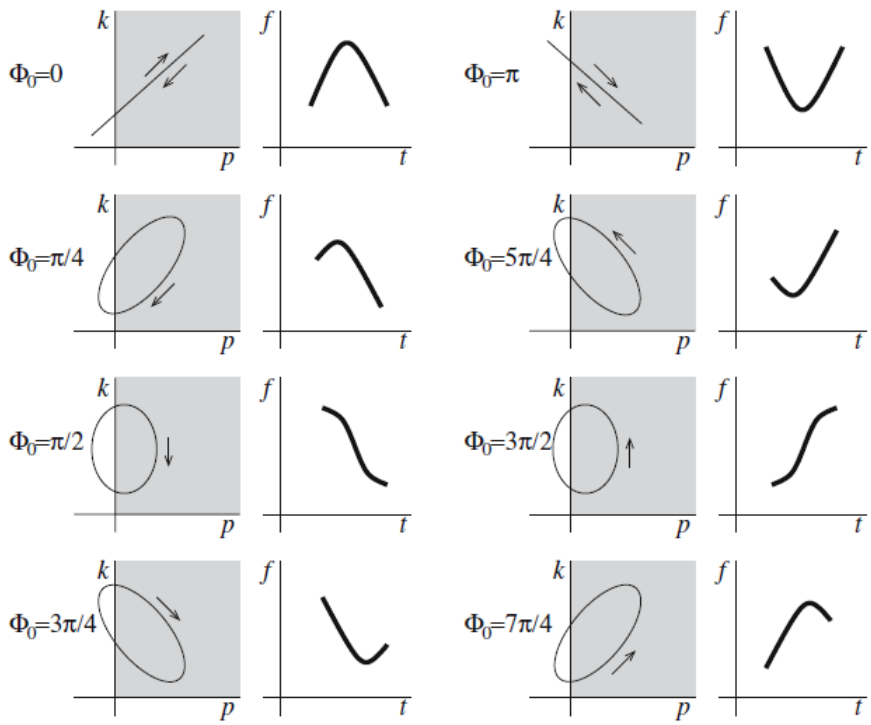

Figura 2.3: Diversidad de sílabas según la diferencia de fase ϕ_0. *The Physics of Birdsong*, figura 5.6 [6]

Una descripción adecuada del espacio de parámetros es fundamental para la clasificación de cantos, ya que la constante de elasticidad k está relacionada con la frecuencia del canto y la presión p es reflejo de la amplitud de la onda. También es necesario resaltar que tanto la elasticidad por unidad de masa como la presión son parámetros que pueden tener interpretación atendiendo al individuo que se está estudiando.

2.2.2. Inicio de la oscilación

Se puede añadir un término adicional de fuerza f_0 al oscilador de van der Pol que explica el inicio del canto como si se tratara de un disparador [4], según se muestra en la ecuación 2.4.

$$\dot{y} = -kx + py - \beta y - \gamma x^2 y - f_0 \tag{2.4}$$

El ciclo de canto de una sílaba comienza con una respiración por lo que es de esperar que los labios estén relajados ($k = 0$) y que la presión sea nula o incluso mínimamente negativa ($p \leq 0$). Para disparar el aumento de la presión y comenzar el canto se abre la siringe y f_0 aumenta hasta un valor mínimo ($f_0 = k\sqrt{\frac{p}{\gamma}}$), donde p ya es suficiente como para comenzar la oscilación. Entonces, f_0 vuelve a su valor original nulo y desde este momento la oscilación queda dominada por el oscilador de van de Pol con $f_0 = 0$. Cuando se llega al final de la vocalización, de forma análoga, f_0 crece y decrece rápidamente, tal que la siringe quede lo suficientemente abierta como para que la presión desaparezca y se cierre el ciclo. La figura 2.4 muestra oscilaciones que incluyen el término f_0.

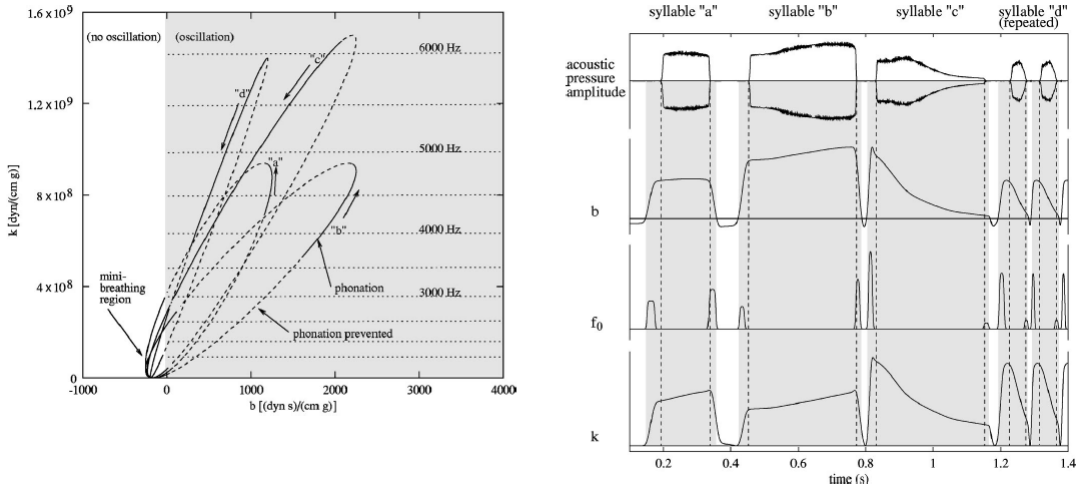

Figura 2.4: Ciclo de este modelo. Izquierda: Espacio de parámetros para cada una de las sílabas. Derecha: Gráfica con la amplitud, presión (b), el disparador (f_0) y la constante elástica, para cada sílaba [4].

En el desarrollo posterior del trabajo no se tendrá en cuenta el término f_0, ya que las condiciones iniciales numéricas de la integración no son nulas, pero es conveniente explicar de forma teórica como comienza el canto para condiciones iniciales nulas.

2.3. Síntesis con series de Fourier

Tal como se explica en el capítulo 2.9 del libro de F. Stremler *Introducción a los sistemas de comunicación* [9], es posible representar una señal en forma de serie de Fourier exponencial compleja, donde F_n son los coeficientes del desarrollo en serie:

$$\begin{aligned} f(t) &= \sum_{n=-\infty}^{\infty} F_n e^{j2\pi\nu_0 nt} \\ \nu_0 &= \frac{1}{t2-t1};\ (T = t2 - t1) \end{aligned} \tag{2.5}$$

Los coeficientes y sus conjugados se pueden calcular con las siguientes integrales en las que se supone que la señal analizada es real.

$$F_n = \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} f(t)e^{-j2\pi\nu_0 nt}dt \tag{2.6}$$

$$F_n^* = \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} f(t)e^{j2\pi\nu_0 nt}dt = F_{-n} \tag{2.7}$$

Ahora se pueden definir nuevos coeficientes:

$$a_0 = F_0 \tag{2.8}$$

$$a_n = [F_n + F_{-n}]) = 2Re(F_n);\ (n \neq 0) \tag{2.9}$$

$$b_n = [F_n - F_{-n}]) = -2Im(F_n) \tag{2.10}$$

Así, se puede expresar la señal como:

$$f(t) = a_0 + \sum_{n=1}^{\infty} a_n + cos(2\pi\nu_0 nt) + \sum_{n=1}^{\infty} b_n sen(2\pi\nu_0 nt) \tag{2.11}$$

La serie de Fourier se puede representar de una forma más compacta como:

$$f(t) = \sum_{n=0}^{\infty} c_n cos(2\pi\nu_0 nt + \phi_n) \tag{2.12}$$

donde:

$$\begin{aligned} c_n = 2|F_n| = 2\sqrt{F_n F_n^*};\ (n \neq 0) \\ c_0 = F_0;\ (n = 0) \end{aligned} \tag{2.13}$$

$$\phi_n = arctan(\frac{Im(F_n)}{Re(F_n)}) \tag{2.14}$$

Como se puede observar, conociendo el módulo de los coeficientes y su desfase se puede expresar la señal. Cuando la frecuencia es diferente para cada ventana de análisis se debe estudiar la frecuencia instantánea.

