Politecnico di Torino

Master of Science in Biomedical Engineering

# Recognition of atrial pathologies in the elderly through the study of the P wave morphology in the Electrocardiogram

*Master's Thesis*

*By*

Francesca Di Cintio

July 2018

Supervisors:

Prof. Dr. Ing Filippo Molinari (Polytechnic of Turin)

Dr. Ing David Lombardi (Cardioline)

***Daring ideas are like chessman moved forward they may be beaten but they may start a winning game  cit. Goethe***

*A chi mi ha donato la Vita*
*A chi mi ha donato l'Amore*
*A chi mi ha donato l'Amicizia*
*A chi mi ha donato il Sapere*
*Grazie.*

# Declaration

The work in this thesis is based on research carried out at the Department of Cardiovascular Diseases, ''San Giovanni Addolorata'' Hospital, in collaboration with Cardioline, Biomedical Company, supervised by Dr. David Lombardi and Prof. Filippo Molinari. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text.

# Preface

This master's thesis has been written in the Department of Electronics and Telecommunications (DET-BioLab), of the Politecnico di Torino, in collaboration with Cardioline Spa.

Thanks to the meeting with this company I had the possibility to work on this project.

This paper was prepared in collaboration with my supervisor Prof. Filippo Molinari, whit the supervision of the project by Dr. Fabio Rangoni and Dr. David Lombardi.

I would like to thank my supervisor Prof. Filippo Molinari, my supervisor at Politecnico di Torino, for all his support in helping me with the problems that I confronted during this work, Dr. David Lombardi and Dr. Fabio Rangoni, chair of the Company.

# Abstract

The global population aged 60 years or over is more than twice large than in 1980, and this number will increase in the future. In Italy, after Japan, the elderly constitute an important part of the total population, with a number of people over 60s which is expected to increase in the following years.

In our country there is the highest proportion of elderly citizens (22.4% of the total population was more than 65 years hold in 2017) among European countries. For this reason aging and chronic diseases can be considered two of the major concerns that the public health care system has to face up to in the next years. Working for solving aging related problems will help to reduce health care cost and improve citizens quality of life. [*sit 1*]

In particular, cardiovascular diseases (CVD) in aged people can be considered one of the more frequent and predictable problems which comes out in aged people.

One-third of deaths all over the word is caused by Cardiovascular diseases. "Past periods of decline in cardiovascular disease mortality marked a remarkable achievement for public health and medical care around the world," said Dr. Christopher Murray, director of IHME and study co-author. "Governments, advocacy groups, clinicians, and communities should look to this new evidence when developing programs and policies that could reduce the burden of cardiovascular disease and save more lives." [1]

Automatic analysis of Electrocardiograms (ECGs) signal provides an important support for diagnostic classification, bulding up ECG databases which could help to predict some heart disease.

A recent study, named 'The Predictor' investigates the correlation between the P-wave morphology and the heart failure. It has demonstrated that it is possible to predict the heart disease insurgence analyzing the ECG signal. [2]

Some cardiac abnormalities can be identificated by the extraction of a series of significant ECG parameters. A research, conducting by an equipe of doctors at the ''San Giovanni Addolorata'' Hospital in Rome, revealed six principal classes into which the P-wave morphology can be classified. They found also a correlation between P-wave duration and the presence of an heart related problem.

This study wants to develop an algorithm which uses high-resolution digital electrocardiograms (1000 Hz, 20 bits) for automatic detection and the classification of P-wave. It can be used as a support for physicians because of its ability to give results in a faster and repeatable way than previous signal analysis techniques. The main goal is to give a early diagnosis that could avoid the sudden heart failure in aged patients.

The algorithm system has been implemented and simulated in Matlab. The algorithm was implemented used recordings of electrocardiograms track from twenty-three patients. Each file was given in xml format, they included general information about the patient, some characteristics about sampling methods and the traces of the twelve leads. All the files were kindly provided by Dr. Rafal Baranowski, collaborator of Dr. Antonio Bayés De Luna. In this files there were all different kinds of P-wave morphology which could be present in the ECG signal.

Subsequently, 1000 data were provided by a Company in the Biomedical sector, used to test algorithm's robustness on a large dataset.

The key points of the algorithm are

-Baseline estimation and denoising using sparsity and the Savitzky-Golay filter

-Wavelet Transform

-The MSE and Analysis of P morphology

Once the data has been processed, the six classes, in which the P waves are divided, can be recognized.

# Summary

# List of Figures

# List of Tables

# Abbreviations

| | |
|---|---|
| AV | Atrioventricular |
| ECG | Electrocardiogram |
| IAB | Inter-atrial blocks |
| BB | Bachmann's bundle |
| AF | Atrial Fibrillation |
| DWT | Discrete Wavelet Transform |
| CVD | Cardiovascular Disease |
| HF | Heart Failure |
| CHF | Congestive heart failure |
| LRA | Lateral right atrium |
| CS | Coronary sinus |
| LAE | Left atrial enlargement |
| IACT | Interatrial conduction time |
| PSD | Power spectral density |
| LM | Levenberg–Marquardt |

# Introduction

This study is focused on  the creation of an algorithm for the identification and classification of the P wave.

In detail, this research work aims to find a method that allows for distinguishing two main wave morphologies:

- Normal P wave morphology
- Abnormal P wave morphology, atrial pathology index.

This paper has been divided into seven main chapters. In the first chapter a brief introduction is made on the heart, explaining its main characteristics and functionalities. Furthermore, the main arrhythmias are discussed with more focus on atrial fibrillation and heart failure.

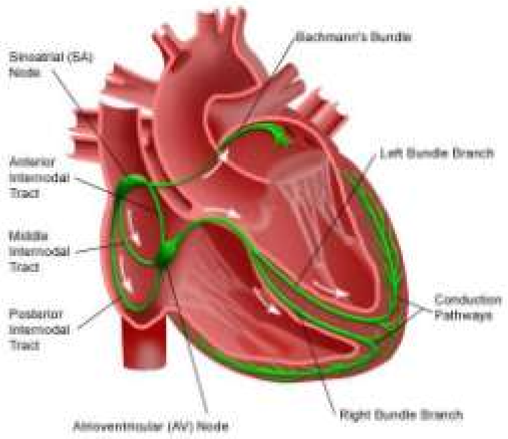
In the next chapter the Bayès syndrome will be introduce, it relates the interatrial block (IAB) to the development of supraventricular arrhythmias. The IAB can be observed through the morphology and duration of the P wave in the ECG. This new discovery leads to the need for P wave arises recognition and classification . In the third chapter it is concisely explained what the electrocardiogram is and the leads which have been used in this study. Chapter four explains the overall structure of the algorithm and which are the main steps. In the following chapters these steps are described with more details. Chapter five explains the method to identify the R wave in the ECG signal. Chapter six analyzes how the median of the beats is calculated and the chapter seven explains how the P wave window is found, which parameters are analyzed for the classification of the wave. In particular, which parameter is most significant for separation between the Normal P wave and the main morphology  subject of research.

The results obtained from the evaluation of the algorithm's performance have been reported at the end.

# Chapter 1

# Heart

In this chapter, there is a brief description of how the heart works and the main heart related diseases, involved in this work.
Heart structure, conduction, and propagation of the electric impulse are discussed for better understanding the physiological problems that it could present. Furthermore, the main arrhythmias which affect the heart are also described.
In figure 1.1 conduction mechanisms representation better explains how the heart operates:




*1.1 Conduction pathways started from the sinoatrial node ( Note that 80–85% of interatrial conduction occurs through Bachmann's bundle)[9]*

Particular attention has been given to the Bachmann's Bundle, that are correlated with the Bayés' syndrome. This syndrome connects the P-waves morphology and its duration with supraventricular arrhythmias insurgence, which common cause of interatrial block.

## 1.1 The heart: Anatomy of the cardiac system

The heart is an involuntary organ that operates like a pump to guarantee the correct flow of oxygenated blood. Heart muscle is shaped as a truncated cone. It is 12 cm high and 10 cm wide. Heart surface presents two ruts, which divide the organ in transversal side, coronary groove, and in right and left heart through interatrial groove. It has a weight between 200 and 300 grams for an adult and it changes based on gender. [3]

It is located in the anterior mediastinum between the two lungs. The heart is externally covered by a membrane, the pericardium, which makes the heart appear shiny. Under this membrane the myocardium and the endocardium are located. The myocardium is the muscle tissue that deals with the real contraction of the heart. The heart is responsible for venus blood oxygenation through the pulmonary district and its redistribution toward the various districts. It is composed by striated musculature, although it is not a voluntary muscle. It is a hollow organ, consisting of four chambers, two atriums, right atrium and left atrium, which are normally not connected to each other, and two ventricles, right and left. They are connected to the atrium through two valves. The right atrium is connected to the right ventricle via the tricuspid valve, they form together the right heart. The main purpose of the right heart is to pump venous blood in the pulmonary circulation. The right ventricle communicates with the pulmonary artery via the lung semilunar valve. The left atrium is connected to the respective ventricle via the mitral valve, they form the left heart. it pumps the arterial blood in the systemic circulation. The left ventricle is in turn connected to the arterial district via the aortic semilunar valve. The atrium are the receiving chambers and the ventricles are the discharging chambers.

The heart contracts about 3 billion times over the time of a subject whose activation starts from the so-called breast-atrial node.

The sinus-atrial node, or sinus node, is an elongated structure located between the superior vena cava and the right atrium, whose pacemaker cells have the capacity to spontaneously depolarize at regular intervals the membrane.

The pulse spreads through the atria and reaches the atria-ventricular node (AV node). The latter constitutes the structure that allows propagation of the impulse between atria and ventricles.

The pulse is subsequently transmitted toward the two ventricles through the His beam and the Purkinie system, respectively. They consist of specialized cells that allow for very rapid pulse conduction.

His beam is divided into two branches, right branch (clearly identifiable in the subjects) and left branch (less organized and with great variability from individual to individual).

The pulse spreads to the atria and reaches the atria-ventricular node (AV node). The latter constitutes the structure that allows the propagation of the impulse between atria and ventricles.
The pulse is subsequently transmitted in the two ventricles through the His beam and the Purkinie system, respectively, consisting of specialized cells that allow to have a very rapid pulse conduction.
His beam is divided into two branches, right branch, clearly identifiable in the subjects, and left branch, less organized and with great variability from individual to individual.
The subdivision of these bands into smaller diameter bands forms the Purkinie system. These districts are highlighted in green in the figure 1.1, in which the Bachmann beam and the internodal traits are also shown in the same way.
Proper conduction occurs when the transmission takes place in an orderly manner through atrial tissue, AV node, His bundle, branches and Purkinie nets. Conductivity is the ability of cells to transmit excitation and affects the excitable cells of the heart. Instead, automatism is an exclusive feature of the pacemaker

cells, i.e. their ability to self excitation, unlike the other myocardial cells. The delay between atriums contraction and ventricles contraction is due to conduction signal slowed down by the sino-atrial node. The delay allows for optimization of blood filling the chamber in the diastolic phase ensuring an adequate systolic throw by the ventricles. The impulse, which arises spontaneously in the sinus node, spreads throughout the myocardium.

### 1.1.1 Bachmann's bundle

In the recent years, the anatomy of the interatrial electrical system was studied in details, with a special emphasis on the Bachmann's region. This because of IAB is directly related to a block in this area.

In 1963 Thomas N. James described three intra-atrial pathways: the anterior, the middle and the posterior tract.[4] The anterior tract is divided into two ways, the internodal traits and the Bachmann's bundle. They leave the sinus node in anterior direction and give off a secondary branch at the superior vena cava level. The sinus node passes through the anterior interatrial band, for this reason it is a fundamental linkage between the two atriums. The internodal traits, instead, continues in the right atrium and finishes in the AV node. Bachmann's bundle represents a faster way to transmits the impulse from right atrium to left atrium.

Some researchers have demonstrated that preferential conduction corresponding to the internodal tracts. A possible reason is related to their significantly higher velocity (1.7 m/second) rather than the surrounding myocardium (0.4 m/second). This was demonstrated *in vivo* , they have assessed the interatrial conduction velocity in a canine heart.

It is a discussion topic the possible reasons why this favourite way, it could due to the presence of specialized conduction tissue or because of the anisotropic orientation of the muscle fibers.

BB across the interatrial path, compared to the distal extensions, is larger,. It has a median measurement of 4 mm in thickness and 9 mm in height. It is described as trapezoidal shaped because of its short lower length (3 mm) and longer upper length (10 mm). [5]

A disorder in the natural anatomy of BB, such as interruption of the parallel orientation of muscle fibers, can cause atrial tachycardias development.

Some studies demonstrated that if BB is interrupted, the wave duration increases. This behavior is represented in a normal ECG with biphasic P wave morphology due to the conduction delay.

## 1.2 Heart diseases: Arrhythmias and Heart Failure

In USA, geriatric patients represent more than 85% of all the deaths caused by cardiovascular diseases. The mortality of cardiac arrhythmia and Heart failure (HF) are much higher in the elderly population (people aged ≥ 65 years) that the younger population. HF is caused by the inability of the heart to perform the normal contractile function of pumping sufficient blood which satisfies the systemic metabolic needs. It is not always easily recognised because in some cases the disease can be asymptomatic. HF usually develops following a cardiac muscle injury, for example as a result of a myocardial infarction, excessive cardiac stress due to untreated arterial hypertension or as a result of chronic valvular dysfunction. The electrocardiogram of many patients suffering from heart failure shows that an alteration of the heartbeat propagation causes modifications in the cardiac mechanical contractile activity. This provokes a dissolution of contraction and therefore a worsening of the contractile capacity of the heart.[*sit 2*]

Atrial fibrillation (AF) and congestive heart failure (CHF) are frequent co-morbidity in the elderly population. The probability that AF precedes CHF or the other way around is similar.

The main arrhythmias are described below, more focus is given to the AF.

Arrhythmias are type of heart diseases, which concern heart rate variations out of the normal range (60 to 90 beats per minute (bpm)). This physiologic range can change based on gender, age and health state. [6] Arrhythmias are divided in two main groups, tachyarrhythmias and bradyarrhythmias. In the first one the heart rate becomes lower than the normal heart rate(approximately 70 bmp). The other one is characterized by a heart rate. Arrhythmias can sometimes endanger people life, especially if a myocardial dysfunction is present in the subject. This pathological condition can prevent the heart from a correct cardiac blood output in the presence of bradycardia or tachycardia. In the case of bradycardia the frequency is less than 60 bpm. In athletes and young subjects it is a physiological condition but in some others it may be a warning of the presence of a sinus node disease. In case of a sino-atrial block, there is a conduction alteration between the sinus node and the atria. The problem of bundle branch blocks is linked to the evolution towards a progressive compromise of the conduction system that can lead to the complete AV block.

There are three types of sino-atrial block. In the first case, called first degree, there is a slowing down of the conduction which is not detectable by the ECG. The second-degree block is splitted into type 1 and type 2: the former is characterized by P wave missing (equal to twice the duration of a P-P cycle), instead the latter consists of an elongation of the sino-atrial conduction time. In the last category, atrial activity is absent.

When there is an abnormal diffusion of the pulse between atria and ventricles, an atrioventricular (AV) block is present. Also in this case there is a subdivision in three degrees of block. In the grade I block, there is a slowing down of the conduction, and therefore, a longer duration of the P-R interval. The grade II block is divided into type 1, in which there is a progressive increase in the R-P interval, on the contrary the P-P interval is constant, and type 2 in which the pulse does not reach the ventricle alternately. The third-degree AV block is known as complete since there is a cessation of conduction.

The branch blocks, which concern conduction through his beam, can be complete or incomplete. there is an alteration at the level of one of the two slings, right or left, which can be a total absence of conduction in the case of complete blocking or a slowing down in the second case.

Tachyarrhythmias, in contrast to those described above, consist of an increase in heart rate that exceeds the maximum value of the reference range of 100 bpm. Among the main tachycardias that can occur in a subject are the extrasystoles, which are characterized by a premature beat that rises in different points compared to the sinus node. in this way the regular heart rhythm is lost. the extrasystoles can be atrial or supraventricular, which are divided into ventricular and junctional.

In the case of an atrial extrasystole going to analyze the P wave it can be observed some differences with respect to the wave p in a normal sinus rhythm. The P wave appears premature, besides if negative in V1 there will be an involuntary involution of the right area of the heart, while if it appears negative in D1, D3 and aVF is a probable extrasystole of the left atrium.

Junctional extrasystoles are less common than the atrial and ventricular ones and depending on where it rises, it can spread to the upper part of the heart or to the ventricles.

It may involve both districts, observing the ECG we note that there is a premature QRS complex not preceded by P wave.

From the hemodynamic point of view, the ventricular extrasystole stops the diastolic filling of the ventricles and prevents a correct ventricular filling volume due to the atrial systole, the systolic stroke will be reduced. On the contrary, the range following the extra-systolic event will have a greater range.

Unlike the extrasystole when it comes to tachycardia, it refers to a heart rhythm greater than 100 bpm. Based on the site of onset it can have:

- Sinus tachycardia, borns in the atrial sinus node and is generally a physiological response

- Atrial tachycardia, presents P waves identical to each other but different from the physiological ones. The frequency is between 130-220 bmp.

Paroxysmal supraventricular re-entry AV, in which there is a return of the pulse that, in addition to passing from the atrium to the ventricle, also spreads in a retrograde direction until to reach the ventricles again and cause tachycardia.

Among the diseases of the circulatory system it can also be found atrial and ventricular flutter, depending on which of the two districts it gets involved. There is a regular rhythm in which the atrium / ventricle, but with an irregular and faster activation.

A more detailed description of atrial fibrillation is necessary, because this pathology is the most common sustained cardiac arrhythmia and is linked to the onset of major illnesses, such as heart failure and stoke. [7]

Over 6 million Europeans suffer from this arrhythmia, of whitch 600,000 Italians, and its prevalence is estimated to at least double in the next 50 years because the population is getting old. [*sit* 5] Atrial fibrillation (AF) is becoming progressively more prevalent with population aging. Generally occurs in the presence of structural defects and alterations of the walls of the chamber. As regards atrial fibrillation, the atrium are activated chaotically and with frequencies of 400-650 bpm. This is caused by the formation of some re-entry circuits. The irregular contraction of the heart can occur very slowly or rapidly. Fibrillation Atrial causes, in fact, stroke and heart failure, but also bradycardia and tachycardia. [8]

AF can be classified based on the duration of the arrhythmia episodes and not referred to a single episode, because it cannot be fully categorized. Classification of clinical AF subtypes is:

-"Paroxysmal" AF refers to episodes that generally terminate after no more than a few days.

-"Persistent" AF, in which a cardioversion is necessary to reactivate sinus rhythm.

-"Permanent" AF cannot be converted to sinus rhythm.

The onset of AF derives from various causes, among the most frequents the myocardial infarction (from 5% to 10% of patients with infarction) and cardiothoracic surgery (up to 40% of patients), hypertension and congenital heart diseases.

AF provokes electrical remodeling that favors the occurrence of focal triggering mechanisms such as multiple wavelets re-entry. A problem in the atrial conduction can be transferred on the left ventricular, causing inappropriately a rapid and irregular ventricular rhythms and at the end the patient heart failure.

The increased mortality associated with AF is in part attributable to stroke, the main problem remains the disability caused by stroke derived from atrial fibrillation.

# Chapter 2

# Bayés' syndrome: Correlation between IAB and supraventricular arrhythmias

As described previously interatrial block (IAB) is a conduction problem that prevents or delays of electrical activation between the atrium. Bachmann's Bundle dysfunction is the cause of this kind of disease. It is the preferential path for the propagation of electrical impulse through the atrium. Some anomalous referred to the connective tissue, as an increase of fibrous tissue due to ageing and subsequently loss of elasticity, can cause an errata conduction in the Bachmann's Bundle, that is directly related with IAB.

Deterioration of electrical conduction from right atrium to left atrium is a disturbance of conduction that is given much importance and taken into consideration by researches.

For a long time anyone has considered a correlation between interatrial block and supraventricular arrhythmias. In addition , in some cases, the interatrial block is often confused with an LAE [9], and it has not always been correctly understood that they are two separate entities. A correct classification of the P wave, not only from the analysis of the duration but also of the morphology, simplifies the division of the two pathologies. It has been observed that an enlargement of the left atrium cuases a P-wave duration variation, which will be greater than 120 ms, but it presents a different morphology in the leads at the contrary respect to the IAB.

The P-wave morphology in V1 is an help to distinguish better the two different pathologies; the lead often presents a P-wave negative mode that is less evident than in cases of LAE. The interest of studying and identifying the IAB has increased in recent years, because of IAB is linked to some diseases may cause deadly effect. A large number of researches have been published regarding the prevalence of IAB and its associations with the risk of AF, ischemic stroke, and cognitive decline. The presence of first-degree IAB and so of prolonged interatrial conduction time(IACT) is correlated to an increased PR interval duration, besides an extension of P-wave duration.

A study, for demonstrating the relationship between P-wave morphology and duration, was conducted by the Geriatric Cardiology Section of the Spanish Society of Cardiology [10], the International Society of Electrocardiology, and the International Society of Cardiovascular Pharmacotherapy, on three groups of patients. The first one presents normal P-wave  morphology, the second one partial IAB and the last one advanced IAB. . Moreover the prevalence of interatrial block (IAB) is high in the elderly. For this reason the aim of the research is to assess the impact of IAB on the risk of AF and stroke during 3 years of follow-up in the geriatric patients. A series of patients aged ≥70 years are examined.
Through a standard ECG analysis and a depth study of P-wave morphology and duration, the results have been demonstrated the hypothesis that IAB is a significant risk factor for AF and stroke.

The news introduced by Bayés [11]is about the correlation between IAB and supraventricular arrhythmias, in particular AF, the development of this one is due to the third-degree IAB onset.  For this reason this disorder termed 'Bayés syndrome'. Bayés divides the atrial blocks into inter- and intra- atrial blocks: in the first one the conduction disorder is between the two atrium, while in the second one is inside the same atrium. Furthermore, he divides atrial blocks into IABs and others atrial block, he gives more importance to the first kind of pathology because of its association with AF and stroke. In particular he considers that the IAB is a delay to the propagation due to a conduction problem in BB.

A study to demonstrate the relationship between P-wave morphology and duration was conduction by the Geriatric Cardiology Section of the Spanish Society of Cardiology, the International Society of Electrocardiology, and the International Society of Cardiovascular Pharmacotherapy. The morphology is divided in three groups, the first one presents normal P-wave morphology, the second one partial IAB and the last one advanced IAB. Through a standard ECG analysis and a depth study of P-wave morphology and duration, the results have been demonstrated the hypothesis that IAB is a significant risk factor for AF and stroke.

## 2.1 The power of P wave with ageing

As previously mentioned, atrial fibrillation is one of the most common cardiac diseases. Several studies have shown an correlation between the progress of age and the frequency which this arrhythmia arises. A study was conducted in Australia [12] with the aim to verify the relationship between aging and AF. The study was performed on three groups of patients, based on age. There is a significant increase of the duration of the P wave passing from group C, made up of younger subjects, to group A, in which the components are older than 60 years.

The study conducted on animals demonstrates how a remodelling of the atrium is due to ageing, there is an anatomical and structural change; the conduction between atria is slowed due to the presence of interstitial fibrosis. Due to ageing, the heart muscle tissue changes, and there is a progressive replacement of the elastic tissue with the fibrous tissue. Hayashi et al. [13] confirm the theory according to which both the interatrial conduction and the duration of the P wave increase over time together with the increase in the probability of the onset of AF.

# Chapter 3

# ECG: electrocardiogram

In this chapter, it is described the electrocardiogram, a types of exams useful as it allows an in-depth analysis of the cardiac cycle, allows to recognize any pathologies present in the cardiac system and remains completely external and non-invasive for the patient. It requires the use of a variable number of electrodes, from a minimum of four to a maximum of 10, through which the biological signal from the heart is taken.

## 3.1 Electrocardiogram

The electrocardiograph is a galvanometer, an instrument that measures intensity current, in this case the signal picked up by the heart electrical activity.

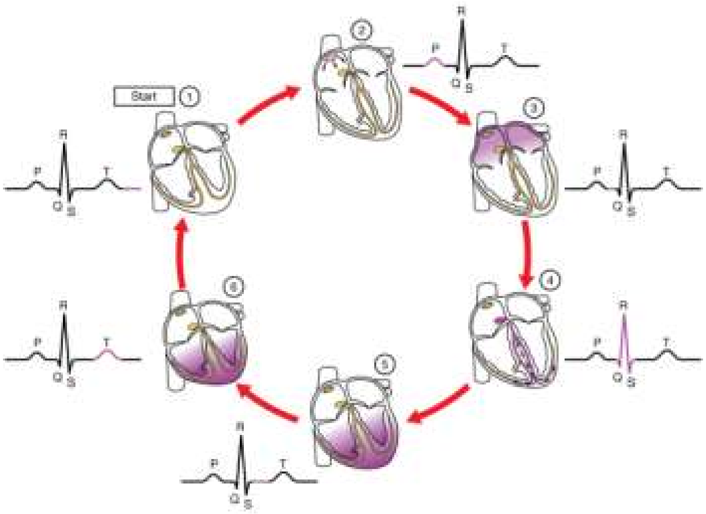
The electrocardiograph is a not invasive exams and gives a lot of useful and important information to understand better the patient's healthy state and the presence or not of heart disease without creates no complaint to the patient.

This machine gives back an electrocardiogram, ECG machines generally can register several potential differences at the same time according the location and number of electrodes located on the body. The ECG is represented on a millimeter paper, each square has a side of 1mm, that slides with a 25 mm/s standard velocity; in this case 1mm corresponds in time to 40 ms. (Figure 3.1). [14]

Usually in the recording of ECG there is, in the first part, a square wave that is used to give a calibration, having a note amplitude to 1mV. For this reason the great interest shall be gived to the calibration. The heart electrical conduction can be projected into different lines. The lead in electrocardiography is a line that connects two observation. Each lead reveals the magnitude of the electrical conduction in the direction of that lead at each instant of time.

It is important to first understand the physiological basis of the ECG and the principal characteristic of standard ECG to be able to find and extract the useful information. Each heart cycle is constituted by the same depolarization/repolarization period, before there is atrium depolarization following by ventricles depolarization both the atria repolarization and for last ventricles repolarization. ECG signal is a pseudo-periodic signal, pseudo is referred to the link between cardiac cycle to heart rate, that establishes the first one. The standard ECG is characterization by a sequence of waves, marked by the letters P, Q, R, S, T and U, each of which represents a particular heart phenomena. The baseline voltage of the electrocardiogram is known as the isoelectric line, that generally is measured as the portion of the pathway between the T wave and the next P wave.

The figure 3.1 explain the heart activity and its link with the ECG waves:



*3.1 Cardiac Cycle: correspondence between heart activity and electrocardiographic track (https://brilliantnurse.com/nclex-cardiac-electrical-conduction-system)*

*)*

In detail it shows:

-P wave, is the first wave that emerges in a cardiac period. It represents the atrial depolarization and so the initial of cardiac period. Its duration, generally, is less than 120 ms, between 50 and 120 ms with an amplitude of about 50-150 μV ( about 2.5 mm). It is possible to distinguish two different components, the first corresponds to the activation of the right atrium and the second to the left atrium. The P wave is composed of two main vectors referring to the depolarization of the right atrium and the left atrium. More details about P-wave will be given in the next paragraph.

-QRS complex, reflects the depolarization of the right and left ventricles and is the most prominent feature of the human ECG. It is composed by three waves that characterize the ventricular depolarization. Q and S are negative, which Q is the first wave in the complex, R is represented with positive deflection and it is the wave with greater amplitude, between 500 μV and 3 mV; as a rule, it is used to determinate heartbeat. It has a duration between 60 to 100 ms. The interval from the beginning of the QRS complex to the apex of the T wave is referred to as the absolute refractory period.

-T wave, the last wave in the cardiac cycle and represent the ventricular repolarization.

-U wave, it has a low amplitude, and frequently is completely absent. Probably it is caused by the repolarization of the interventricular septum.

The ECG is moreover characterization by intervals and segments . This parameters is necessary to distinguish the normal or abnormal space between two electrical events.

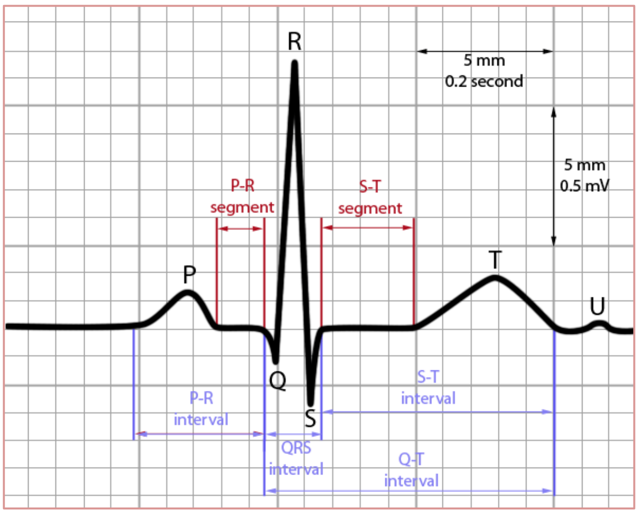
They can be summarized as following:

• PR interval is measured from the beginning of the P wave to the beginning of the QRS complex. The PR interval reflects the time the electrical impulse takes to travel from the sinoatrial node through the A-V node and entering the ventricles.

• PR segment connects the P wave and the QRS complex. This coincides with the electrical conduction from the A-V node to the His bundle and the bundle branches and then to the Purkinje Fibers

• QT interval is measured from the start of the QRS complex to the end of the T wave. This is the time from the beginning of the excitation of the ventricles to the end of their relaxation.  QT interval should be analyzed because a prolonged QT interval  is a risk factor for ventricular tachyarrhythmias and sudden death.

• ST segment connects the QRS complex and the T wave. The ST segment represents the period when the ventricles are depolarized. Since there is a priori no electrical propagation the segment is isoelectric. • ST interval is measured from the end of the QRS complex (J point) to the end of the T wave. • RR interval is measured between two peaks of successive R waves and represents the cycle of ventricular repolarization. It is associated with the cardiac period.




*3.2 Normal ECG track: PQRST sequence with the different intervals reported (http://www.medicine.mcgilf)*

## 3.2 The 12 Lead

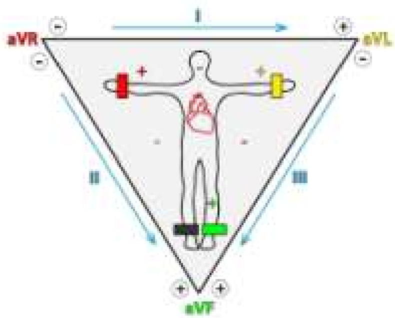
The standard 12-lead ECG system is composed as follows:

• 3 Einthoven bipolar limb leads: I, II and III,

• 3 unipolar limb leads: aVR, aVL and aVF,

• 6 unipolar precordial chest leads: V1, V2, V3, V4, V5 and V6.

This kind of ECG signal detection allow to have a three-dimensional representation of the electrical activity of the heart. [15] Willem Einthoven discovered that only three electrodes are sufficient to monitor depolarization of the heart.




*3.3 Einthoven triangle (http://www.ippocrateshop.com/blog)*

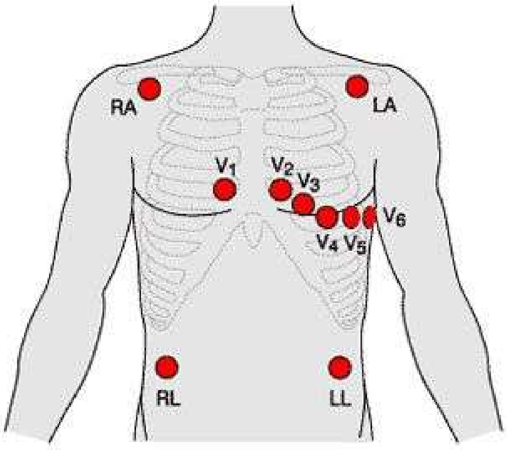
This electrodes are attached on each arm and one on the left leg. The standard limb leads are:

• Lead I: potential difference between the left arm (positive electrode) and the right arm (negative electrode),

• Lead II: potential difference between the left leg (positive electrode) and the right arm (negative electrode),

• Lead III: potential difference between the left leg (positive electrode) and the left arm (negative electrode). It is possible to calculate as a difference between the others.

The figure shows the position of the electrodes for this leads and the leads that are described below:




*3.4 Electrodes position: there are 10 electrodes to obtain 12 leads, RL is the reference electrode.(http://www.firstaidforfree.com/recording-a-12-lead-ecgekg)*

At a later time limb leads and six chest leads has been added.

It is sufficient, of all the 12 leads, 10 electrodes; in fact 6 leads derive from the six precordial electrodes and the other six derive from the 3 peripheral electrodes, three are bipolar leads and three are unipolar. The last electrode is the reference electrode, which is considered to have zero voltage.

The three bipolar leads are obtained as follows:

• aVR : potential difference between the right arm (positive electrode) and the center of the heart,

• aVL : potential difference between the left arm (positive electrode) and the center of the heart,

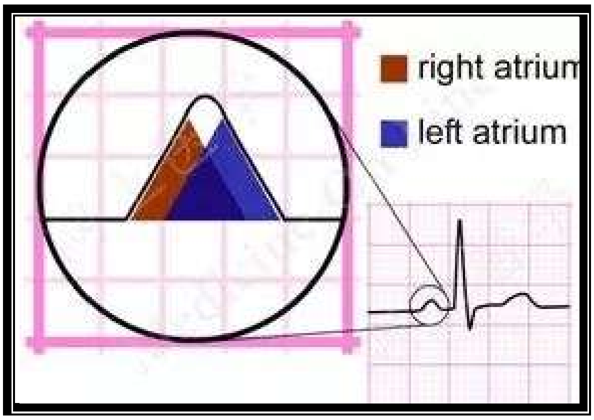
• aVF : potential difference between the left leg (positive electrode) and the center of the heart.