Frecuencia instantánea

Cuando el argumento del coseno de 2.12 tiene una frecuencia que es diferente para cada instante, la señal podría expresarse con un solo término que depende de la frecuencia instantánea ν_i [9, 12]. El coseno de la señal se expresa con una fase que depende de dicha frecuencia instantánea:

$$f(t) = c_n \; cos(\phi(t)) \tag{2.15}$$

$$c_n = 2|F_n| \tag{2.16}$$

La frecuencia instantánea se define como la variación de la frecuencia en el tiempo y, por tanto, se expresa como la derivada de la fase total del coseno:

$$\omega_i(t) = \frac{\partial \phi(t)}{\partial t} \tag{2.17}$$

Entonces el argumento del coseno $\phi(t)$ depende de la frecuencia instantánea como su integral:

$$\phi(t) = \int \omega_i(t)dt = \int_0^t 2\pi\nu_i(\tau)d\tau + \phi_0 \tag{2.18}$$

Donde se entiende que el resultado de la integral es la frecuencia en cada instante más el desfase inicial.

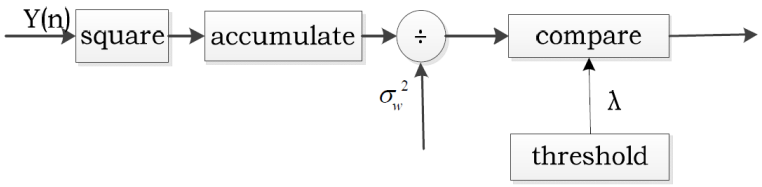

Figura 2.5: Diagrama de flujo de la detección de eventos en el dominio temporal [13].

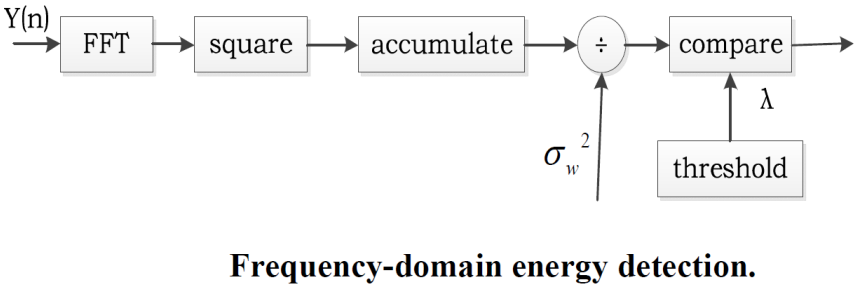

Figura 2.6: Diagrama de flujo de la detección de eventos en el dominio de frecuencias [13].

2.4. Método de detección de eventos

La detección de eventos de sonido permitirá distinguir en qué momento el pájaro produce una sílaba de canto. El procedimiento utilizado en el artículo *Environmental Sound Recognition Using Double-Level Energy Detection* [13], propone una detección de energía doble para la existencia de señales en un ambiente con ruido. Debe haber detección de energía en el dominio temporal y en el de frecuencias.

La detección de energía en el dominio temporal, está relacionada con la intensidad del sonido en un instante de tiempo respecto a su entorno y sigue el diagrama de la figura 2.5. Para ello, se toma el módulo al cuadrado de la señal y se integra para obtener la energía, después se divide entre el valor del ruido σ en unidades de energía, es decir, al cuadrado. A continuación, se establece un límite λ para el número de veces en las que

debe superar la señal al ruido para que exista señal:

$$\frac{\sum_n |x(n)|^2}{\sigma^2} > \lambda \tag{2.19}$$

De forma análoga, la detección en el dominio de frecuencias sigue un procedimiento similar al del dominio temporal (figura 2.6), la diferencia reside en que primero se efectuará la transformada de Fourier de los datos. Luego se calculará su energía, se dividirá por el ruido y se establecerá un umbral a superar.

Capítulo 3

Metodología desarrollada

Abstract

In order to present a set of results there are some steps to accomplish and systems to built: First, it is necessary to find a source of data sound that will be used to test the system. Then, the data is analysed searching for events with birdsong and the parameters (fundamental frequency, spectral density and phase) are obtained. With these parameters, it is possible to synthesize an accurate approximation of the original signal. Finally, it is necessary to analyse the result of the parameters when they are included in the van der Pol oscillator in a particular way.

ooo

3.1. Adquisición de la señal

Para desarrollar el trabajo y demostrar la validez de cada apartado, se utiliza un archivo de sonido con el canto correspondiente al canario (*Serinus canaria*) subcampeón de canto de España en 2011 [5]. Este será el registro de canto objeto de análisis y procesamiento en los apartados posteriores. Por otra parte, se construirá un sistema de códigos escritos en lenguaje de programación *python*, para el tratamiento y análisis de los datos. De forma complementaria se usó la aplicación *Audacity* bajo una licencia de software libre.

3.2. Detección de eventos

Estudiar adecuadamente los parámetros subyacentes del canto de ave requiere analizar los datos de sonido que se tengan, elaborar un registro donde se indiquen las sílabas del canto, es decir, donde se apunte en que instante de tiempo se encuentra cada evento. Además será necesario elaborar cierto refinado de los datos, que descarte las regiones que solo contengan ruido.

De acuerdo a esta idea, se ha elaborado un sistema capaz de detectar los eventos de un archivo de sonido de forma automática, tal que se escriba un fichero con etiquetas temporales para cada evento.

La detección de eventos sigue de forma sintetizada el procedimiento utilizado en el artículo *Environmental Sound Recognition Using Double-Level Energy Detection* [13], tal como se explicó en el apartado 2.4, aunque sólo se utilizará la parte de detección de eventos en el dominio temporal. Se interpretará como canto aquella parte de los datos cuya energía supere un umbral determinado, o expresado de otra forma, tal que la señal sea superior al ruido en una magnitud adecuada. Para calcular la energía instantánea se aplicará un producto de convolución al cuadrado del valor de los datos con una ventana de Hann[1] de 256 píxeles (equivalente a 0.0058 segundos), como se muestra en la ecuación 3.1. Se utiliza una ventana de Hann para suavizar la señal y evitar discontinuidades al principio y final de los datos.

$$E(t) = |x(t)|^2 * Hanning(t) = \int |x(\xi)|^2 \cdot Hanning(t-\xi)d\xi \tag{3.1}$$

A continuación, se dividirá el resultado de la convolución entre el valor de ruido en unidades de energía σ^2. De esta forma cuando los datos sean solo ruido, el valor de la fracción $\frac{E(t)}{\sigma^2}$ será la unidad. El cálculo del valor de ruido se realiza al final de esta sección.