In the precordial leads the ECG signal is obtained from the difference between each of these electrodes (V1-V6) and the central terminal.

## 3.3 P-wave characteristics

P wave is related to the atrial depolarization. The main electrical vector is directed from the sinoatrial node toward the A-V node, and spreads from the right atrium to the left atrium. The first part of the P wave derives from the depolarization of the right atrium, the first to activate, while the second from the depolarization of the left atrium. A wave, for each depolarization, is associated. In a no pathologic heart the two activations take place almost simultaneously. There is an overlap of the two waves.

The P wave generally has small wave amplitude, less than 2.5 mm in the limb leads and 1.5 mm in the precordial leads, with a duration less than 120 ms. This duration denotes a good electrical conduction between the two chambers. P waves should be upright in leads I and II, inverted in aVR, and sometimes biphasic in lead V1.[*sit 3*].
In this case the right and left atrial waveforms have opposite direction.
In a normal conduction of the heart the P wave has a monophasic shape, always positive in lead II, as shown in the figure:




*3. 5 The component of the P wave given by the right atrium is colored in blue and the component of the left atrium in brown, under normal conduction conditions, the sum is a linear wave, (https://lifeinthefastlane.com)*

It is generally also positive in AVL, AVF, I, V4, V5, V6 and negative in AVR.[*sit* 4]

The P wave morphology is influenced by some parameters:

-the origin of the sinus rhythm

- the passage of conduction in the interface between the two atria

- the size of the chamber, on which the delay in completing the depolarization of the chamber depends.

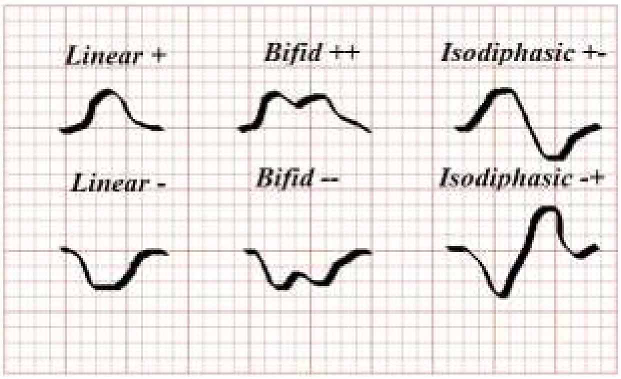
The changes, caused by pathology, disease or, as explain before, by aging, in the electric activity modifies the morphology of the P-wave.

If an enlargement of the right or left chamber occurs, there is a change in the morphology of the P wave. Atrial abnormalities are most visible in the leads II, III, aVF and lead V1, as the P waves are most prominent in these leads.

More recently studies, has been shown that a prolonged P-wave duration is a marker of incident AF. A study conducted by Copenhagen University Hospital [16], included 285,933 individuals, confirmed the correlation between a longer duration of the P wave (>120ms) and the AF. Furthermore, the study investigated the association between the duration of the P wave and the risk of death due to cardiovascular causes and presumed ischemic stroke causes. At the end of the observations on the patients, on December the 31st, 2011, if not previously ended for other causes such as the patient's death, the connection between AF and P-wave duration, and between P-wave duration and ischemic stroke is confirmed.

Furthermore, a study conducted by the MAVERIC, in Boston [17], showed that the presence of a dome followed by a spike can be used as a marker to identify the IAB. The end of the research on approximately one hundred patients showed that 94% of those affected by IAB had a P wave with the morphology described above.

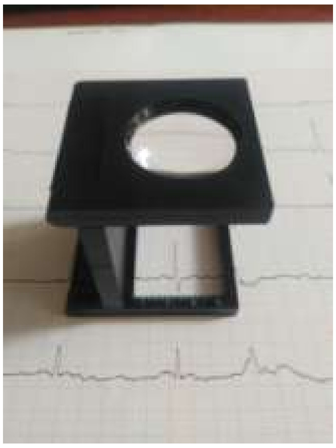
In this work the Company wants to recognize six P wave classes with the morphologies shown in the figure:




3.6 The six classes in which the P wave morphology is divided

This request derives from a study conducted at the Care-Lab, Rome on a dataset of patients who have been studied to observe the morphology of the P wave. The satisfactory results on the relationship of the morphology with the insurgency of deadly pathologies could not be reported.

The individuation and classification of the P wave occurs as shown in the figure in a completely manual way:



3.7 Method to the manual localization of P wave in the ECG track

A millimetric magnifying glass is used to observe the P wave, and identify Onset and Offset in more simple way. One of the six classes was assigned manually by the operator. This method requires a large amount of time, hence the idea of making the identification and classification of the P wave automatic.

# Chapter 4
# Overall structure of the Algorithm

The algorithm was implemented with the aim of identifying and subsequently classifying the P wave obtained from the median of the beats present in the recording. The idea is born from the desire to create an automatic method that replaced the work done manually by health professionals.

As described previously, to identify the presence of any anomalies in the morphology of the P wave, an operator locates the window containing the P wave and assigns a class.

The purpose is to define an Onset and an Offset for delimiting the P wave window automatically.

Some intermediate steps were necessary starting from the original signal to obtain the median, from which the P wave window is located.

The main steps of the algorithm are as follows:

-Reconstruction lead

-ECG filter ( Baseline Remove ,Savitzky-Golay filter)

-R wave detection (Wavelet Transform)

-Median Calculation

-P wave capture

-P wave classification

In the following chapters, the approach used in the steps will be described in detail. A dataset of 1,000 signals was provided, used for verification and validation the algorithm.

## 4.1 Reconstruction Lead

1,000 ECG recordings have been supplied by the Cardioline company to verificate the algorithm on a sufficiently set of data. The signals are sampled at a frequency of 500 Hz, each signal has 5000 samples, 10s of recording. Each data file contains the 12 ECG leads; in some cases, the leads, that derive from others, are not present in the file. Due to this defect, if the leads are null, it was necessary to reconstruct them. A Reconstruction Lead function was created for reconstructing the missing leads.

The following formulas are used, depending on which one is absent:

-Lead III

$$III = II - I \tag{4.1}$$

-Lead AVR

$$AVR = \frac{-(I + II)}{2}$$

-Lead AVL

$$AVL = \frac{(I - II)}{2}$$

-Lead AVF

$$AVF = \frac{(II - I)}{2}$$

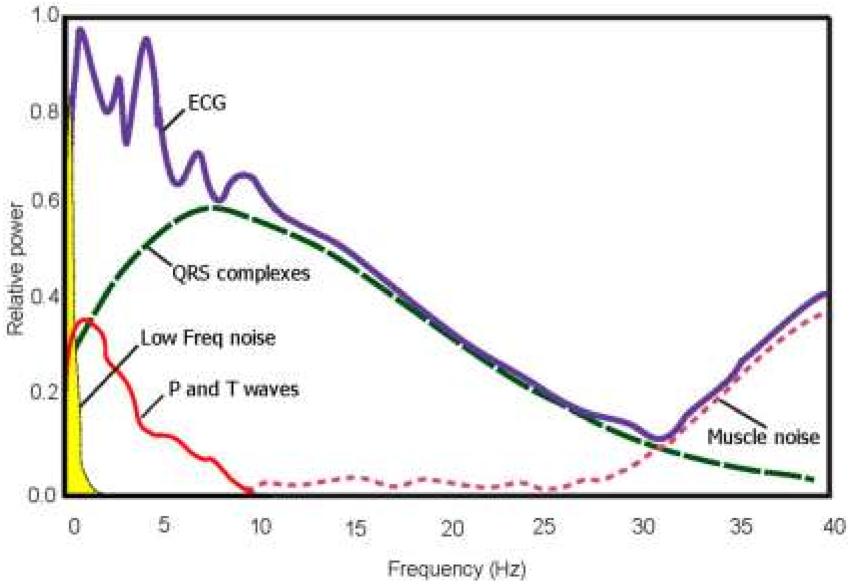
## 4.2 ECG filter: Introduction

During the acquisition of the ECG signals, it is problematic to have recordings without artefacts and noise. The ECG signals given are deteriorated by noise, so pre-processing is necessary to obtain a final signal that could be processed.

Figure 4.1 shows where the main frequency components of the ECG signal are located. The QRS complex extends up to 35-40 Hz, while the P wave and the T wave have lower frequencies, maximum 11 Hz. There may be low frequency noises that cause slow drift in the signal, moreover movement artifacts that have the frequency components allocated in bands higher than components of interest.

The final signal could lose some own features.

The ECG signal, unlike other biomedical signals, presents a repetitive pattern that distinguishes it from the baseline wandering and the noise. It is important to keep its main pattern in the filtering.




*4.1 Distribution in frequency of the ECG components and noise, it is noted that the ecg has the majority of the components of interest at frequencies lower than 25 Hz (https://res.mdpi.com)*

The first step of the algorithm consists in removing the slow drifts, that do not allow to identify the baseline on the signal, and the noise present in the recording.

Reducing noise and baseline estimation are critical operations in the ECG.

A special attention must be given to the pre-processing, in order that signal's characteristics and its morphology are preserved.

The traditional filtering approach, in many cases, could eliminate not only the noise, but also the frequencies of interest , considering that the complex QRS frequency range is extended in a wide band, between 0 and 20 Hz [18] [19] .

Among the various methods that are found in the literature for the removal of the baseline, based on the Wavelet or Fourier Transform, a new approach, Empirical Mode Decomposition (EMD), has been introduced. [20]

The EMD approach decomposes the signal into several components, used sparse and redundant representation model.

The advantage of this method is that removes the frequencies given by the noise and the baseline wandering, without creating a significant distortion in the signal.

### 4.2.1 BEADS (Baseline Estimation and Denoising Sparsity)

The algorithm used for the baseline removing was taken from a study done on the Chromatographic signal. [21] The purpose is to both reduce noise and eliminate the baseline. The idea is to decompose the complex signal into sufficiently distinct components. In this case the signal is decomposed recognizing three main components:

- the baseline, which is approximated to a low-pass signal.
- the peaks of the signal, the second and third order derivatives, as sparse signals.
- random noise, considered as a residual.

This method, called BEADS (baseline estimation and denoising sparsity), employs the Majorization-Minimization (MM) approach.

The algorithm, suitable for use on long data sets, has been formulated as a convex problem; this ensures that the algorithm always converges to a single solution.

The following solution is sought:

$$x^{(k+1)} = \arg\ \min_x G(x, x^{(k)}) \tag{4.2}$$

Where k is the iteration counter. For each iteration a certain condition must be verified. In particular a convex optimization problem must be solved. It is a problem in which the objective and constraint functions are convex, which means they satisfy a inequality.
In BEADS is:

$$G(x,v) \geq F(x) \qquad for\ all\ x$$
$$G(v,v) = F(v) \qquad for\ x = v \tag{4.3}$$

The signal is modeled as the sum of three components, described above:

$$y = x + f + w \tag{4.4}$$

Where:
*y= signal given;*
*x= sparse-derivative signal;*
*f= baseline wandering;*
*w= noise;*
To estimate x the following optimization equation is used:

$$x = \arg min_x \left\{ F(x) = \frac{1}{2} \|\boldsymbol{H}(y-x)\|_2^2 + \sum_{i=0}^{M} \boldsymbol{\lambda}_i \sum_{n=0}^{N-1} \boldsymbol{\varphi}([\boldsymbol{D}_i \boldsymbol{x}]_n) \right\} \tag{4.5}$$

As the algorithm has been thought, it is possible to select some initial parameters. the $\boldsymbol{\lambda}_i$ , that are present in the formula 4.5, are regularization parameters, a high value of these makes $\boldsymbol{D}_i\boldsymbol{x}$ more sparse. Furthermore, the parameter $\boldsymbol{r}$ could be selected. This constant defines the penalty function $\boldsymbol{\varphi}([\boldsymbol{D}_i\boldsymbol{x}]_n)$.

Depending on the selected of $\boldsymbol{r}$ value it could be possible to choice an asymmetric or a symmetric penalty functions. The difference is that, in the first case a different penalty is given between the positive and negative values, in the second case the penalty is independent of the sign of the x values .

The final formula, that in this work it is used ,is defined as follows:

$$F(x) = \frac{1}{2}\|H(y-x)\|_2^2 + \lambda_0 * \theta(x) + \lambda_1 \sum_{n=0}^{N-1} \varphi([D_1 x]_n) + \lambda_2 \sum_{n=0}^{N-1} \varphi([D_2 x]_n) \tag{4.6}$$

In the paper [18] there are several definition of the differentiable symmetric penalty function $\varphi$ . In this work, it is decided to describe this function as:

$$\varphi(x) = |x| - \varepsilon * \log(|x| + \varepsilon) \tag{4.7}$$

While the penalty function $\theta(x; r)$ is identified as following:

$$\theta(x; r) = \begin{cases} x & x > \varepsilon \\ f(x) & |x| \leq \varepsilon \\ -rx & x < -\varepsilon \end{cases} \tag{4.8}$$

$\varepsilon$ is equal to 1*10-6 . f(x) is a fixed second-order polynomial function used to consider the case x=0, more in detail when x has values between $\pm\varepsilon$, a range around zero.

Taking into account $\Theta$ , the asymmetric penalty function, if r=1 it becomes $\Theta(x; r) = \varphi_A(x)$, the penalty function is symmetric.

In this specific case, it is chosen:

Number of Iteration= 30;

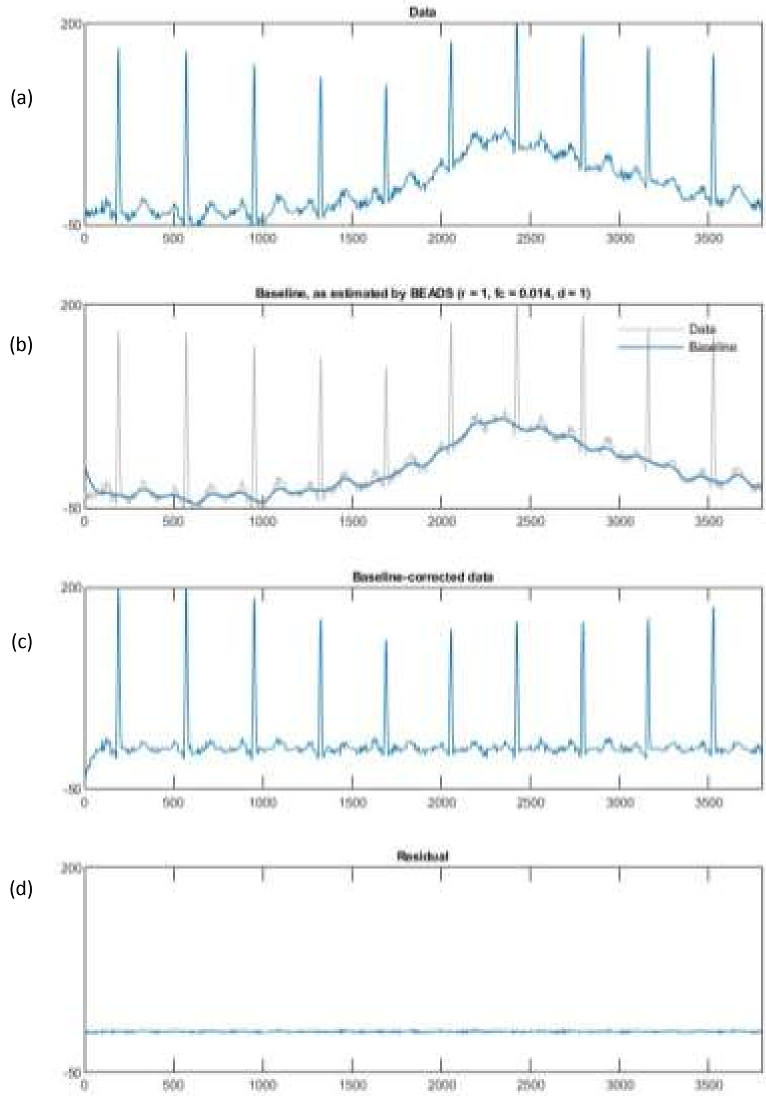
$r$=1;

$\lambda_0 = 0.1;$

$\lambda_1 = 1;$

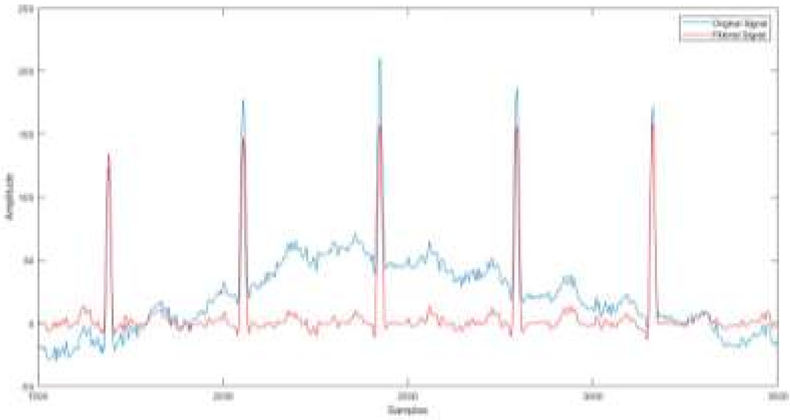
$\lambda_2 = 0.8;$

As shown in figure 4.2, the original signal is affected by slow drifts; subsequently, with the application of the BEADS method, the baseline wandering is identified and removed. The residual noise is also reported in figure 4.2, obtained by subtracting the signal and the baseline given at the end of the 30 iterations.