[1]La ventana de Hann, o de coseno elevado se suele denominar ventana de Hanning por la confusión debida a la analogía entre los nombres de dos autores distintos, Julius von Hann y Richard Hamming y a que la primera vez que se referenció esta ventana fue Blackman y Tukey en 1959 donde decían "hanning a signal is to apply the Hann window to it".

Finalmente, se impone la condición de que el valor del cociente antes descrito debe ser mayor que un valor establecido E_{Umbral}, que representa una energía umbral a superar, como se expone en la ecuación 2.19.

$$\frac{E(t)}{\sigma^2} > E_{Umbral} \tag{3.2}$$

Por simplicidad se estableció $E_{Umbral} = 1$, ya que de esta forma se tiene una comparativa directa entre el ruido y la señal, de tal forma que solo se produce detección cuando la energía instantánea supera al ruido como se expresa en la ecuación 3.3.

$$E(t) > \sigma^2 \tag{3.3}$$

Además, se implementa una función en el código de etiquetado que no tenga en cuenta picos de sonido aislados, de tal forma que si la etiqueta tiene una duración muy pequeña será descartada (estableciendo un tiempo mínimo para la detección) o si aparecen intervalos de señal que cumplen la condición umbral y están lo suficiente cerca entre sí (estableciendo un tiempo mínimo de discriminación) serán anotados como un solo evento, en lugar de como varios.

En definitiva, de esta forma, los segmentos con canto recibirán una etiqueta con su momento de inicio y fin para poder trabajar con ellos *a posteriori*, ya que se pretende obtener el valor de la frecuencia fundamental de ellos. La figura 3.1 es un ejemplo visual de lo que es detectado y lo que no. En el apéndice A quedan a disposición del lector algunas líneas del código desarrolladas para este apartado.

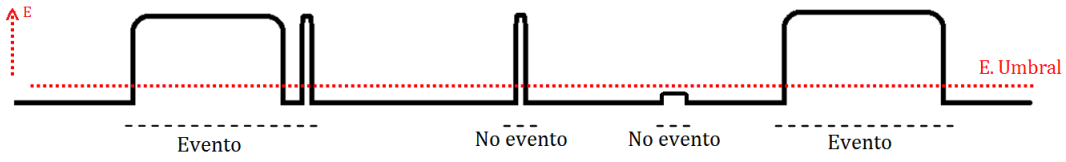

Figura 3.1: Funcionamiento esquemático del detector de eventos.

Cálculo del ruido

Durante el etiquetado de la señal se compara la energía de los datos con cierta σ de ruido que es necesario calcular. Una idea inicial podría ser obtener la desviación estándar de algún segmento formado solo de ruido en los datos, sin embargo, se puede utilizar un proceso menos arbitrario y que es posible generalizar a más casos. Si se realiza un recuento de los valores de la señal que aparecen para cada energía, es de esperar que la mayor parte de los datos se encuentre en los niveles de ruido. Si se representa el número de cuentas frente a la energía, aparece una variación brusca de la pendiente de la curva allí donde se supera el umbral del ruido y queda solo la señal con canto. De esta forma se puede obtener una estimación de la magnitud del ruido en los datos, aproximadamente: $\sigma^2 = 40$ unidades de energía.

3.3. Estimación de la frecuencia fundamental

Para encontrar las frecuencias fundamentales se ha decidido aplicar un algoritmo iterativo que permita aumentar la resolución en frecuencia y en tiempo de cada evento. Se definirá la densidad espectral de una ventana de muestras con datos suficientes para estimar la frecuencia fundamental. En este caso, la frecuencia se encuentra en torno a $\nu_f = 2000$ Hz y se utilizará una ventana de N=64 píxeles (0.00145 segundos).

Para representar la ventana muestreada se utilizará la expresión 3.4 donde cada muestra de la ventana ha sido representada por medio de una delta de Dirac. A continuación se resolverá su transformada de Fourier expresada en 3.5, para las frecuencias entre 1000 y 4000 Hz, para así obtener la consecuente densidad espectral para cada frecuencia de este intervalo, véase 3.6 y 3.7.

$$f(t) = \sum_{n=0}^{n=64-1} x_n \delta(t - n\Delta t) \tag{3.4}$$

$$F(\nu) = \frac{1}{T}\int_{-\infty}^{\infty}\left(\sum_{n=0}^{n=63} x_n \delta(t - n\Delta t)\right) e^{-j2\pi\nu t} dt \tag{3.5}$$

$$F(\nu) = \frac{1}{T}\sum_{n=0}^{n=63} x_n e^{-j2\pi\nu n\Delta t}\Delta t \tag{3.6}$$

Siendo $T = N\Delta t$, con N=64 píxeles, la densidad espectral queda:

$$F(\nu) = \frac{1}{64}\sum_{n=0}^{n=63} x_n e^{-j2\pi\nu n\Delta t} \tag{3.7}$$

El método computacional recorre el intervalo de frecuencias donde se espera encontrar el resultado, entre 1000 y 4000 Hz con una resolución de 0.1 Hz. El máximo de la densidad espectral corresponderá con la frecuencia fundamental. Este proceso se aplica a su vez a cada ventana de 64 píxeles, con un desplazamiento de dicha ventana de un píxel al siguiente.

Se deduce entonces que este cálculo aporta tres parámetros de utilidad: la frecuencia fundamental ν_0 (siendo sus múltiplos las frecuencias armónicas), el módulo de la densidad espectral máxima $|F(\nu_0)|$, de la ecuación 3.8, y el desfase de dicha densidad espectral en la ecuación 3.9, $\phi(\nu_0)$.

$$|F(\nu_0)| = \sqrt{F(\nu_0)F^*(\nu_0)} \tag{3.8}$$

$$\phi(\nu_0) = \frac{Im(F(\nu_0))}{Re(F(\nu_0))} \tag{3.9}$$

3.4. Síntesis de la señal

Una vez obtenidos estos parámetros, siguiendo la explicación del apartado 2.3, es posible recuperar la señal original de forma sintetizada, expresándola como un solo término:

$$x(t) = 2|F(\nu_0)| \cdot cos(\phi(t)) \tag{3.10}$$

Esta señal se denominará sintetizada ya que se compone de los parámetros esenciales extraídos de la señal original. Su amplitud será dos veces la densidad espectral máxima

en cada instante, mientras que $\phi(t)$ depende en cada momento de su correspondiente frecuencia instantánea y el desfase inicial. El argumento del coseno se construye integrando la frecuencia fundamental instantánea:

$$\phi(t) = \int_0^t 2\pi\nu_0(\tau)d\tau + \phi_0 \tag{3.11}$$

Como se acaba de explicar, conocer los parámetros ν_0, $|F(\nu_0)|$ y ϕ_0 permite trabajar con los valores numéricos de una señal sintetizada, en lugar de con todos los datos de la grabación original. En definitiva, se han extrapolado aquellos valores que construyen el canto de forma estructural. Esto permite a su vez construir un espacio de parámetros similar al propuesto en la sección 2.2.1, que da una idea del tono y amplitud de la onda, por lo que ofrece la posibilidad de clasificar y cuantizar el canto. Además, permitirá resolver numéricamente la ecuación de van der Pol. En el apartado 4.2 se expondrán gráficamente los resultados de este método.

3.5. La ecuación de van der Pol

Tal como se explicó en el apartado 2.2, el modelo de van der Pol describe la dinámica del sistema y aporta soluciones que reproducen la señal si se introducen los parámetros adecuados. Es posible integrar la ecuación de van der Pol numéricamente en el intervalo de tiempo del evento, introduciendo para cada instante un valor de *k, p* y γ.