*4.2 (a) original signal affected by slow drifts (b) highlighted in blue the baseline found by the algorithm (c) Filtered Signal (d) Residual is equal to the difference between the original signal and the sum of the found baseline and of the filtered signal*

Subsequently, after the extraction of the baseline wandering, the original signal was compared to the signal returned by the algorithm. This step is necessary to verify that the morphology is not modified and that only the slow drifts is eliminated by the algorithm. The result can be observed in figure 4.3



*4.3 Comparison between the Original Signal and Filtered Signal, the morphology remains unchanged*

A further instrument that allows to verify the filtering operation is the PSD. The signals have a repeatability pattern, but their variation in the future cannot be known exactly . Since what it could be calculated is an estimate of the power. It is talked about the spectral density of power. The discrete time signal is considered to have E{x(t)}=0, [22] a mean value equal to zero. The mean value for all signals, in fact, has been removed. The PSD is calculated starting from the formula:

$$P_D(f) = \frac{1}{2P+1}\left|\sum_{n=i-p}^{i+p} P_{xx}(f_n)\right|^2 \tag{4.9}$$

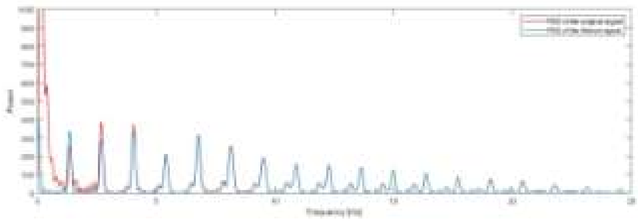
Where

$$P_{xx}(f_n) = T * \sum_{n=-M}^{M} x[n] * e^{-j2\pi f nT} \tag{4.10}$$

A rectangular window equal to half the length of the signal was considered.

The choice of the window's length, in general, is done by evaluating some parameters, such as: spectral resolution and statistical variance ; its shape should be based on a flatten and leakage effects. In this case the representation of the PSD has the unique purpose of showing qualitatively the action of the applied filter. For this reason the use of a Daniell periodogram , without the overlap and with a rectangular window, such as described before, is sufficient.

It is possible to observe, comparing the power spectrum in the two cases , as shows in figure 4.4, that the low frequency components, characteristics of the lens drifts, are deleted.



*4.4 Comparison between the PSD of original and filtered signal, the PSD of signal after filtering lost low frequency caused by the noise*

## 4.2.2 Savitzky-Golay filter

The Savitzky-Golay filter is a filter based on the method of least squares. This filter is a FIR, that is, has a finite impulse response. It was proposed for smoothing and calculate derivates of all kinds of data.

It is suitable for the application in this field because there is an improvement of the information contained in the data, without a great computational complexity. [23] The mathematical operation performed by this filter is described [24] by the following formula:

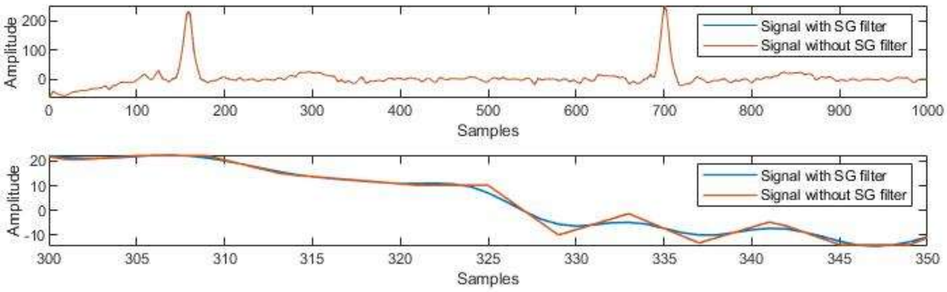
$$Y_j^* = \frac{\sum_{i=-m}^{i=m} C_i Y_{j+1}}{N} \tag{4.11}$$

Y indicates the original ECG signal and Y* is the output of the filtering . $C_i$ is the coefficient for the i-th signal value of the smoothing window, N is the smoothing window length and it is used to normalize the result. Index j takes values from 1 up to the signal length. [25] The filter creates a polynomial whose coefficients are calculated using the least squares method. The degree of the polynomial is given by the chosen order.

Furthermore, it is an applicable filter for treating the ECG signal ,as it does not change significantly the morphology of the signal.

Two input parameters are required. The first parameter, **f**, defines the filtering window. The smoothing window length must be odd. Generally, a too high value of the first parameter increases the smoothing effect, but flattening the peaks in the signal. In the following case, the frame length is set equal to 11. It is obtained a matrix B of the polynomial coefficients of size fxf. The second parameter, the polynomial order, must always be less than the frame length. It is an integer specifying the degree of the smoothing polynomial,which is typically set in a range from 2 to 4. In this specific case it is pre-set equal to 3.

The figure 4.5 shows the effect of the Savitzky-Golay filter in this application, an enlargment on the image is realize to make the smoothing effect of the Savitzky-Golay filter more understandable:




*4.5 Comparison between signal with and without SG filter*

After the application of both filters the signal noise ratio has been calculated as follows:

$$SNR = 20 * \log_{10}\left(\frac{V_{pp}}{4 * \sigma_n}\right) \tag{4.12}$$

Where $V_{pp}$ is the peak-peak value of the signal and $\sigma_n$ is the noise deviation standard, calculated considering the isoelectric line. The isoelectric line should be null but in reality there is fluctuations around zero due to the noise.

The SNR is an index about how much noise has been reduced compared to the signal. Ideally, it should be infinite, that is, the noise is absent. This, obviously, is impossible to achieve.

In this field the signal is considered as a deterministic signal and the noise as a random signal. So, in the first case, the peak-peak value is calculated and in the second one the deviation standard is considered, multiplied by 4.

The result obtain indicates an increase equal to 20% of SNR.

SNR=18.19 dB *pre filtering*

SNR= 24.16 dB *post filtering*

# Chapter 5

# R wave detection method

## Introduction to R wave detection method

In the state of Art there are different methods for the identification of the R wave.
An algorithm for the detection of ECG characteristic points [26] was implemented by Cuiwei Li and Chongxun Zheng. 600 ECGs were analyzed using the Wavelet Transform, taking into consideration the levels from 1 to 4. This choice is make because for scales equal or greater than $2^5$, the energy of artefacts and noise increases and there is a reduction of the QRS complex energy. Two controls have been introduced, which will be resumed in the following chapters and defined as First Control and Second Control in the present work. The first check consists in delimiting the permit distance of two QRS waves consecutive, set in this case equal to 200 ms. The second check consists to reduce the threshold if there is an R-R interval greater than 150% of the average R-R interval. For the detection of the T and P wave the scale $2^4$ and $2^5$ of the wavelet transform is considered and two windows is used, before and after the detected R peaks. The results showed that 99.8% of the QRS complexes are correctly detected.
A most recent research by the Engineering College Salem, in India, [27] analyzed the ECG signal through the Discrete Wavelet Transform for the feature extraction. The Db4 (Daubechies Wavelet) was chosen among several DWT's families. Before starting with the application of the DWT, a pre-processing was performed, using a median filter, to remove the baseline. In this study it is considered the scales from 1 to 4, because in the scale 4 there is the most energy of the QRS complex. The threshold method on the DWT reconstruction is used in the detection of the R wave peaks, considering that it represents the highest peak in the lead. Also in the study conducted by the Department of Electrical Engineering, at the National Tsing Hua University Hsinchu, in Taiwan, [28] the DWT was used to Noise reduction and R wave detection. They use a combination of DWT and SWT (Stationary Wavelet

Transform). In details the sym5 (Symlet) was chosen among the possible wavelet families. Levels 3 to 5 were considered, which mostly contain the frequency components of the QRS complex. After the reconstruction with the DWT, the Threshold method was used to detect the R peaks. A Control is used to verify if there are some lost R waves. In fact if the interval between two consecutive R waves is more than a set value, a different Threshold is used. This algorithm gives good result, both as regards the Sensitivity(99.70%) and a Positive Predictivity (99.65%).

Another research on the ECG signal, conducted by J.Martinez [29], using again the scales from $2^1$ to $2^4$ of the DWT for the R wave detection have returned good results. A 99.66% of Sensitivity and a Positive Prediction equal to 99.56%.

In several algorithm on the ECG analysis the Wavelet Transform is used in combination with the threshold method. Another example is about the study conducted by [30], where the Daubehies Wavelet (dB6) is used and the scales D3 up to D5 are taken into consideration. A threshold is used to eliminate the lower waves,that are called pseudo waves, and the R waves are thus identified.

Other studies have instead used a combination of adaptive thresholds for R waves detection. [31] [32] from the research conducted by the Center of Biomedical Engineering in Bulgaria, three types of thresholds are established to detect the peaks even in case of high noise and to avoid losing the peaks with low amplitude. The results for both Sensitivity and Specificity is greater than 99.6%.

Taking as a reference the articles described above, an algorithm based on the Wavelet Transform and Threshold method has been developed. This methods have been enriched with additional steps to make the R wave detection algorithm more robust and efficient.

## 5.1 R wave detection

In order to identify the heartbeats and consequently calculate the median for each lead is necessary to detect R peaks. As already reported in the introduction, the

QRS complex is representative of the depolarization of the right and left ventricles. The R wave indicates the highest peak of the signal, a greater amplitude between 500 µV and 3 mV, which makes its location less complex.

The R waves, present in the signals given, have variable number, based on how many heartbeats are recorded within the 10 seconds of recording. Three separate functions have been realized to find all the R waves present in the signal. The different steps can be grouped into three large consecutive blocks:

-R wave detection with the threshold method

-R wave detection with the wavelet transform method

-Verification of the real presence of the R wave in the leads

### 5.1.1 Threshold Method

In the first phase a simple threshold is used to identify the highest peaks, which represent the R wave.

Processing one lead at a time, the value of the signal was compared to a certain point with the chosen threshold equal to:

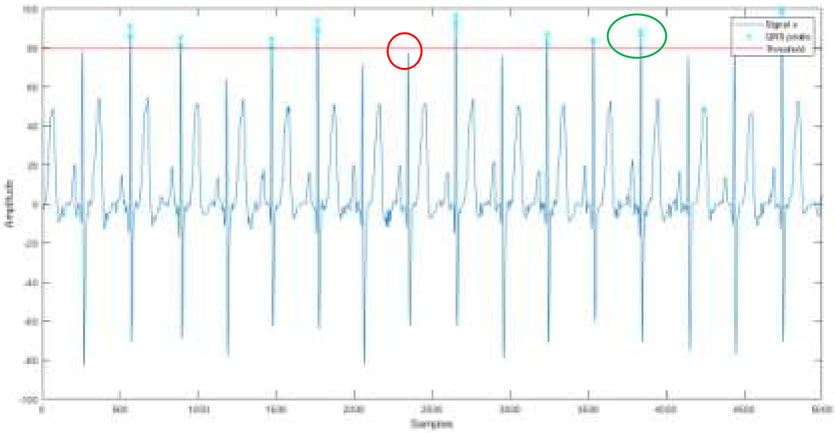
$Thres = 0.8 * \max(x)$ *if the |Maximum| is major than |Minimum|*

$Thres = 0.8 * min\,(x)$ *if the |Maximum| is minor than |Minimum|*

Where x is the ECG lead.

After setting the threshold, on a vector called QRS all the points of the signal x are saved that are greater (or lower in the second case) of the threshold.

For greater clarity the figure 5.1 show one of the two cases:



*5.1 Highlighted in red there is the lower R wave of the threshold which for this reason is not recognized as such, in green the case is highlighted in which more points are taken in the QRS complex which have a value greater than the threshold*

From the figure it is able to see how not all R waves are found.

Since it is not possible to capture all the R waves with a simple threshold, controls have been added to identify the waves correctly. as an example, in the two cases, two R waves were shown that, if they did not exceed the threshold value, were not recognized as such. Moreover, in green, those waves were highlighted in which more fiducial points of the QRS complex are taken. The goal is to finally obtain a vector with all R waves, therefore steps have been added with the following :

-$FirstControl = (Sample\ Rate * 0.060)\ samples$

-$SecondControl = 1.5 * median(RR\ interval)\ samples$

Where RR interval is the distance between the values include in the vector QRS, after the $FirstControl$.

Whether the distance between two consecutive points in the vector QRS is less than a constant value, the points belong to the same QRS complex. Bearing in mind that the QRS complex has a duration between 60 and 100 ms, if it is considered a value equal to 60 ms it is sure that the points in the range belong to the same QRS complex. The $FirstControl$ is used to eliminate all the points belonging to the same QRS complex, keeping only the maximum and saving it in a vector R wave.

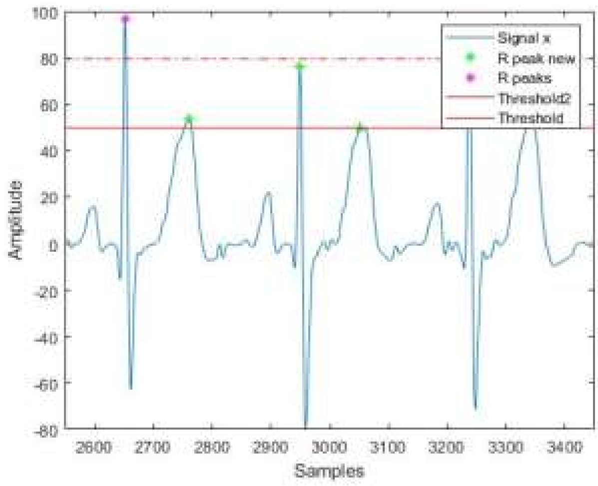
When the distance between two consecutive points, ie between two R waves is greater than an imposed value ($SecondControl$), probably a R wave has been lost. A new threshold is considered to catch the lost R wave:

$Thres2 = 0.5 * \max(x)$ *if the |Maximum| is major than |Minimum|*

$Thres2 = 0.5 * min(x)$ *if the |Maximum| is minor than |Minimum|*

At the end of this step a vector containing all the R waves identified by the threshold method is obtained.

This method presents some problems, as shows in the figure; some waves, that do not belong to the QRS complex, can be detected, due to the presence of a high T wave.




*5.2 Starting from the original signal one can perceive the error of recognizing as R wave higher T waves*

Starting from these problems, it is implemented a different method, more effective in the R wave detection. The capture of the R wave is, in fact, at the base of the P wave detection algorithm, so it is essential that it be performed accurately.

### 5.1.2 Wavelet Transform Method

The only threshold method is not sufficient to detect all R waves correctly and efficiently. A second step is introduced, which starting from the R waves found and from the signal reconstructed by the discrete wavelet transform method. The peaks are identified on the reconstructed signal. In this way the algorithm is more performing.
The second function, based on Wavelet Transform, will be describe in detail after. In the following paragraph a brief view of what is and how the Wavelet Transform works will be.

#### 5.1.2.1 Discrete Wavelet Transform

The Wavelet Transform is a time-scale representation, useful for the analysis and processing of signals. it is a powerful analysis tool when you need to work with non-stationary signals, which you want to get information about both the time domain and the frequency domain. the wavelet transform consists of a filter bank and has an excellent time-frequency localization capability. the operation consists in the decomposition of the signal on different scales related to frequency components, providing a multiresolution representation.

The Wavelet Transform can be continuous or discrete. in this case the Discrete Wavelet Transform (DWT) is employed. Specifically, the MODWT, the maximum overlap discrete wavelet transform, is used in this work. DWT and MODWT have the same starting formula.

The Discete Wavelet Transform of a signal x (t) is defined as follows:

$$W(a,b) = \sum_{i \in Z} f(i) * \psi_{n,k}(i) \quad (5.1)$$

.

where $\psi$ is the mother wavelet, **a** is a scale factor ,corresponds to frequency information, and **b** is the location parameter, attributed to the shift of the wavelet function through the signal.

The functions used in the transform are derived from the mother wavelet through

$$\psi_{n,k}(i) = \frac{1}{\sqrt{a^n}} \psi\left(\frac{i - kba^n}{a^n}\right) \quad (5.2)$$

a translation **b** and scaling **a**. $\psi$ is defined as follows:

It can be defined a group of Wavelet functions, Wavelet families, between which it

is possible to select a mother wavelet. The most common Wavelet functions are: Haar, Daubechies,Symlets and Coiflets. The DWT consists of a filter bank through which the decomposition of the signal takes place. There are two kinds of filtes:

- LPF: low pass filter
- HPF: high pass filter

-

The number of blocks of low and high pass filter is depended on the level chosen. The high pass filter provides information on the signal while the low pass filter return in output an approximation of the signal that is input to the filters of the next level. This result is called the first level of Wavelet transform.