La integración se realiza con la función *odeint* de la librería *scipy* de *python*. Es necesario aportar las condiciones iniciales, que se obtendrán de 3.10 para el primer valor de la frecuencia fundamental. Además, se aporta un método iterativo que interpola los valores de la presión y la constante elástica, que calcule de forma automática el valor de dichos parámetros para cualquier instante de tiempo que solicite la función *odeint*.

En la integración de la ecuación de van der Pol se omite el término β, que será incluido

dentro de la presión, que quedará redefinida:

$$\begin{gathered} \dot{y} = -kx + (p - \beta)y - \gamma x^2 y \\ Siendo\ (p - \beta) = p' = p \end{gathered} \tag{3.12}$$

Los datos sintetizados en el apartado 3.3 son la frecuencia fundamental, la densidad espectral y el desfase, que seguirán alguna relación con k, p y γ.

La relación entre la frecuencia fundamental y la constante elástica ya es conocida:

$$k = (2\pi\nu_0)^2$$

Sin embargo, la relación de la presión y el rozamiento γ con los parámetros obtenidos no es tan clara y la ecuación de van der Pol no aporta soluciones teóricas que permitan estimar la relación entre ambos. Las referencias bibliográficas del apartado 2.2 y los ajustes numéricos ayudan a esclarecer las claves para conocer p y γ:

- Se toma la presión como el cuadrado de la amplitud de la oscilación por un factor constante de ajuste α. Se estima que los valores de la presión se encuentran en torno a 2000 Pa [3, 4, 8], esto ayudará a seleccionar α adecuadamente:

  $$p = \alpha(2|F(\nu_0)|)^2$$

  La amplitud se toma de la ecuación 3.10.

- El rozamiento γ se toma como una constante, está relacionado con la posición y depende del tejido de las paredes del conducto y se estima que su valor variará poco en los intervalos de tiempo que dura una señal.

Los valores de α y γ han sido obtenidos calculando la función de coste, que es un ajuste entre la amplitud de la señal sintetizada y las diferentes soluciones de la integración del oscilador de van der Pol, que busca la distancia mínima entre ambas.

Se utiliza un método iterativo que fija un valor para γ y recorre los valores de α en un intervalo, hasta encontrar su mínimo local más pequeño. Después se fija α y se

recorre γ. A continuación, se realiza un mapeado del entorno de estos mínimos iniciales de α y γ, buscando aquella región en la que se encuentre el valor mínimo del coste. La repetición iterada de este proceso conduce al par de valores para α y γ que ajustan mejor la solución de van der Pol a la señal.

En el apartado 4.3 se presentan las soluciones del oscilador de van der Pol para los α y γ óptimos. En el apéndice B se exponen algunas líneas del código desarrolladas según esta sección y que dieron como resultado las imágenes del apartado 4.3.

Capítulo 4

Resultados y discusión

Abstract

After all the previous tasks are done, it is possible to introduce and discuss the results. The synthesis of the signal works successfully and provides the possibility to include the parameters in the van der Pol equation. It is necessary to adjust the value of these parameters before including them in the van der Pol oscillator, then, the equation is integrated numerically and the solutions are compared with the original signal. Finally, there is a brief explanation of why the fundamental frequency has jumps in some sound events.

ooo

4.1. Evento utilizado

El evento que servirá de modelo para exponer los resultados en los apartados 4.2 y 4.3 corresponde a la segunda sílaba del archivo de sonido utilizado [5]. Su forma se muestra en la figura 4.1. En el apartado 4.4 se usa de ejemplo la sílaba trigésima. En la figura 4.5 y figura 4.6 se utilizan las cinco primeras sílabas de canto.

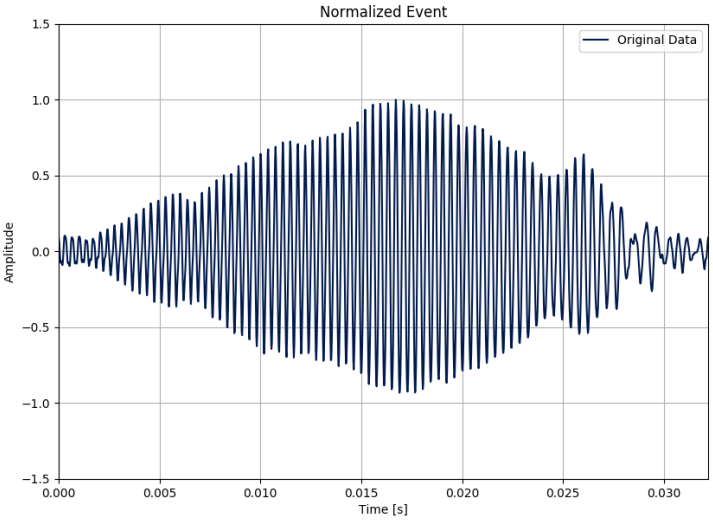

Figura 4.1: Segunda sílaba del archivo de sonido, que será utilizada de modelo.

4.2. Señal sintetizada

Es conveniente comprobar que los parámetros calculados a partir de los datos, devuelven la señal original de forma sintetizada, como se explica en el apartado 3.4. Para ello se realizará un contraste visual entre ambos, como se muestra en la figura 4.2. Se observa que la señal sintetizada y la original tienen una envolvente similar y cierto desfase hacia el final de la oscilación. Este desfase aparece porque en los datos originales de sonido, la oscilación no es regular, tiene ruido, otras oscilaciones acopladas y varía mucho en poco tiempo. La señal sintetizada es una aproximación de la señal original y estimará la frecuencia contemplando todas las irregularidades antes descritas, por tanto se obtendrá una oscilación promediada y más suave, que no concuerda con la original pero permite trabajar con datos redondeados.

En la figura 4.3 se observa que la frecuencia fundamental encontrada sigue adecuadamente el espectrograma de la señal original, siendo los armónicos los múltiplos de esta.

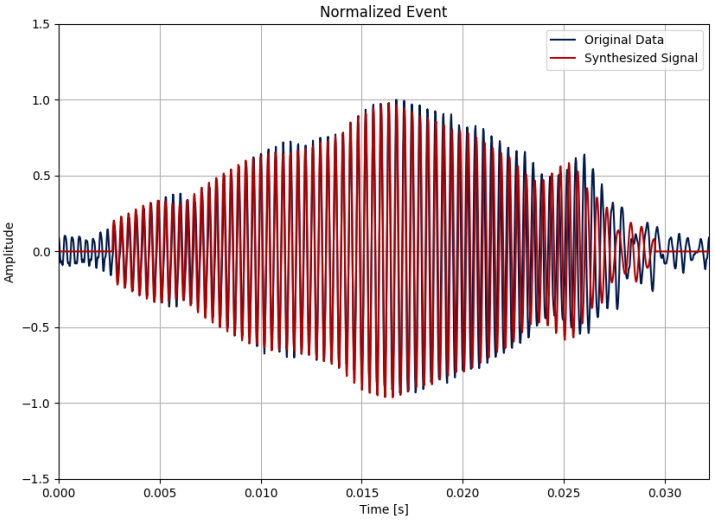

Figura 4.2: Comparativa entre la señal original y la señal sintetizada

Además, se puede construir un espacio de parámetros como el mostrado en la figura 4.4, con la frecuencia fundamental y la densidad espectral, que más adelante darán idea de la constante elástica y la presión. En la figura 2.3 aparece un ciclo (abierto en este evento) similar al indicado en la sección 2.2.1, aunque no se ajusta a una elipse perfecta.

El espacio de parámetros es particularmente útil para reconocer cada tipo de canto y poder realizar su clasificación, ya que alberga información sobre la amplitud y la frecuencia de la onda en cuestión.

También se puede estudiar el canto del pájaro como un conjunto de varias sílabas, de esta forma se aprecia con más perspectiva la repetición de sonidos y se comprenderá mejor cualquier posible clasificación. Por ejemplo, se muestran los cinco primeros eventos del canto con la señal sintética en la figura 4.5 y su espectrograma en la figura 4.6. Se puede observar el parecido que mantienen las cinco sílabas en la forma de la envolvente y en la variación de la frecuencia. Así es posible entender de forma más global la estructura del canto del ave, sus patrones y repetición de determinadas sílabas.

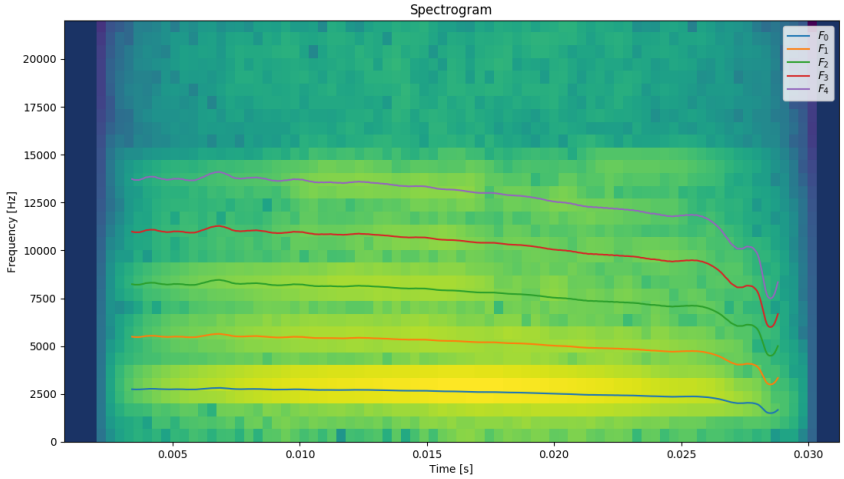

Figura 4.3: Espectrograma de la señal original (fondo) y la frecuencia fundamental y sus cuatro primeros armónicos

4.3. Integración de la ecuación de van der Pol

Las soluciones de la ecuación de van der Pol reproducen la señal si se introducen los parámetros adecuados, tal como se explica en el apartado 3.5. Para hallar los valores óptimos se realiza un mapeado de α y γ calculando la función de coste para cada par de valores, como se presenta en la figura 4.7. Los valores obtenidos tras el ajuste son $\alpha = 5560 \pm 10$ y $\gamma = 25210 \pm 10$, para este evento. Además, con $\alpha = 5560$ parece que la presión se encuentra en los valores esperados según la bibliografía.

Una comparativa entre la señal y la simulación con el oscilador de van der Pol expondrá con claridad hasta qué punto se ha logrado hacer coincidir la oscilación, como muestra la figura 4.8.

Se observa que el oscilador de van der Pol aproxima la forma de la envolvente y progresivamente se desfasa de la señal, tal y como se aprecia entre las figura 4.9 y

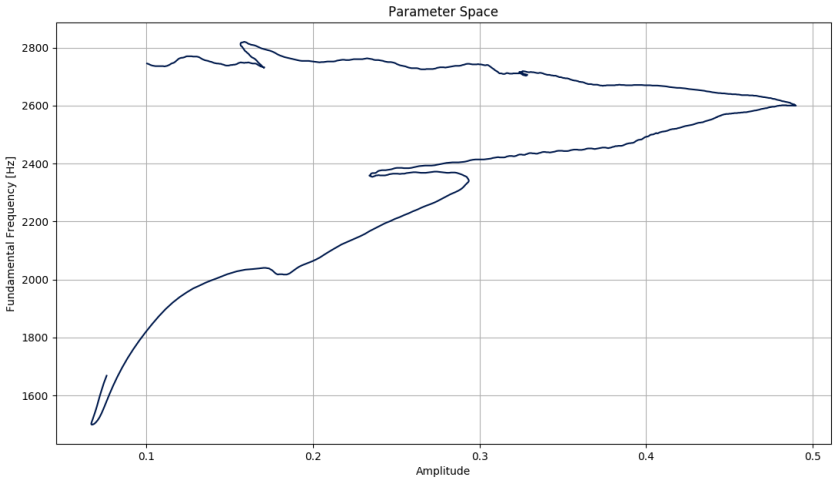

Figura 4.4: Espacio de parámetros.

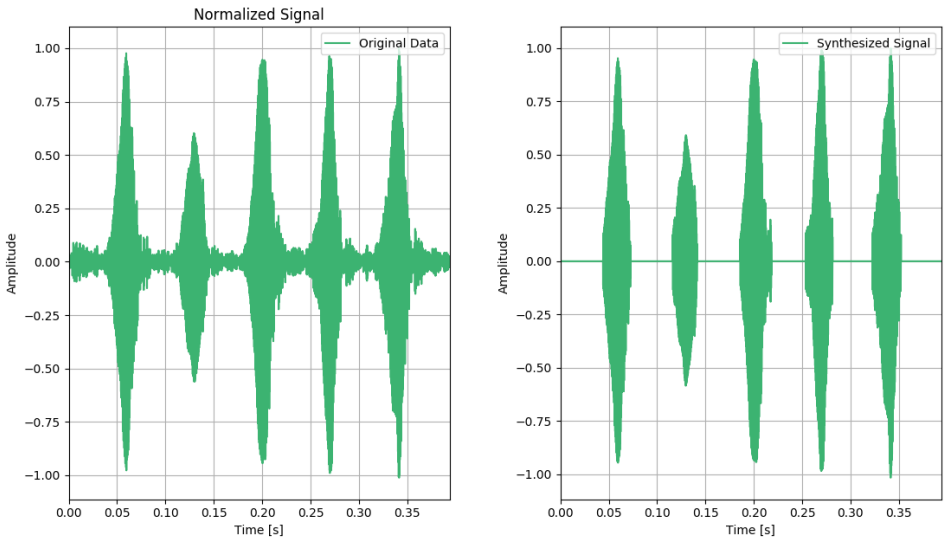

Figura 4.5: Señal original y señal sintetizada de las cinco primeras sílabas de canto del ave.

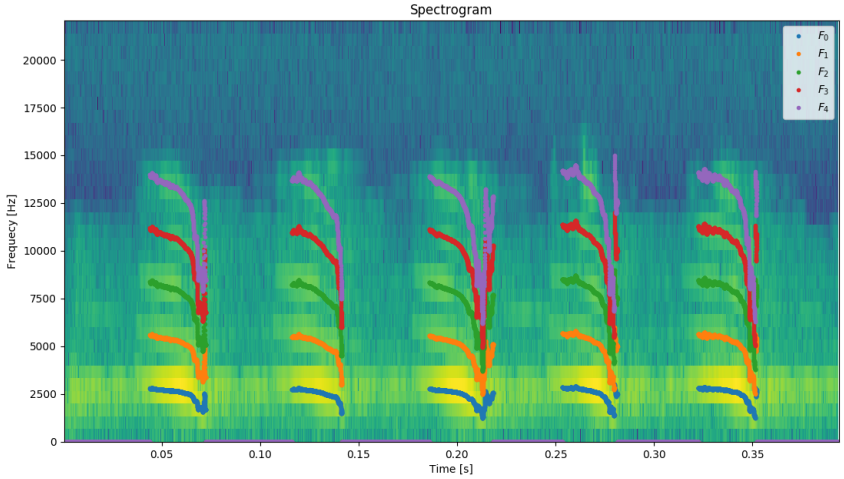

Figura 4.6: Espectrograma con las frecuencias fundamentales de las cinco primeras sílabas de canto del ave.

figura 4.10. Una de las consecuencias de este desfase es que el ajuste para optimizar α y γ pierde precisión, al no coincidir la amplitud de ambas oscilaciones en la posición adecuada. Se estima que las causas de este desfase provienen del error acumulado, de la falta de precisión en los parámetros obtenidos con la señal sintetizada, en el propio ajuste realizado para α y γ o en los límites numéricos de la integración.