What has been illustrated is also valid for MODWT, only the downsampling by the factor of 2 does not occurs. The MODWT is redundant because it is not orthonormal. It is not necessary that the samples are powers of two, the MODWT is in fact defined for all the samples.

The applications are the same, with some additional characteristics of the MODWT, for which it was preferred to the traditional DWT.

DWT and MODWT have different normalizations for filters,and there is no downsampling by 2n in the MODWT. The widths of MODWT e DWT filters are the same, but in the MODWT the filters are renormalized as follows:

$$\tilde{h}_{n,k} = h_{n,k} * \frac{1}{2^{\frac{j}{2}}}$$

$$\tilde{\mathrm{g}}_{\mathrm{n,k}} = \mathrm{g}_{\mathrm{n,k}} * \frac{1}{2^{\frac{\mathrm{j}}{2}}} \tag{5.3}$$

The main reason for choosing MODWT is that it solves the problem of alignment more precisely than the DWT, as showns following

As shown in the table, each level of decomposition of the Wavelet transform corresponds to a exact frequency range [33].

| Level | Frequency Range (Hz) |
|---|---|
| cD1 | 90-180 |
| cD2 | 45-90 |
| cD3 | 22.5-45 |
| cD4 | 11.25-22.5 |
| cD5 | 5.63-11.25 |
| cD6 | 2.81-5.63 |

| | |
|---|---|
| cD7 | 1.41-2.81 |
| cD8 | 0.7-1.41 |
| cA8 | 0-0.7 |

*Table 3 Correspondence between details and frequency range*

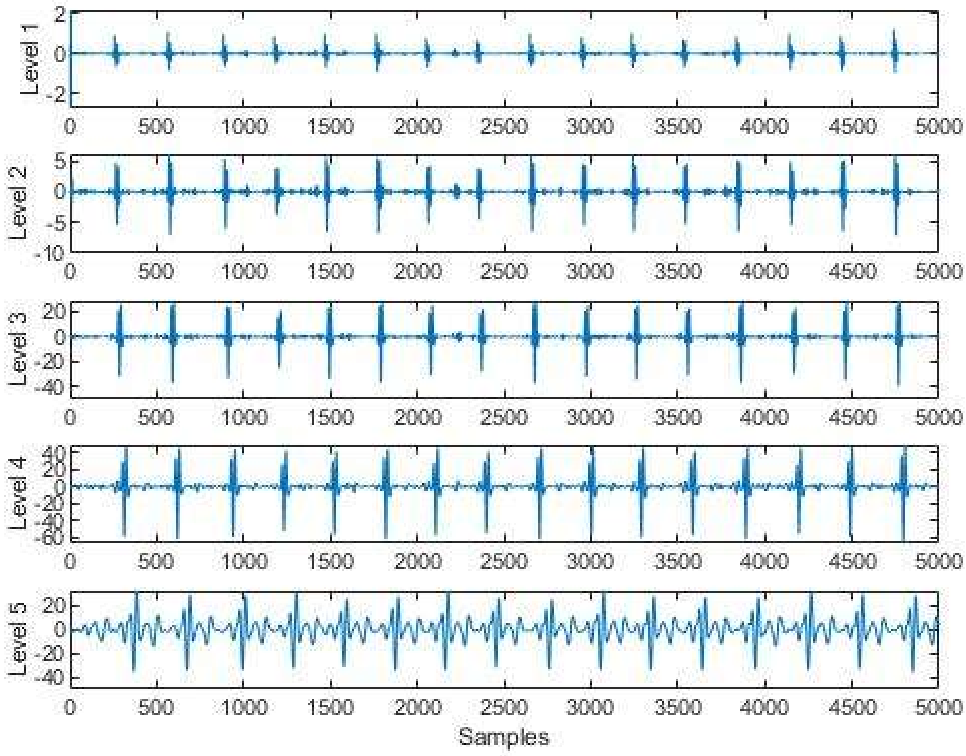
The DWT is used in this field in the ECG signal, so the range of interest corresponds to the level 4 and 5.

### 5.1.2.2 Maximum Overlap Discrete Wavelet Transform Method

The previous step that returns the vector containing the R waves is useful for deriving the average interval between a QRS complex and the other. It is thus possible to define a time interval in which the QRS complex is located, taken as a reference for the next step. A Dictionary for matching pursuit is defined, started from the QRS complex taken.

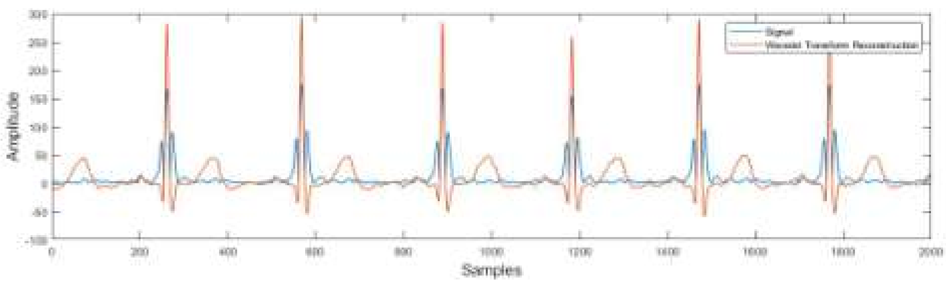
Subsequently the decomposition and reconstruction of the signal was carried out using MODWT. a 5-level decomposition was performed and levels 4 is taken for signal reconstruction. The reason of this choice is described in the previous chapter. The reason for this choice, as well as from the experimental observation, is justified by the several researches. They confirme that the most energy of the QRS complex is founded in scale 4.

The decomposition that occurs on the 5 levels is shown in the figure 5.3, it can seen how, in each decomposition, different frequency components are described.



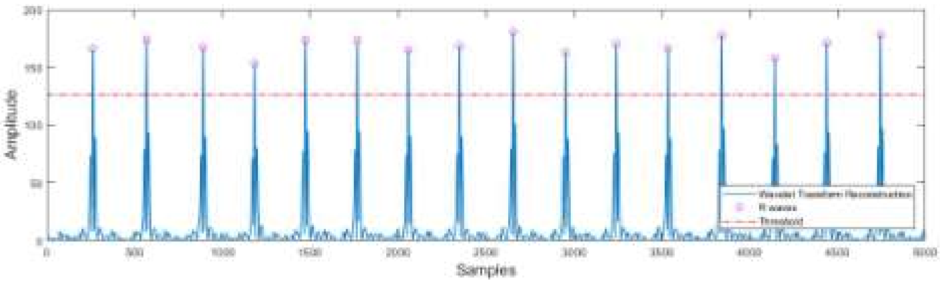

5.3 Representation of the WT decomposition of the ECG signal on 5 levels

A new signal y is reconstructed which mainly highlights the frequencies concerning the QRS complex. The signal y is compared with the original signal to verify that the original complex QRS is aligned with the signal y. The figure 5.4 shows the comparison:




5.4 Comparison between original signal and its reconstruction

It is possible to see how the reconstruction, starting after the decomposition on the 5 levels, returns a y signal that has peaks corresponding to the QRS complex of the original signal, while, elsewhere, there are only fluctuations around zero. This makes easier to identify the peaks of the R wave. Considering the module of y. as shown in figure 5.5 the problem referred to the figure 5.2 is solved. It is also possible, for the same reason, to set up a lower threshold.




*5.5Detection of R wave by WT*

The Thresold is set up as follows:

$$Thres = 0.7 * \max\left(|y|\right)$$

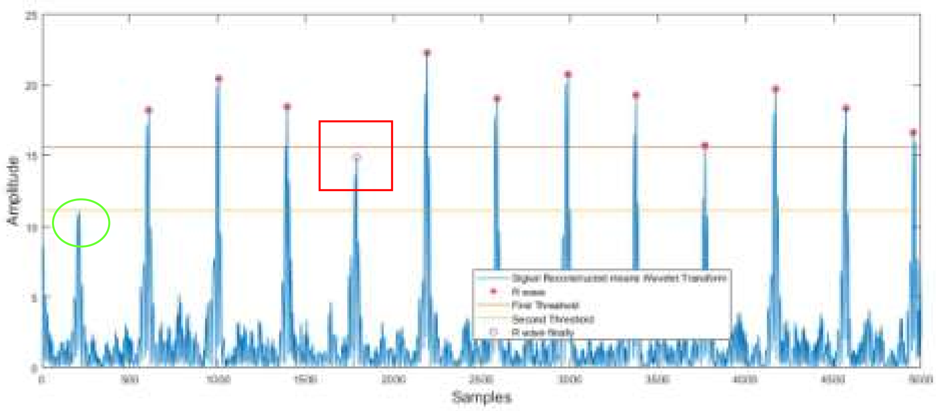
To identify waves A function present in matlab was used. the function makes it possible to identify all the points of $|y|$which have a height greater than the threshold and which are located at a certain distance.

In this way the $FirstControl$not necessary, as it was sufficient to give the value $(Sample\ Rate * 0.060)\ samples$ in input to the function present in Matlab. The output already includes all those points of $|y|$ greater than the threshold and which are at some distance from each other.

The SecondControl is, instead, maintained: if the distance between two consecutive R waves is greater than the SecondControl probably a R wave has been lost. . If the RR interval between two consecutive waves is greater than SecondControl, a new threshold is used:

$$Thres2 = 0.5 * \max\left(\ |y|\right)$$

Considering the interval between two R waves, a peak is catched if it is greater than the second threshold, it is a R wave that could not be recognized with the first threshold. The figure 5.2 shows what has just been said. The case in which the R wave is lower than both thresholds, and for this reason does not recognize, is highlighted in green. this problem is in most cases solved with ismember.. As happens in the example shown:




*5.6 The absolute value of WT is integrated in blue, the Thresholds 1 and 2 are in red and yellow, a R wave lower than the others is taken in the second control, while the R wave in green is never taken with this method*

At the end of the $SecondControl$, a new vector, containing all the R waves of the lead, is output.

## 5.2 Is Member

The last step consists, as described in [23], in taking into account that in the different leads of the same recording the QRS complexes can not differ more than 90 ms. This step is introducted for two main reason:

-solve the problem of non-recognition of R waves in some leads

-recognition of false R wave in a lead

So a check on all R waves founded in the leads is implemented.

The vector containing the R waves belonging to a lead is taken into consideration and is compared to the R waves of the other leads. To do this, a function already exist in Matlab tools is used. The function verifies that the values contained in a vector are also present in the other, unless a difference that is set by the operator. A tolerance of 90 ms was chosen.

The function returns two vectors, vector A and vector B, after the comparison between i-th lead and j-th lead. The vector A contains in i position the positions of the R waves of j which are present in i.

The vector B contains only values equal to 1 or 0:

-1 : the wave R n-th of j is present in i

- 0: the wave R n-th of j is not present in i

The vector A, if the wave R n-th of j is not present in i, is set =0.

Once the vectors A and B are obtained for all comparisons, it is obteined a Number of Leads x Number of R waves array. If a wave is only present in another lead or is not present in any of the others (value in the vector A = 0 in all the comparison), it will be deleted. A Number of Leads x Number of R waves array containing all the R waves of the 12 leads is given in output.

Afterwards, it was possible to calculate, once all the R waves are found, the average distance RR [34]

$$RRmean = \frac{\sum_{n=1}^{N}(IndexR_{n+1} - IndexR_n)}{N} \tag{5.4}$$

.

Where N is the number of R waves and $IndexR_n$ is the n-th R wave position.

After this calculation the heart rate $HR$ can be evaluate in Hz as follows:

(5.5)

$$HR = \frac{60}{RRmean} * Fc$$

.

Where $\boldsymbol{Fc}$ is the Sample Rate, equal to 500 Hz.

The evaluation of the Heart Rate is useful to recognise the tachycardia cases. In the tachycardia patient, the P wave is presented as an incision on the T wave[sit 6], this kind of analysis is not interesting in this study. In the tachycardia due to atrial fibrillation, there are a lot of fluctuation, that it does not permit the recognize of P wave

# Chapter 6
# Medians Calculation

The final scope of this work is to automatically detect P wave on the median of the beats. The reasons to consider the median and not the single beat are:

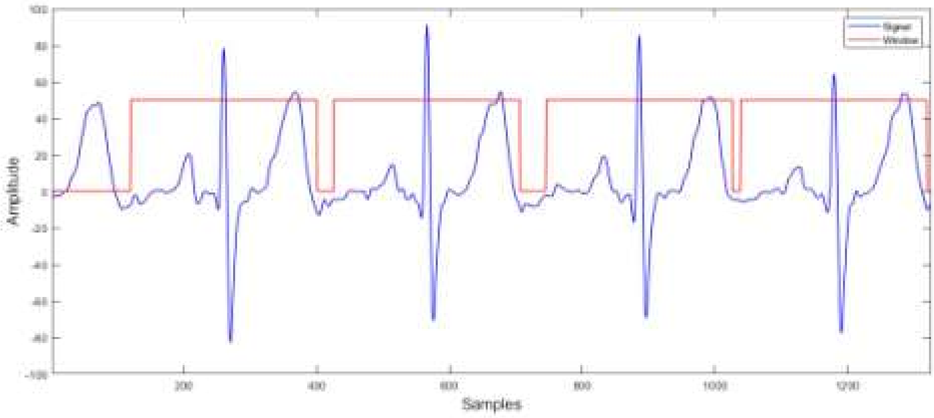
-Reduce the noise on the signal, making easier to detect the P wave

-Mediate the P waves allows to take into account the main P morphology and to eliminate the outliers due to artefacts or residual noise.

After founded all the R waves present in the recording it was sufficient to consider a rectangular window, that included all the heartbeat(PQRST), centered in the R wave. The constructed window is used to separate the beats and to calculate the median of them.
In detail this part of the algorithm has been realized with the following main steps:
6.1 creation of a rectangular window, based on the heart rate, ie based on the R-R interval, centered in R wave position.
6.2 Cross correlation calculation between the separated beats to obtain a better alignment between the beats
6.3 Median calculation of the realigned beats
6.4 interpolation to obtain all the medians of equal length

## 6.1 Window Construction

The function *Window Construction* uses the average distance between the R waves to construct a signal composed of 0 and 1. The window has a value equal to 1 at the beat, while in the isoelectric line it assumes null value. The figure 6.1 shows the window construction, highlighted in red.




*6.1 The detection of beats used to build the window to divide the signal into individual beats*

In this phase the average heart rate is also calculated starting from the average distance between the R waves. The following formulas are used:

$$Average\ Heart\ Rate = 60 * \frac{F_c}{RR\ mean} \tag{6.1}$$

Where $F_c$ is the Sample Rate and $RR\ mean$ is calculated as in formula (5.4).
A zero vector of the length of the signal, called Window, is initialized.

From time to time, for the window construction, a i-th R wave position is taken into account. The distance between the i-th R wave and its consecutive is examined:

-If the distance between the R wave and the i-th+1 R wave is greater than $RR\ mean$ a window of value 1 for a length equal to $RR\ average$ in correspondence with the i-th R wave is constructed;

-If the distance between the i-th R wave and the i-th+1 R wave is less than $RR\ mean$ , a window of value 1 for a length equal to the difference between the two R waves less a couple of samples is considered.

## 6.2 Cross Correlation beat to beat

The window is used to separate the beats of the signal. The signal is divided into epochs(for each beat), considering that when there is an ascent front of the window signal from 0 to 1 there is the beginning of a beat and when the window takes again null value, from 1 to 0, there is the end of the beat.

A matrix with size equal to Number of beats x beat's length is output at this operation. The matrix contains all the beats that have been identified starting from the initial signal.

Therefore it is proceed with the calculation of the cross correlation on double pass:

- first pass: a cross correlation between a beat and all the others is performed. for each comparison a counter is incremented if the delay is less than a certain threshold. The operation is performed for each beat. You will get a vector containing N counters, where N is the number of beats. The maximum value of the vector will correspond to that beat which is probably more representative of all the signals.

-second pass: the cross correlation is performed by the reference beat obtained in first pass with all the others. The beats that in the comparison have a delay greater than a set threshold will not be considered for the calculation of the median. The remaining beats will be realigned as follows:

```
If delay <0
beat(h,:)=[beat(h,abs(delay):end of beat),beat(h,1:abs(delay)-1)];
else delay>0
beat(h,:)=[beat(h,end of beat-delay:end of beat),beat(h,1:end of beat-
delay-1)];
```

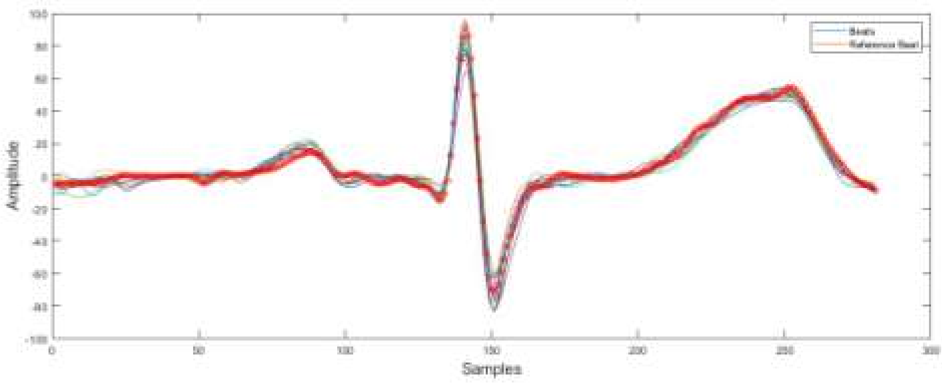
where *beat(i,:)* is the i-th beat derived to the initial signal.