Por otra parte, la ecuación de van der Pol no genera la segunda frecuencia armónica, dado que no existe ningún término cuadrático capaz de generarla. Por ende, existe otra discrepancia entre la solución de van der Pol y la señal, que supondrá cierta diferencia en la energía entre ambas y presumiblemente explicaría parte de la diferencia de amplitudes entre ambas ondas.

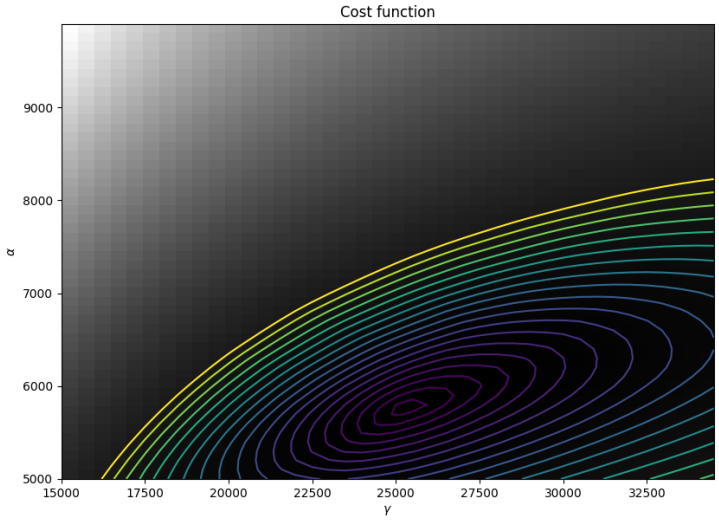

Figura 4.7: Mapa de la función de coste en función de los parámetros, con mínimo en $\alpha = 5560 \pm 10$ y $\gamma = 25210 \pm 10$.

4.4. Búsqueda de la frecuencia fundamental

La búsqueda de lafrecuencia fundamental tiene ciertas limitaciones, ya que en las gráficas producidas aparecen oscilaciones bruscas en el valor de la frecuencia.

El método propuesto en la sección 3.3 no es completamente fiable, ya que se basa exclusivamente en buscar el valor máximo de la densidad espectral y suponer que este valor coincidirá con el de la frecuencia fundamental. Sin embargo, en determinadas ocasiones aparecen escalones abruptos en la frecuencia, la mayor parte de las veces entre la frecuencia fundamental y el primer armónico, véase figura 4.11. Esto es debido a que en espectrogramas con poca definición se puede encontrar la mayor densidad espectral en el primer armónico, dando lugar a estos saltos. Sería posible evitar estos saltos si se implementa un método que realice el *tracking* de la señal más cuidadosamente, por ejemplo, con un método que siga la frecuencia fundamental en un entorno cercano, asegurando así cierta suavidad en su obtención.

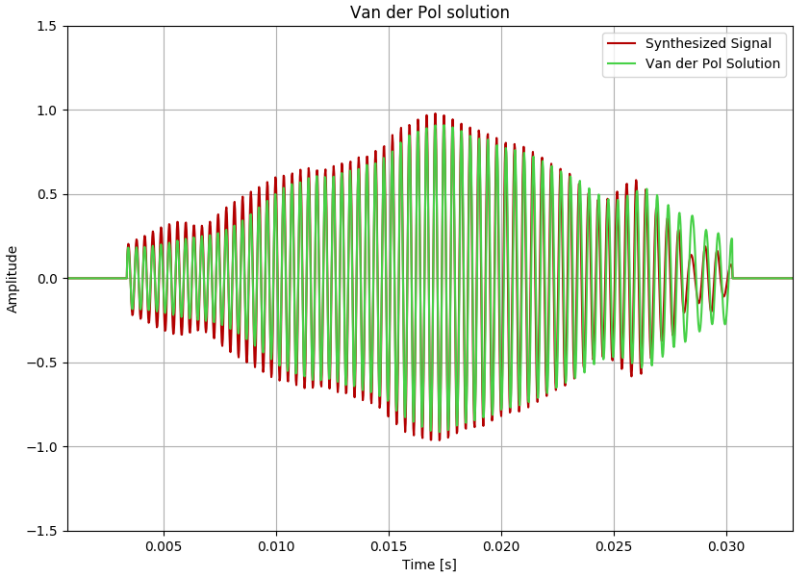

Figura 4.8: Comparativa entre la señal sintetizada y la solución del oscilador de van der Pol. Para $\alpha = 5560 \pm 10$ y $\gamma = 25210 \pm 10$

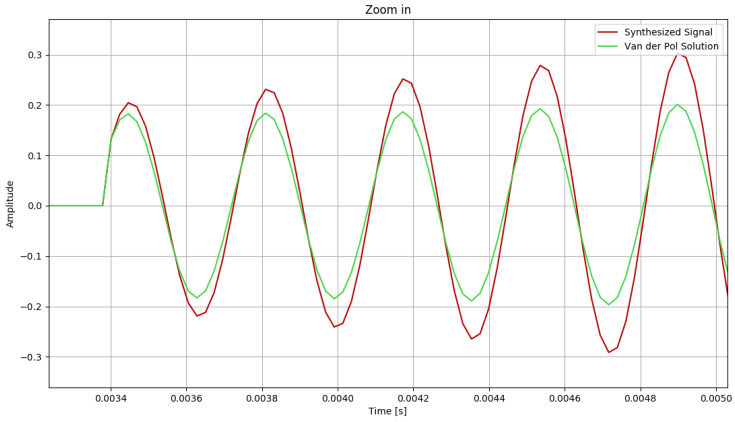

Figura 4.9: Señal sintetizada y oscilación de van der Pol superpuestas. Inicio de la oscilación en fase.

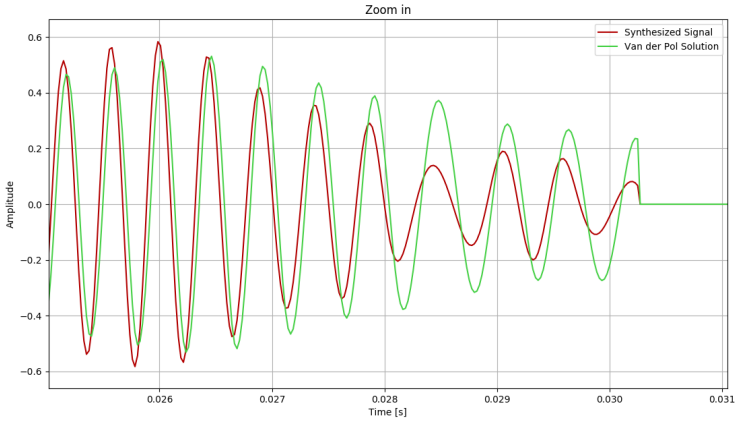

Figura 4.10: Señal sintetizada y oscilación de van der Pol superpuestas. En las últimas oscilaciones aparece un desfase.