For understanding better the meaning of the *delay* a brief description about the cross correlation is useful.

The cross correlation is a standard method of estimating how much two series are correlated. The cross correlation *C* at delay d is defined as:

$$C(d) = \frac{\sum_{i=1}^{N}[(x(i) - mx) * (y(i-d) - my)]}{\sqrt{\sum_{i=1}^{N}(x(i) - mx)^2}\sqrt{\sum_{i=1}^{N}(y(i-d) - my)^2}} \qquad (6.2)$$

For all delays d=0,1,2,...N-1 the result in a cross correlation is twice the length as the original series. x(i) and y(i) ,in this case, are the i-th beat and j-th beat. I and j can be assumed value = 1,2,…N, where N is the total number of beats. mx and my are the means of the corresponding beats.




*6.2 All beats are realigned with respect to the reference beat before the median calculation*

## 6.3 Median calculation of the realigned beats

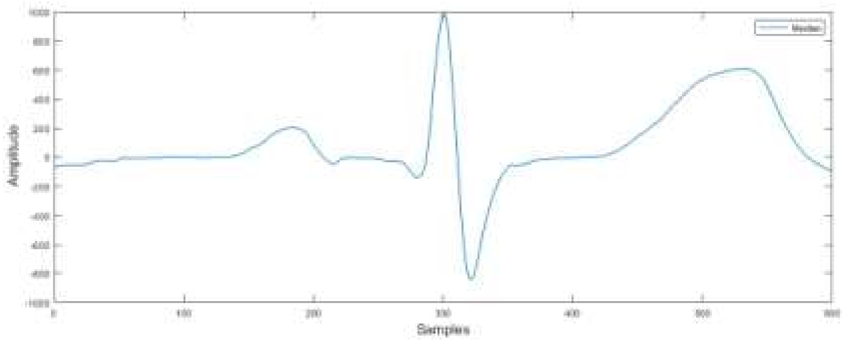
The median is a descriptive parameter, it represents the value that divides the distribution area into two halves..The median is generally used for skewed distributions. It was preferred to the average because it is less sensitive to outliers. The concept can be understood better with an example:

Considering a distribution of numbers *x = [12,15,20,23,24,27,200]* in which the value 200 is an anomalous value. The mean and the median are, respectively, 45.5

and 23. This simple example can be reported in this field of application. In the signal, in which there are on average a twelfth of beats, there are some beats that due to, for example, an artifact, are not representative of the signal pattern. Using the median these beats, if not previously eliminated with the cross-correlation method, will not affect the result. The most representative beats will have more weight in the final result, since the pattern is more repeated in the 10 s of acquisition.

The median is calculated among the realigned beats. After realignment a matrix A, of size *Number of beats* x *Length of beat,* is obtained. The median is performed along the first dimension, ie between the PQRST cycles that are found in the signal:

- Choice dim = 1, the median(A,1) returns a row vector containing the median of the elements in each column.
- An example of a median is shown in the figure:




*6.3Median of Signal*

## 6.4 Interpolation

The interpolation consists in finding a function able to represent the input signal on a variable number of points, without modifying the morphology of the signal. In Matlab, in order to perform interpolation, it is necessary to define the resampling frequency and the number of points on which the signal is to be represented.

The function has been constructed in such a way as to allow the operator to set the two parameters:

- Resampling frequency

-Number of points on which to represent the median

The operation is used for obtaining the output medians on the same number of points,without the morphology is changed.

# Chapter 7

# P wave Detection

The main part of the work proposed in the paper regards the detection of the P wave. To be able to identify the P wave correctly, the physiogical ECG pattern was taken into account. As previously said, the cardiac cycle of a healthy subject is constituted by the PQRST sequence. It is must taken into account, in the construction of the function to identify the P wave window, the following key points:

-the QRS complex is generally preceded by the P wave

-the PR interval (time between the beginning of the P wave and the beginning of the QRS complex) has a duration between 120 and 200 ms

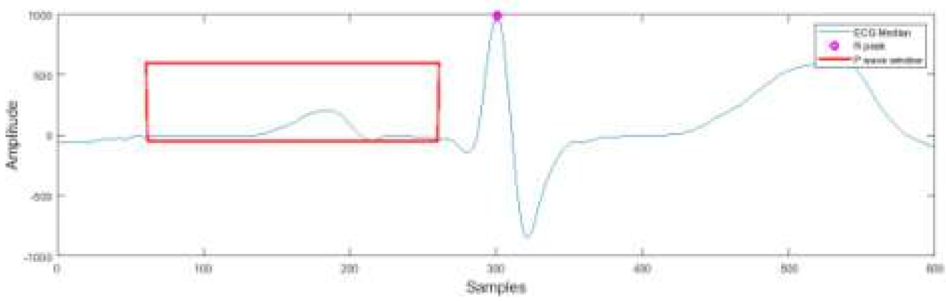
-the R wave is simple to identify, therefore it is a good reference point for identifying the P wave associated.

In the study conducted by A.Vitali [35]for the estimation of morphological features in the ECG, the most interesting waves is identified and analyzed. To find P wave a window is used. Taking as reference point the R wave peaks a window with a length equal to 200 ms, located at a distance of 50 ms from the reference point is catched. Subsequently the P wave will be the highest peak within the interval. The same method is also used by S.Sarkar [36]in a work about the construction of an apparatus for the recognition of atrial arrhythmia. To individuate the P wave he considers a window allocated with distance not specified respect the R wave peak.

## 7.1 P wave Window

This idea has been reused in this field. Starting from the medians obtained for each lead, a window of 200 ms is obtained in which the probability of detecting the P wave is high. The localization of the window has been realized taking as reference the maximum peak of the median, which corresponds, as previously said, to the R wave. Moving to the left of 45ms we are outside the QRS complex. Considering that the QRS complex has a duration between 60 and 90 ms and the peak of the R wave is about in the center of the complex, for at least 50 ms before the maximum it stills in the QRS complex.

Figure 7.1 shows the window constructed for the detection of the wave P, starting from the peak of the R wave :




*7.1 A window, with a fixed distance from R peak, is considered to identify the P wave*

The onset is define as the point between the TP isoelectric line and the start of the P deflection, the offset is the junction between the end of the P deflection before the beginning of PR segment.

In order to find the onset and the offset of the P wave, it was decided to proceed on two parallel branches and subsequently join the results:

- Zero Crossing Function
- Threshold Method based on the Noise deviation standard

### 7.1.1 Zero Crossing Function

Initially all the peaks are searched in the window, positive and negative. The set threshold is equal to half the maximum in the search for positive peaks and equal to half the minimum in the search for negative peaks.

Once the negative and positive peaks were identified, it was seen if the highest positive peak, in absolute value, was higher or lower than the highest negative peak;

it is assigned a flag as follows:

```
If | maximum peak of window|> |minimum peak of window|
Flag=1; %positive wave
else maximum peak of window|< |minimum peak of window|
Flag=0; %negative wave
```

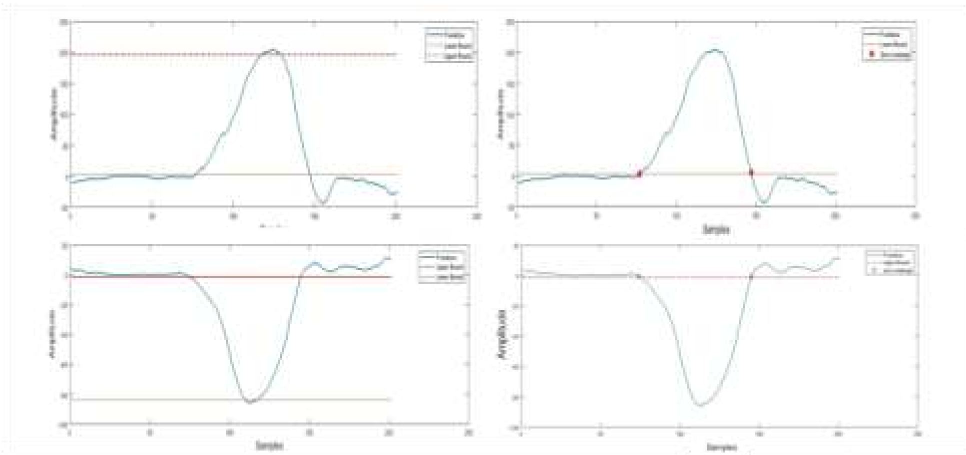
The attribution of the flag is necessary for the subsequent search of the zero crossings.

It was necessary to consider the lower and upper bound and not zero as a reference. Although the consistent preprocessing previously done, the zero can not be considered in all cases as a reference isoelectric line.

For this reason, it was decided to define in each case the reference line with respect to which the curve crossings should be found.

According to the flag, and therefore to the distribution of the wave, one of the two levels will be considered to determine the zero crossing.

As shown in figure 7.2 it is possible to calculate, considering the distribution of the wave P, a lower bound and an upper bound.



*7.2 Lower Bound, Upper Bound and zero crossings are represented both in the case of a positive wave (flag = 0) and in the case of a negative wave (flag = 1)*

In this case a flag = 1 is attributed to the wave, the onset and the offset of the wave will be considered the points of intersection between the curve and the Lower Bound.

For a better understanding, when the wave is negative, flag = 0 has also been reported. As can be seen from the lower part of the figure 7.2 the Upper Bound must be used to calculate the onset and offset of the P wave.

The points of crossover are found to evaluate the points of the vector x around the isoelectric line (`midRef`). The vector x contains the points of the signal located outside a transition region. If the value before and after i-th point are outside the transition region, then i-th x is the transition point:

```
If the transition is upward%look for the first positive crossing of
midRef
iX =find(x(1:i-1)<= midRef & midRef < x(i+1:end));
else %look for the first negative crossing of midRef
iX = find(x(1:i-1)>= midRef & midRef > x(i+1:end));
end
%%__________________________

% the value of the midRef dipend to the flag
If flag=1 % the wave is positive
midRef is equal to the LowerBound
else flag=0%the wave is negative
midRef is equal to the UpperBOund
end
```

where MidRef is a reference line that can be assume, how the pseudo code shows, different values.

For conciseness it is talked about zero crossing, this implies, as previously specified, that the crossing is evaluated respect to the Lower or Upper Bound.

In this step two controls are introduced. The first control is used to eliminate the crossings line which are due to fluctuations around zero and do not belong to the P wave. To do this a threshold has been set

$$Threshold\_tight = 20ms$$

This threshold was chosen on the basis that the minimum duration of a P wave is 60 ms.

Considering a vector containing N crossings, for i=1,2...N:

```
If crossing(ith+1)-crossing(ith) < Threshold
crossing(ith)=crossing(ith+1)+ crossing(ith)/2; %between two close
%values the average of the two is kept and the two values eliminated
delete crossing(ith+1)
end
```

The second check, on the contrary, verifies if the distance between two consecutive crossings is more than another threshold:

$$Threshold\_wide = 20ms$$

A P wave duration greater than 160 ms is impossible, so if the difference between onset and offset is more than $Threshold\_wide$, a new zero crossings research is done. Going to increase the isoelectric line tolerance it is looking for two closer crossings of zero.

After these two controls the first path returns the vector containing the zero crossings.

### 7.1.2 Threshold Method based on the Noise deviation standard

The second method with which the crossing is evaluated is taken from the propose to [37]. In the study conducted by F: Censi et al. to derive the temporal duration of the P wave a thresholding method is used. The onset of the P wave is found by analyzing 20 consecutive points. If this points are highest than 3 times the standard deviation of the residual noise the first of them is taken as the onset. The Offset is obtained with the same criteria, but starting from the QRS complex and coming back.

This method has been revisited and readapted. P-wave Onset and Offset are automatically taken as the first point, among 5 consecutive points, higher than a set threshold . In the Onset research the first half of the signal is considered, while Offset is defined going backward from the QRS, considered the last half of the P window. The threshold is equal to the residual noise standard deviation. Residual noise was measured in the isoelectric segment between P-wave and T wave.

The residual noise, and consequently its standard deviation, have a different value from signal to signal. this makes the threshold more specific for the single case but has required the implementation of a new function, *TP segment*, for the identification of the isoelectric line between the P wave and the T wave. In the following subsection it will be described in detail.

---

**TP segment Function**

In the physiological path of the ECG the T wave of a beat is followed by the P wave of the next beat and between the two curves there is an isoelectric line. The aim is to be able to identify this segment for extrapolating the residual noise. The median is typically composed of a single beat, then from a succession of P wave , a QRS complex and a T wave. For this reason it is necessary to do some operations to obtain the TP segment. The peak of the R wave is taken as the reference point. Beginning from this point two windows are used to center P and T wave. The window to the left of the R wave serves for catching the P wave and the right window for the T wave.

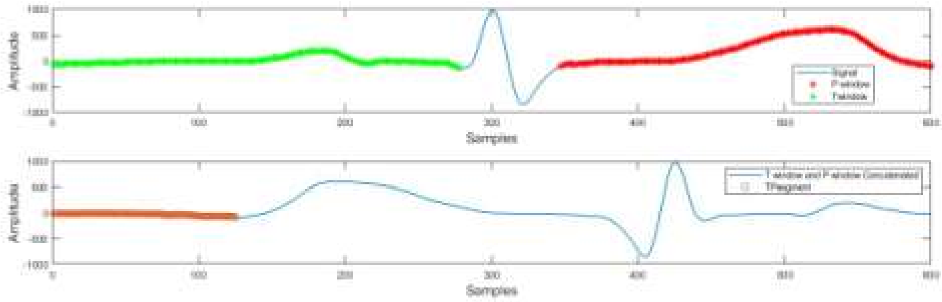
The two segments obtained are concatenated. In this way it is possible to have a T wave followed by the P wave as in the physiological path.

It is considered two points as references:

- maximum of the P wave

- maximum of the T wave

moving a certain number of samples with respect to the two maximums, the isoelectric stretch will be seen as shown in the figure:

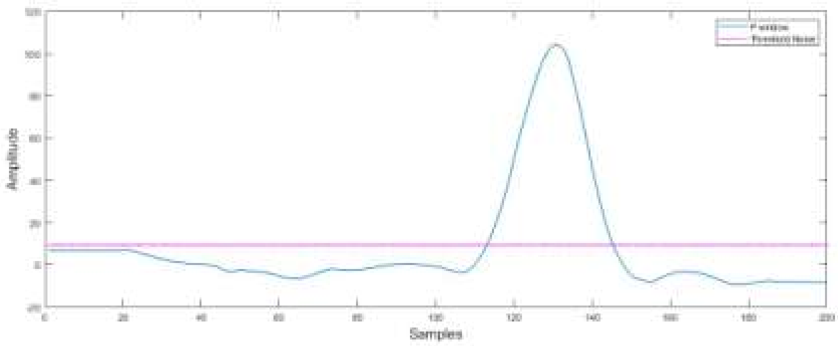



*7.3Above: the P wave is highlighted in red and the T wave in green; below: the identified TP segment is highlighted*

After identification of the TP segment, the standard deviation will be calculated, which will be used as the threshold, called Threshold_Noise, for the function to find the onset and the offset.

---

Once the threshold has been identified, the points of the P window are analyzed compared to it as the figure 7.4 shows:




*7.4 Identification of the onset and the offset of the P wave through the threshold*

In the Onset search it is started from the first point of the window and continue forward. In the Offset search, on the other hand, it is proceeded backwards, starting from the end of the window.

The following pseudo-code is useful to clarify the concept:

```
%Onset research
for i from 1 to the Maximum point of the wave
     if x(ith)>Threshold_Noise %comparison in Amplitude
        count_on=count_on+1;
     else
        count_on=0;
     end
    if 5 consecutive points greater than the threshold are counted
    ONset=i;
    end
end
%Offset research
for i from the end to the window to the Maximum point of the wave
     if x(ith)>Threshold_Noise %comparison in Amplitude
    count_off=count_off+1;
     else
     count_off=0;
     end
     if 5 consecutive points greater than the threshold were counted
     OFFset=i;
     end
end
```

if the Onset is not found, the value equal to the beginning of the window is assigned. When no point is recognized as offset, athe Offset is equal to the end of the window.

**Final Output: method for determining onset and offset at the end of the two parallel paths**

The second method used is inserted because the Onset and Offset search through zero crossing presented a problem. In the case of a biphasic P wave, the offset was often considered to be the second zero crossing and not the final point of wave. For this reason this second function was implemented and the results were subsequently joined.