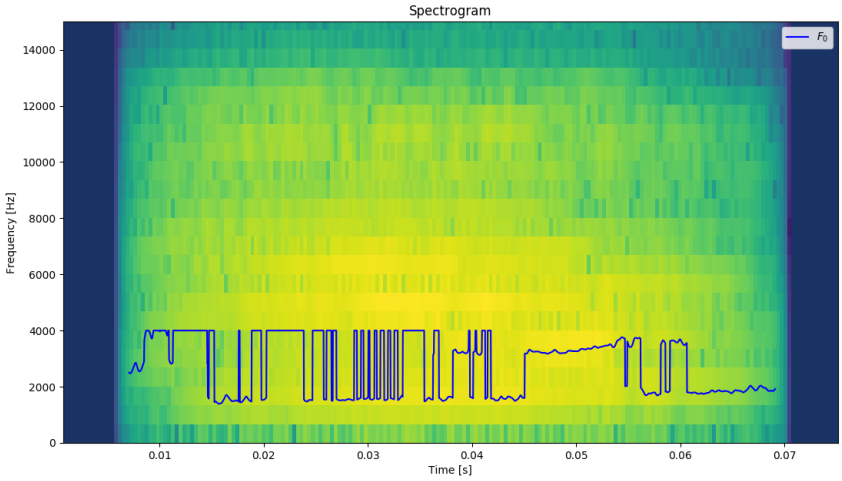

Figura 4.11: Ejemplo de un mal seguimiento de la frecuencia fundamental. La frecuencia fundamental se encuentra a unos 2000 Hz y su primer armónico a 4000 Hz, la falta de definición en el espectro causa los saltos y toma el armónico como si fuera la frecuencia fundamental.

Capítulo 5

Conclusiones y perspectivas de futuro

Abstract

To close the research, it is a good idea to present a list of future works and what are the lacks of the method. There are a lot of possible improvements to add and potential applications to the field. Finally, the conclusions are exposed, trying to explain the results and summarizing the achievements.

ooo

5.1. Líneas futuras de trabajo

El trabajo desarrollado puede ser ampliado ya que muchos de los métodos utilizados son susceptibles de mejora.

En el apartado de adquisición de la señal, sería posible obtener la señal de una fuente mejor conocida, en la que se sepa con claridad como se grabó y trató el archivo de sonido original. Esto posibilitaría trabajar con información adicional, tener un mayor control y opciones de personalización. Aunque *Youtube* se ha revelado como una muy buena fuente de datos sonoros de animales, hay otras fuentes como, por ejemplo, la British Library Sounds donde se pueden conseguir registros de los que extraer información sobre

distintas especies.

Aunque la detección de eventos funciona correctamente, se podría añadir la detección en el dominio de frecuencias, como se describió en el apartado 2.4, para aquellos casos con poca resolución o mucho ruido en el dominio temporal.

En la parte de cálculo del ruido, la adquisición del valor de ruido no es estricta ni está automatizada, sería conveniente implementar un método que busque el valor de ruido de forma continua, que calcule su valor de forma precisa y lo implemente a la propia detección de eventos para filtrar la señal adecuadamente.

El seguimiento de la frecuencia fundamental funciona buscando el valor máximo de la densidad espectral, sin embargo, tal como se explicó en el apartado 4.4, este método conduce a errores y saltos. Es posible solventar este defecto si se realiza un *tracking* de la frecuencia fundamental que utilice de forma iterativa el valor anterior de la frecuencia para estimar el siguiente en un entorno cercano. De esta forma el seguimiento de la frecuencia fundamental será más suave y sin oscilaciones abruptas.

La integración de la ecuación de van der Pol conduce a cierto desfase respecto a la señal original, corregir este aspecto depende del método de integración utilizado y de los valores de los parámetros introducidos, que requieren un estudio más profundo. La relación de la presión y el rozamiento γ con los parámetros obtenidos de la síntesis de la señal ν_0, $|F(\nu_0)|$ y $\phi(\nu_0)$ necesita una aproximación más rigurosa y una mejor justificación teórica. Además, la amplitud de la solución para los parámetros óptimos no coincide con la amplitud de la envolvente de la señal, entre otros factores, esto se puede deber a que el oscilador de van der Pol no genera la segunda frecuencia armónica. Si se justifica una ampliación de la ecuación que introduzca un término cuadrático quizás se podría generar un segundo armónico, que presumiblemente ajustaría mejor la amplitud. Un ejemplo es la propuesta de hacer que la elasticidad no sea lineal, lo que permitiría introducir los armónicos pares de los que carece la ecuación de van der Pol.

Es necesario estimar y estudiar el error que se produce al aproximar los datos con

la señal sintetizada. También se debe calcular la desviación de la solución de van der Pol respecto a la señal sintetizada, y por tanto, el error respecto a los datos originales. Además, se podría calcular la magnitud de las pérdidas por la falta del segundo armónico en la ecuación de van der Pol.

Uno de los potenciales más interesantes de este trabajo es la posibilidad de clasificar eventos. Si se elabora un método automático que compare el espacio de parámetros de cada uno y establezca criterios de similitud, se obtendría un sistema que favorecería la catalogación, el censo y el estudio de la conservación biológica.

Si se quisiera seguir la línea de trabajo de este estudio, sería interesante perfeccionar el modelo propuesto o estudiar otros diferentes para otras aves u otro tipo de vocalizaciones. Se podría adaptar el sistema a otras especies animales y generalizar este método de tratamiento de sonidos. Además, el estudio de la no linealidad de la oscilación de van der Pol ayudaría a explicar las relaciones de los parámetros de la ecuación, que aún requiere una investigación más profunda.

5.2. Conclusiones

Para cerrar el trabajo parece adecuado recapitular los hitos superados y elaborar un compendio del trabajo aportado:

- Se ha construido un sistema capaz de detectar eventos, analizarlos por separado, extraer sus parámetros fundamentales y sintetizar la señal. La señal sintetizada ofrece la ventaja de poder trabajar con datos refinados, que permiten estudiar y manipular la información esencial de cada evento. Esto abre la puerta al análisis del espacio de parámetros, la clasificación y la simulación de la señal con modelos matemáticos, como el oscilador de van der Pol.

- La síntesis de la señal presentada en el apartado 4.2 aproxima satisfactoriamente a los datos originales. De esta forma se comprueba que la composición de la señal

mediante series de Fourier [9] explicada en el apartado 2.3, genera adecuadamente el canto del pájaro si se introducen los datos para la frecuencia fundamental y la densidad espectral adecuados.

- Los parámetros obtenidos en la síntesis pueden representar la presión y la constante elástica, por tanto, su espacio de parámetros indica el flujo y el estado del canto en cada momento, véase la figura 4.4. Tal como se predijo en el apartado 2.2.1, figura 2.3, la imagen muestra la presencia de un ciclo, la respiración (aumento y descenso de la presión) y el flujo del proceso de canto del pájaro parece coincidir con lo explicado en la bibliografía. Sin embargo, la figura no se ajusta a una elipse, por lo que se puede deducir que las expresiones cosenoidales 2.2 y 2.3, propuestas en [6], podrían no representar con exactitud la oscilación de la presión y la constante elástica de datos experimentales.