The second technique has the problem of having a threshold that depends to another function. If the function of the TP segment identification makes errors in catching the isoelectric tract, the residual noise will subsequently be calculated incorrectly. The Threshold_Noise wrongly identified does not allow to identify the real Onset and Offset.

If at the end of the two functions it is obtained the reasonable values of onset and offset for the both functions, the first result is preferred, having shown greater stability in the observations.

Reasonable value is referred to a value that gives a P wave length between 60 and 160 ms. When the onset of one of the two functions has value 1 or the offset has value end of window, the output values of the other function are taken. If both results of the two functions return:

Onset = 1

Offset = end of window

Probably in the evaluate lead the P wave is not present or it is drowned up by the noise. Therefore, the mean values of the other leads are taken as offset and offset.

If in all leads the onset have value 1 and the value offset is equal to the end of window it is expected to be in the absence of the P wave.

**Ginput**

Since the algorithm is realized to be subsequently used in a graphical user interface (GUI) the ginput function has been introduced. The GUI will be created to give to the operator a rapid and usable interface , that shows the output of the algorithm and the results of the P wave classification. This function, present in the matlab toolbox, is used to allow the user to visualize the signal and a dialog box permits to choice if he wants to reset the onset and the offset . The median of a single lead and the onset and the offset of the P wave, highlight in red, are visualized. Manually the user can reselect the two points delimiting the P wave before proceeding with classification.

# Chapter 8

# P wave Classification

The focus of this work is to find parameters which enable the distinction between pathological and healthy patient.

The algorithm must be able to classify the P wave with respect to six classes required by the company.

the six real classes are shown in the following figure:

to each class has been attributed a value for language practice, we will talk about:

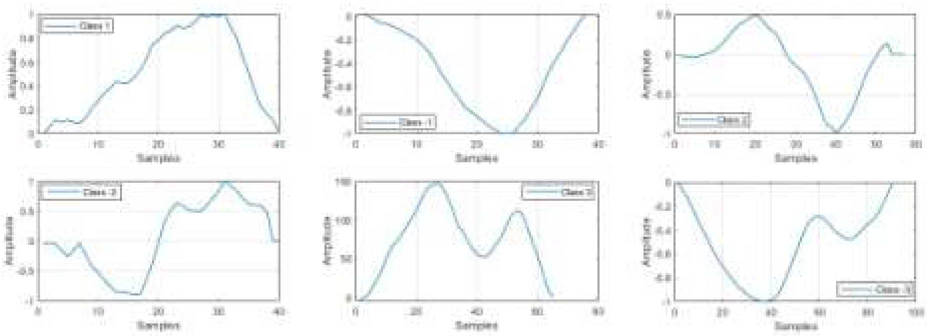
- class 1: positive linear wave
- class -1: negative linear wave

-class 2: isodiphasic wave with the first positive peak

-class -2: isodiphasic wave with the first negative peak

-class 3: positive bifid wave

-class -3: negative bifid wave



*8.1 Twenty of each type reference wave are selected, an example is reported of the six types*

The separation between class 1 and class 3 is the most important feature of this study. Class 3 is an indication of an atrial problem, it can be predictive of the onset of atrial fibrillation with related serious issues for the subject .

Bayés syndrome recently discovered as a correlation between the IAB, as mentioned before, so the appearance of atrial fibrillation and a specific morphology of the P wave are the most important key point of the study.

Scientific and technical studies conducted on the classification of ECG waves present in the literature were, analyzing 23 elderly subjects, has taken as a guideline. Subsequently, deepening and refined the methods, only those parameters that demonstrated an effective division capacity of the two classes are preserved.

## 8.1 Overview about Classification Methods for the P wave

In the study conducted by T.Hardahl et al. [38]the analysis of ECG curvature was performed considering the two statistical moments Skewness and Kurtosis and the temporal parameter. This type of study was conducted analyzing the T wave and using the combination of further information. In the case of the P wave it was observed by the same Bayes that it is not possible to distinguish the class 1 from the class 3 from the temporal duration because a partial atrial block has a duration greater than 120 ms, but a linear morphology. it is in the advanced atrial block, that the wave P, further to have a longer duration, presents the typical bifid morphology that it is wanted to identify. [39]

Another study, conducted by the same T. Hardahl, [40] on the T wave takes into account the radius of curvature of the wave and the local minima and maximums.

The main idea proposed is to analyze the derivatives of first and second order to obtain information about the concavity of the wave.

The approved method of derivatives is useful to know the curvature of the wave and to possibly identify non-physiological curvature due to noise.

In the research on the P wave parameters that allow to recognize the risk of atrial fibrillation by L.Clavier et al. mean and variance of the duration of the P wave and of the half of the area of the P wave from the starting point have been obtained. [41]

The idea stems from the fact that if the wave is elongated or bifid should return an A/T, where A is half the area and T is the total duration of the wave, greater in these cases than in the linear one. The negative results of this method further confirmed that a time domain analysis of the P wave for classification would be deviant and incorrect.

The method used by F: Censi et al. for the analysis of the P wave morphology it was very interesting for this work. [42] Starting from the following definition of Gaussian function:

$$y = \sum_{i=1}^{n} a_i * e^{-\left(\frac{(x-b_i)}{c_i}\right)^2} \tag{8.1}$$

Where the parameters a b and c depend on the initial P wave. It is tried to obtain, with the approximation error minimization method, a Gaussian model that approximates, as best it can, the P wave examined.

On the basis of the approximation obtained and the crossing of the zeros it is possible to classify the P wave. This method is a good starting point from which to take a cue. The main problem of the study conducted is that it does not take into account the morphology of class 3. In the case of bifid wave the approximation with the Gaussian becomes problematic.

## 8.2 P wave classification algorithm

The function from the classification in the six classes is performed in two main steps:

**1)**Comparison between the P wave examined and a set of reference P waves by MSE

**2)**Use some morphological parameters for the final classification

All the waves are normalized, so they have values between -1 and 1.

**1:** Twenty P waves are selected manually for each class, so the reference set consists of 120 waves. Each input P wave is compared with the 120 reference waves.

The mean square error is calculated as follows:

$$MSE = \sum_{i=1}^{N} \frac{x(i) - Wave\_type_j(i)}{N} \tag{8.2}$$

Where x(i) is the i-th point of the P wave evaluated. i can take values between 1 to N, N is equal to the length of the input signal.

$Wave\_type_j$ is the j-th reference wave. The first comparison takes place with the first reference wave for each type. This means that at the end of the first comparison there will be six MSE values. The comparison is repeated for the other 19 remaining waves for each type. At the end of this it is gotten a minimum value of MSE that allows to assign a class to the wave examined. This value is used for the next step.

**2:** In the second step there is a subdivision into three subgroups, the six classes are in fact merged into:

- 1 | 3;
- 1 |-3;
- 2 |-2.

This three groups have some similar characteristics, therefore the MSE can have close values for the two classes and could return the wrong class as output. This error can be partially corrected by the morphological analysis of the wave.

Class 1 and class 3 should have the following characteristics:

- the highest peak is equal to 1;

- there are only positive peaks;

- the vector containing the positive peaks in class 1 is of unit length;

- the vector containing the positive peaks in class 3 is of a length is equal to 2.

As regards classes -1 and -3 these present:

- the highest negative peak of -1;

- there are only negative peaks;

- the vector containing the negative peaks in class -1 is of unit length;

- the vector containing the negative peaks in class -3 is or equal to 2.

In classes 2 and -2 the characteristics are:

- have negative and positive peaks;

- the vector containing all the zero crossings have a length equal to 3 (onset + offset + zero crossing);

- the positive peak anticipates the negative peak in class 2;

- the negative peak is before the positive peak in the -2 class.

The conditions must be simultaneously verified in order to assign the class of belonging. If no class is assigned in this step, the class that returns the second minimum value of MSE will be assigned. The annexed pseudo code explains the algorithm about the morphological analysis:

```
if Class_step1 is equal to 1 or 3
  if  ~isempty(pks_pos) && pks_pos(1)=1 && isempty(pks_neg) && number
of peaks=1
    Class=1;
  elseif ~isempty(pks_pos) && pks_pos(1)=1 && isempty(pks_neg) &&
number of peaks=2
    Class=3;
  else
```

```
    delete the MSE minimum found previously and drop the new MSE
    minimum
    allocate the corresponding class to the value found
  end
```

If in the example the wave has only one positive peak, it will be assigned in class 1. If the different parameters are not recognized, it will be assigned, as follows: if the second minimum value of the MSE concerns the comparison with the wave type 2, the class 2 .

## 8.3 Analysis of additional parameters

The main problem with this method is that it does not take into account the presence of noise. In fact, despite the pre-processing, the wave P is a wave with a low amplitude that is often covered by noise. This condition could create for the class1 some oscillations around the maximum peak that cause the incorrect classification. In order to achieve the main objective of this work, that is to distinguish class 1 from class 3, additional parameters have been analysis.

The first dataset, including the 23 patients, was considered in the identification of the relevant parameters of the relevant parameters to split healthy and patients at risk, the 23 patients of the. This choice was made because the initial dataset contains the main characteristics of interest:

-Patients older than 65 years;

-Patients with abnormal P wave.

Taking a cue from the literature reported in the previous paragraph, it is analyzed different parameters, parameters in the frequency domain and additional parameters. To improve the algorithm's performance, in distinguishing class 1 from class 3, the following evaluation are made:

- Observe different parameters in frequency;
- Curve concavity evaluation;

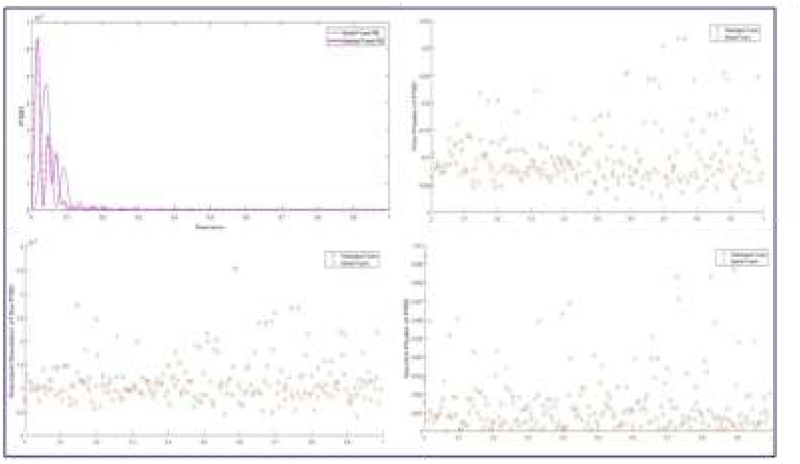
- Residual calculation after the Gaussian approximation method.

The only parameter turned out interesting is the Concavity and therefore the calculation of the first and second derivative of the wave taken into consideration.

## A. Analysis in the Frequency Domain

Since the classification of the P wave is difficult in time domain, it is tried to pass in frequency and to evaluate if there are any differences between the class 3 and the class 1. Taking a cue from what was found in literature it is made a analysis in the frequency domain.

In the first results, the PSD for class 3 shows more peaks and shifts compared to class 1. From this first observation the positions of the first two peaks for both classes is obtained. Furthermore the standard deviation is calculated. The results obtained shows that it is not possible to use these parameters to differentiate the two classes. As it can see in figure 8.1 there is no possibility to select a threshold that allows to create a division of the two classes.



*8.2 Comparison between Normal P wave PSD and Abnormal P wave PSD*

## B. Gaussian Approximation Method with LM

It is decided to use, through the Levenberg–Marquardt algorithm, a Gaussian approximation regarding the P wave .

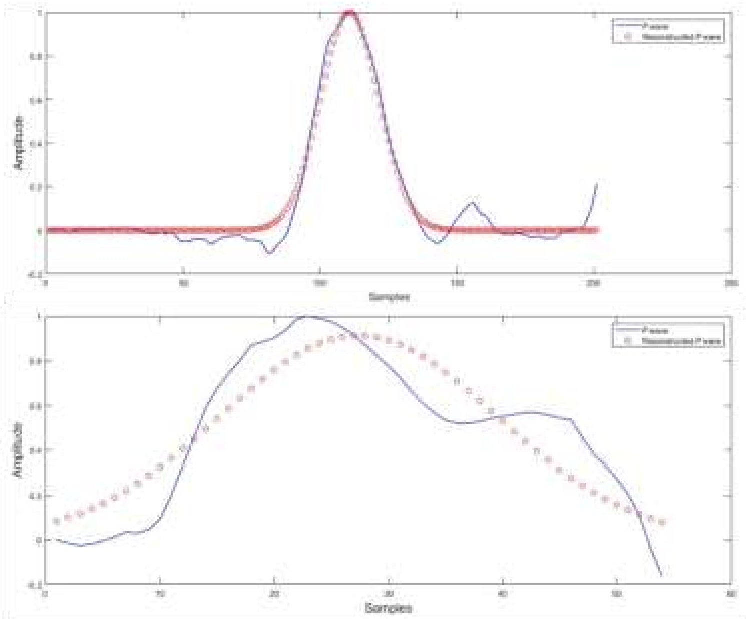
The idea arises from that the linear P waves can be assimilated with a good approximation to a Gaussian as the figure 8.2 shows, while the P waves of class 3 cannot be approximated by a single Gaussian. The residual value, that is the value obtained from the difference between the Gaussian model that the LM algorithm returns at the end of the 10 iterations imposed with the original wave, is used as first approach in this analysis. To make the approximation, formula 8.1 was used and the residual was calculated as follows:

$$Residual = \frac{sum((F(x, xdata) - ydata)^2)}{N_y} \tag{8.3}$$

Residual represents the quantity of the validation data not included in the model.

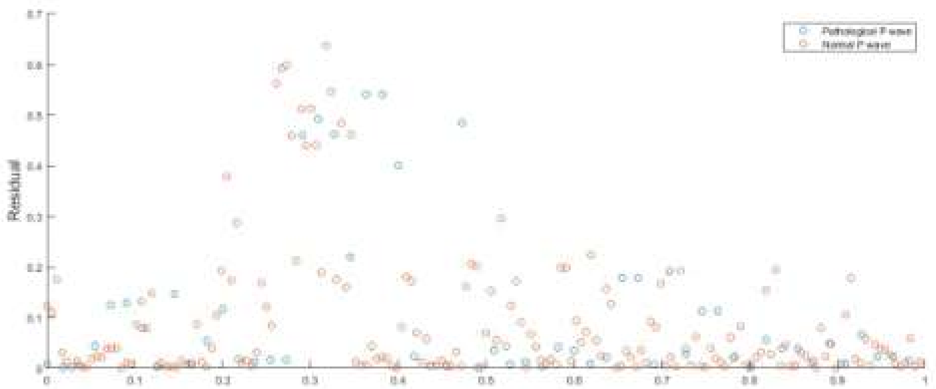
F(x,xdata) is the Gaussian function that approximates the input signal, ydata is the input and $N_y$ is the length of the input signal.

In figure 8.2 the output of the Gaussian approximation with the LM method and the input signal highlighted in blue are represented, it is evident that in the case of linear P wave the approximation with a Gaussian is possible, as compared to the bifid P wave.




*8.3 The approximation with a Gaussian returns an excellent reconstruction in class 1, the class 3 is not approximate to a single Gaussian*

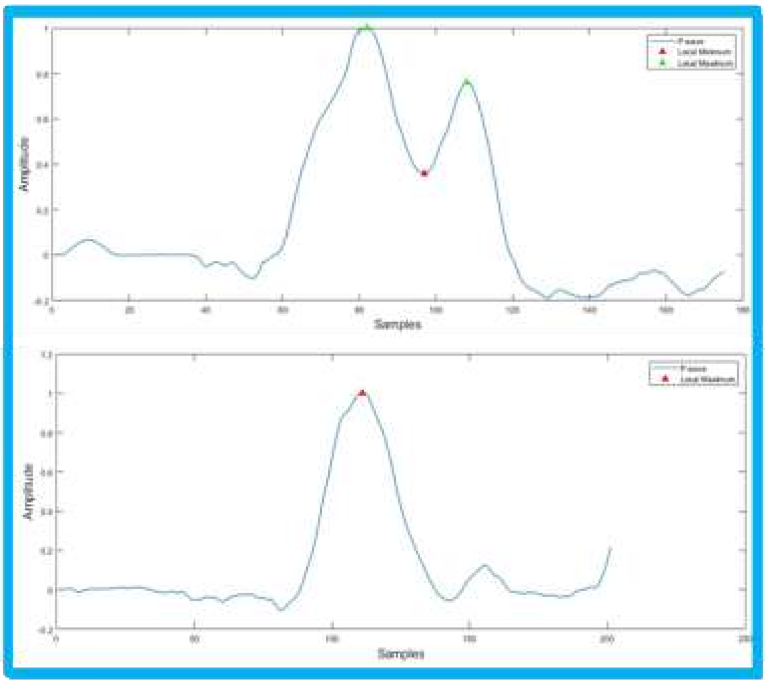
The 23 initial signals  show that it was impossible to make a subjection and to create a threshold starting from this result as shown in the figure 8.3. The residual does not represent a parameter that allows to obtain a correct classification.