- Por último, la idea principal del trabajo residía en estudiar las ecuaciones para la producción de cantos de pájaros, partiendo del modelo del libro *The physics of birdsong* de Gabriel Mindlin [6], para justificar la validez de los modelos propuestos. La producción de cantos de pájaros paseriformes con el oscilador de van der Pol simula aproximadamente la señal, con cierto desfase y desajuste en la amplitud, pero que con un estudio más profundo puede tener un futuro prometedor.

En conclusión, el sistema creado funciona adecuadamente pero podría ser mejorado, los parámetros obtenidos son válidos y la señal sintetizada aproxima correctamente. En cuanto a la simulación de cantos de pájaros con la ecuación de van der Pol, las soluciones obtenidas pueden ser tomadas como una primera aproximación válida.

Bibliografía

[1] Ana Amador y Gabriel B Mindlin. «Low dimensional dynamics in birdsong production». En: *The European Physical Journal B* 87.12 (2014), pág. 300.

[2] Ezequiel M Arneodo y Gabriel B Mindlin. «Source-tract coupling in birdsong production». En: *Physical Review E* 79.6 (2009), pág. 061921.

[3] Tim Gardner y col. «Simple motor gestures for birdsongs». En: *Physical review letters* 87.20 (2001), pág. 208101.

[4] Rodrigo Laje, Timothy J Gardner y Gabriel B Mindlin. «Neuromuscular control of vocalizations in birdsong: a model». En: *Physical Review E* 65.5 (2002), pág. 051921.

[5] Pedro Mazuelas. *Subcampeon de España 2011 Canario Amarillo Nevado cantando 93 Puntos hermano del campeon 94p*. 2012. URL: `https://www.youtube.com/watch?v=jpYukYPAOKc`.

[6] Gabriel B Mindlin y Rodrigo Laje. *The physics of birdsong*. Springer Science & Business Media, 2006.

[7] GB Mindlin. «The physics of birdsong production». En: *Contemporary physics* 54.2 (2013), págs. 91-96.

[8] JD Sitt y col. «Dynamical origin of spectrally rich vocalizations in birdsong». En: *Physical Review E* 78.1 (2008), pág. 011905.

[9] Ferrel Stremler. *Introducción a los sistemas de comunicación*. 1993.

[10] M. Alberto Trevisan y G. B. Mindlin. «Models of birdsong (physics)». En: *Scholarpedia* 4.3 (2009). revision #91505, pág. 6923. DOI: 10.4249/scholarpedia.6923.

[11] Wikipedia. *Siringe — Wikipedia, La enciclopedia libre*. 2017. URL: https://es.wikipedia.org/w/index.php?title=Siringe&oldid=100022965.

[12] Wikiversidad. *Modulaciones angulares — Wikiversidad,* 2015. URL: https://es.wikiversity.org/w/index.php?title=Modulaciones_angulares&oldid=107868.

[13] Xiaoxia Zhang y Ying Li. «Environmental sound recognition using double-level energy detection». En: *Journal of Signal and Information Processing* 4.03 (2013), pág. 19.

Apéndice A

Varios códigos del sistema

Algunas de las definiciones usadas en la detección de eventos de sonido, como se explicó en la sección 3.2.

```
#!\usr\bin\python
# *-* coding: utf-8 *-*
from __future__ import division
import numpy as np
import matplotlib.pyplot as plt
import scipy.io.wavfile as wav

def calc_energia(f,ven=256,dt=1/44100,variance=40.):
        '''Get the energy of a real signal
        Input:
                f = signal array
            ven = size of the convolution window
            dt = sample interval
            variance = noise energy
        Output:
```

```
16              ener: energy array in noise energy units
17          '''
18          f=f.astype('float')
19
20          # Convolution to obtain the mean energy
21          accum=np.convolve(f*f,np.hanning(ven),'same')*dt/ven
22          ener=accum/(variance)
23          return(ener)
24
25  def escribe_etiquetas(ener,file_etiq,dt,umb_ener=1,tmargen=0.0):
26          '''Write the time labels of the event
27          Input:
28                   umb_ener = threshold of energy
29              tmargen = threshold of time, smaller values will be
                    discarded as a signal or as noise
30              file_etiq = incoming sound data file
31          Output:
32              fi = new file with the labels of each detected event
33          '''
34          fi=open(file_etiq,'w+')
35          #Loop for segment selection
36          ant=False
37          ce=[0.,0.,'']
38          n=0
39          mar_det=mar_disc=int(tmargen/dt)  #threshold of detection
                for a valid event / threshold of discrimination for
                two consecutive events
40          for i in range(len(ener)):
41          act=ener[i]>umb_ener
```

```
42          t=i*dt
43          if act and not ant:
44            if i-ce[1]/dt > mar_disc:
45          if ce[1]/dt-ce[0]/dt > mar_det:
46            fi.write('%f\t%f\t%s\n' %(ce[0],ce[1],ce[2]))
47            n+=1
48          ce[0]=t
49          elif not act and ant:
50            ce[1]=t
51            ce[2]='S%d' %n
52
53          ant=act
54
55          #For the final inputs of the loop
56          if ant:
57          ce[1]=t
58          ce[2]='S%d' %n
59          if ce[1]/dt-ce[0]/dt > mar_det:
60          fi.write('%f\t%f\t%s\n' %(ce[0],ce[1],ce[2]))
61          fi.close()
62          return
```

Apéndice B

Varios códigos del sistema

Varias definiciones de una clase (*class*) de *python*, para estudiar las soluciones de la ecuación de van der Pol, tal y como se expone en los apartados 3.5 and 4.3.

```
def interp(self,t):
        '''Interpolation method able to integrate any time value
        Input:
                t = incoming time value
        Output:
                k, p, g = elastic constant, pressure and gamma
                   friction term for the time value entered
        '''
        idt_i=int(t//self.dt)
        idt_f=t%self.dt
        #For the final inputs
        if idt_i > len(self.ff)-2:
                f=self.ff[-1]
                amp=self.dem[-1]
```

```
14          #Interpolation between each pair of values associated
               with the entered instant of time
15          else:
16                  f=(1-idt_f)*self.ff[idt_i]+idt_f*self.ff[idt_i+1]
17                  amp=(1-idt_f)*self.dem[idt_i]+idt_f*self.dem[
                       idt_i+1]
18          #Definitions of the elastic constant, pressure and gamma
               friction term
19          k=(2*np.pi*f)**2
20          p=4*self.alfa*amp**2
21          g=self.gamma
22      return k,p,g
23
24  def pend(self,st,t):
25          '''State of the system
26          Input:
27                  st = Values of the state of the system [position,
                       velocity]
28          Output:
29                  dst = Differential of the state of the system [
                       velocity, van der Pol equation]
30          '''
31      k,p,g=self.interp(t)
32      x, v=st
33      dst=[v,-k*x+p*v-g*x*x*v-g*x*x]
34      return dst
35
36  def __call__(self,a,g):
37          '''Odeint numerical integration and cost function
```

```
38          Input:
39                  a = alpha value to scale the pressure value
40                  g = gamma value from the friction term
41          Output:
42                  cost = value of the cost function for this alpha
                       and gamma parameters
43          '''
44      self.alfa=a
45      self.gamma=g
46      self.sol = odeint(self.pend, self.st0, self.t)   #Integration
47      cost=np.sum(np.abs(self.f*self.f-np.roll(self.sol[:,0]*self.
           sol[:,0],-1)))   #Cost function
48      return cost
```