*8.4  The residuals in the two classes, Red:class 1 Blue:class 3*

## C. Concavity study

The study of the concavity is used to analyze the local minima and maximums. The idea arises from the fact that in the theoretical case, class 1 has only one peak and therefore has a concavity facing downwards with a point of maximum local equal to the maximum height at the apex of the curve. In class 3 the curve will instead present a central deflection with two local maxima and a local minimum as shown in the figure 8.4:




*8.5 The figure above shows the local minimums and maximums in class 3, the figure below shows those of class 1. In class 1 there is only one local maximum that coincides with the global maximum.*

The second derivative can be positive, negative, or zero *[sit 7]*. The second derivative permits to classify critical point and to individuate local maxima and local minima. To recapitulate :

• if $\frac{dy}{dx}(\mathrm{p}) = 0$ and $\frac{d^2y}{dx^2}(p)$ is positive: p is a local minimum;

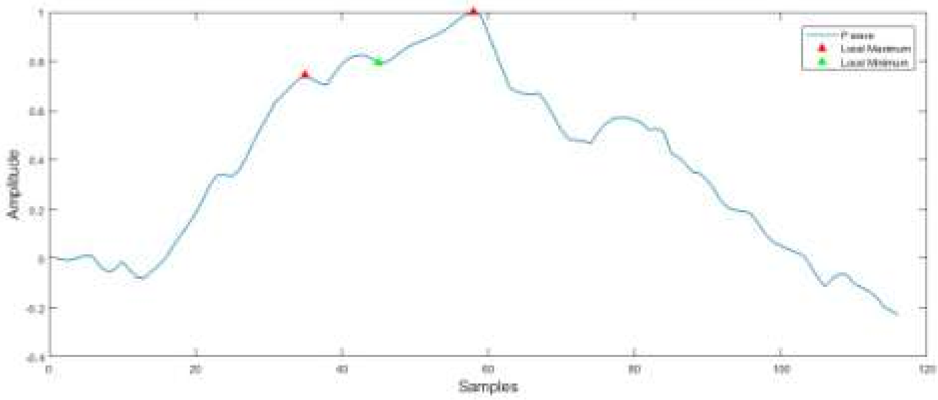
• if $\frac{dy}{dx}(\mathrm{p}) = 0$ and $\frac{d^2y}{dx^2}(p)$ is negative: p is a local maximum;

•if $\frac{dy}{dx}(\mathrm{p}) = 0$ and $\frac{d^2y}{dx^2}(p)$ is equal to 0: no new information are added about the behavior of y(x) at x = p.

The calculation of the local maximum and minimum points has been done as explained in the following pseudo code:

```
First_derivation=diff(x);
dy=[0 sign(First_derivation)];
locdn = find((diff(dy))==2);          % location down
locup = find((diff(dy))==-2);         % location up
Second_Derivation=[locdn locup];
```

This idea has been implemented and the first tests returns promising results. Merely, the problem of noise and oscillations that risk creating misclassifications remains unsolved.




*8.6 A misclassification of the P wave with a Concavity method, the local maximum are highlighted in red and the local minimum in green*

The figure 8.5 shows how the noise can cause the classification mistakes, despite the controls introduced. Two checks were introduced:

The first check takes into account the fact that the local minimums and maximums of interest are located in the highest area ,above a threshold, and so oscillations on the isoelectric due to noise must not be considered.

The second one concerns the proximity between a local maximum and a local minimum that cannot be too close. These two controls, which it is necessary to introduce, need a better setting to solve these misclassifications.

It is wanted to pursue this route going to set up controls and to search for non-traditional filtering methods. Arbitrary filters, used to remove the oscillations under a set threshold in order to eliminate progressively the noise and to leave only the curvature due to the morphology of the wave.

# Chapter 9

# Evaluation of the algorithm performance

In the evaluation of the algorithm's performances three main phases are highlighted, two intermediate ones beyond the conclusive phase, which it is considered necessary to evaluate:

- evaluation of the correct holding of the R wave
- correct identification of the window in which the P wave is situated
- correct classification of the P wave

The first evaluation is important because a correct recognition of the R wave is the core to the subsequent calculations on the heart rate, and so individuation of the single beats for the correct construction of the median.

The identification of the window where the wave P is positioned is a further critical point to be evaluated. Finally, the correct assignment of the class is assessed.

## 9.1 Evaluation of R wave detection

The first check is made on the R wave detection. In a classification process, the terms “true positive”, “true negative”, “false positive” and “false negative” are used to compare the classification predicted with the real classification. In this case the real negatives cannot be defined, it can define:

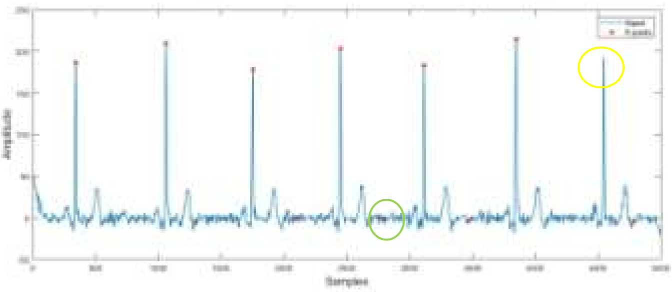
- true positive TP: R waves actually present in the signal;
- false positive FP: the algorithm recognizes points as R waves, but the R wave is not actually present;
- false negative FN: the algorithm does not recognize real R waves.

In the table 9.1 the terms are shown:

| TP | FP | FN |
|---|---|---|
| **153044** | 5272 | 828 |

*Table 4 True Positive, False Positive and False negative*

In the figure 9.1 there are highlighted in green FP and in yellow FN:



*9.1 R wave detection control*

To evaluate performance it is used the criteria of Precision:

$$Precision = \frac{TP}{TP + FP}$$

And Recall

$$Recall = \frac{TP}{TP + FN}$$

It is obtain a Precision=96.67% and a Recall =99.46%.

The error is mainly due to the following factors:

-presence of R waves in the first samples of some leads, which are not present in others;

-presence of noise at the beginning of the acquisition;

-presence of a single peak much higher than the others, probably due to a extrasystole.

The extrasystole consists of an arrhythmia that returns on the path a wave R much higher in amplitude than the rest of the signal. This can sometimes cause false negatives. False positives are mainly originated by too low thresholds that capture peaks belonging to other waves.

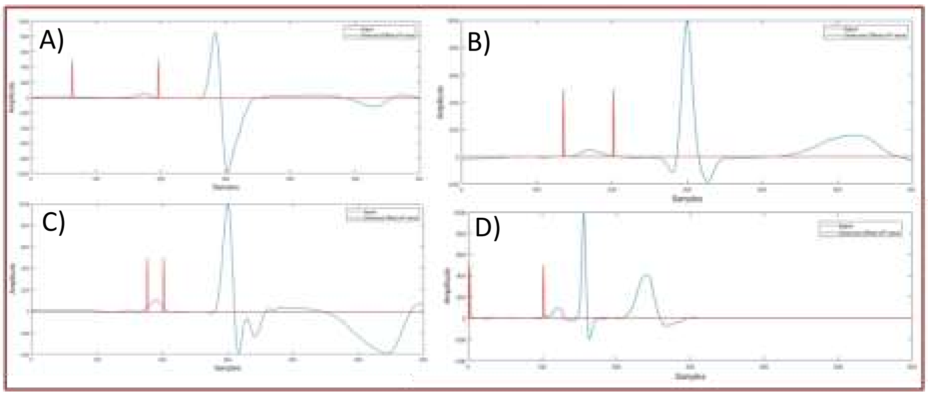
## 9.2 Evaluation of P window detection

The second evaluation is carried out on the search of the P wave window. Based on the observations made, four classes are identified (shown in figure 9.2), between which 1 and 2 are recognied as correct result and 0 and -1 are the wrong ones.

The 1000 signals in the Dataset were analyzed and the membership classes were assigned as follows:

- ❖ 1 = the window includes the wave P (A in figure 9.2);
- ❖ 2= the window is perfectly centered between Onset and Offset of the P wave (B in figure 9.2);
- ❖ -1 = the wave is taken but Onset and Offset less than real (D in figure 9.2);
- ❖ 0 = the window does not include the P wave (C in figure 9 .2).

The figure shows the four cases that are observed and are divided as described:




*9.2 Four types of output are found in the P window detection*

In the class 1 the wave P is evident but not correctly taken onset, offset or both.

Several oscillations around zero do not allow the correct detection of the P wave and consequently the Offset and Offset. In class1 we notice all the instances in which the P wave is not present in the lead.

In class 0 the window is not centered on the wave P, as the figure 9.2 D shows.

This problem can due to two main critical step:

-fixed distance between peak R and the start of P wave window;

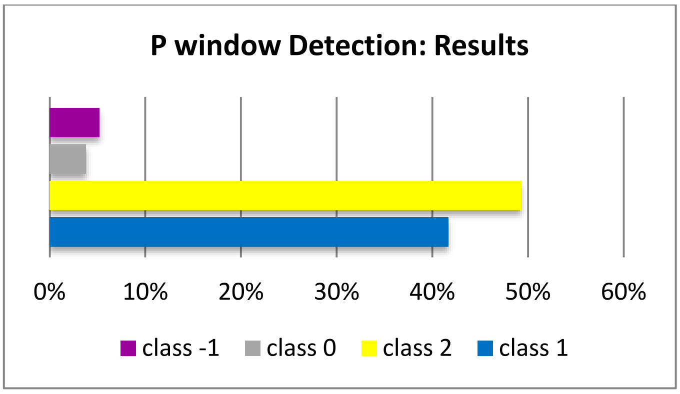
-the wrong R peak detection.

In class 0 the window can include other waves over the wave P, that not allow a correct classification.

The window, in class -1, partially includes the P wave, there is an underestimation of its duration. When the wave is isodiphasic, sometimes only the half wave that forms the P wave is identified (class -1). The main problem is in identifying the threshold value for the identification of the zero.

The P wave, in class 2, is detected correctly.

The results obtained from the algorithm performance evaluations show that in most cases (about 91%) the P wave window is correctly centered and in 9% of the cases part or all of the wave is lost




*9.3 P window detection results*

Unfortunately only 49.3% of the correct detection cases of P window onset and offset are located precisely. These results are mainly due to the oscillations present in the signal that are very large in amplitude. Furthermore, it would be necessary to correct the calculation of the TP segment to have a better threshold. The aim is to improve the threshold determination for the zero crossing calculation, as accurate as possible.

## 9.3 Evaluation of P wave classification

With reference to classification, two checks were made. In the first one was evaluated if the algorithm correctly returned the class of the single wave. The two consecutive steps adopted have proved to be necessary with respect to the only

MSE. In the first case there is a value of corrected classified equal to 74% while in the second case equal to 81% . Subsequently, it is evaluated whether the patient was recognized as pathological or not. For this purpose it is used a parameter that counted how many pathological waves are recurring in the same patient. In fact not in all the derivations it is possible to recognize the bifid wave. If, at least, in 3 leads (generally leads II, aVF and lead V1) the abnormal wave is identified, the patient is recognized as pathological.

In checking the validity of the algorithm, the first test on the 23 initial patients is very satisfactory. In this case the result was excellent because all the six pathological patients were recognized. In evaluating the same on the 1,000 patients, the results obtained are not the desired ones.

| Confusion Matrix | Real Pathological Patient | Real Pathological Patient |
|---|---|---|
| Pathological Patient Predicted | TP<br>65 | FP<br>26 |
| No Pathological Patient Predicted | FN<br>47 | TN<br>803 |

*Table 3 Confusion Matrix about the correct pathological patient recognized*

Through the confusion matrix it was possible to derive:

$$Sensitivity = \frac{TP}{TP + FN} = 60\%$$

$$Specificity = \frac{TN}{TN + FP} = 97\%$$

A very high specificity is obtained, but a too low sensitivity.

In 59 patients the wave P is not visible. In the remaining 941, it can be observed that out of 112 pathological patients, ie with Bifid P wave, 47 are not recognized as such. It was realized that the setting of the algorithm parameters on the part of the

P wave classification is polarized on the first 23 patients. It is therefore necessary to look for non-constant parameters which can be adapted to individual cases.

# Future developments and Conclusion

This thesis proposes the implementation of an algorithm to individuate and classify P wave in the ECG.

Research in this field is motivated by the aim to increase the prevention of cardiovascular diseases in the elderly, reduce diagnosis times and employee operator errors. For this purpose, an automatic algorithm is implemented. Subsequently its application has shown some limitations, already known in the literature, that do not make it excellent in performance and robustness.
The future work is to solve the problems related to the low amplitude of the P wave which is difficult to emerge from the residual noise.

The final aim of the work is the realization of a User Graphical Interface (GUI), in which the doctor is interfaced with the automatic analysis returned by the algorithm. Currently the results obtained are still to be improved. The main problems are given by the fixed thresholds introduced in the controls and steps of the algorithm.
The most critical constants are:
-Constant for determining the P wave window;
-Threshold for determining the crossing of zero;
-Threshold, in amplitude and in time, to determine the local minimums and maximums.
Furthermore, the problem with the noise is still a point on which working. It is wanted to develop a filter that is able to eliminate the oscillations due to noise and to keep only the convexities characteristic of the P wave morphology.

To solve the problem of threshold there will look for a better setting for the thresholds listed above and, in parallel, for implementing an adaptive thresholding method.

# Reference


[1] C. M. e. al., “ Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015,” *Journal of the American College of Cardiology,* 2017.

[2] G. F. Mureddu, “Evaluation of different strategies for identifying asymptomatic left ventricular dysfunction and pre-clinical (stage B) heart failure in the elderly. Results from 'PREDICTOR', a population based-study in central Italy,” *European Journal of Heart Failure,* vol. 15, no. 10, pp. 1102-1112, 2013.

[3] M. D. Felici, Embriologia umana , morfogenesi,processi molecolari,aspetti clinici, Piccin, 2014.

[4] M. J. Van Campenhout, “Bachmann's bundle a key player in the development of atrial fibrillation?,” *Circulation: Arrhythmia and Electrophysiology,* vol. 6, pp. 1041-1046, 2013.

[5] M. J. Van Campenhout, “Bachmann's bundle a key player in the development of atrial fibrillation?,” *Circulation: Arrhythmia and Electrophysiology,* vol. 6, pp. 1041-1046, 2013.

[6] G. Lanza, A. Pelargonio and G. Dello Russo, Aritmie, ELSEVIER S.R.L.

[7] M. P. Maurits A. Allessie, P. Penelope A. Boyden, M. A. John Camm, M. André G. Kléber, M. P. Max J. Lab, M. Marianne J. Legato, M. Michael R. Rosen, M. Peter J. Schwartz and P. M. Spooner, “Pathophysiology and Prevention of Atrial Fibrillation,” *American Heart Association,Inc,* February 6, 2001..

[8] P. Kirchhof, “Outcome parameters for trials in atrial fibrillation: Recommendations from a consensus conference organized by the German atrial fibrillation competence NETwork and the European heart rhythm association,” *Europace,* vol. 9, pp. 1006-1023, 2007.

[9] A. Bayés De Luna, “Interatrial blocks. A separate entity from left atrial enlargement: A consensus report,” *Journal of Electrocardiology,* vol. 45, pp. 445-451, 2012.

[10] Martínez-Sellés, “Rationale and design of the BAYES (Interatrial Block and Yearly Events) registry,” *Clinical Cardiology,* vol. 40, pp. 196-199, 2017.

[11] D. Conde, “Bayés' syndrome: The association between interatrial block and supraventricular arrhythmias,” *Expert Review of Cardiovascular Therapy,* vol. 13, pp. 541-550, 2015.

[12] P. M. Kistler, “Electrophysiologic and electroanatomic changes in the human atrium associated with age,” *Journal of the American College of Cardiology,* vol. 44, 2004.

[13] W. C. M. Y. e. a. Hayashi H, “Aging-related increase to inducible atrial fibrillation in the rat model.,” *J Cardiovasc Electro- physiol ,* vol. 13, pp. 801-8, 2002.

[14] R. R. MejÍa, Interpretazione dell’elettrocardiogramma, Milano: Springer, 2011.

[15] D. A. Rawshani, Clinical ECG interpretation, University of Gothenberg,Sweden.

[16] J. B. Nielsen, “P-wave duration and the risk of atrial fibrillation: Results from the Copenhagen ECG Study,” *Heart Rhythm,* vol. 12, pp. 1887-1895, 2015.

[17] V. A. MD, “Specific electrocardiographic markers of P-wave morphology in interatrial block,” *Journal of Electrocardiology,* vol. 39, pp. 380-384, 2006.

[18] Y. Zhou, “Denoising and Baseline Correction of ECG Signals using Sparse Representation,” *IEEE,* pp. 1-6, 2015.

[19] C. H. Lin, “Frequency-domain features for ECG beat discrimination using grey relational analysis-based classifier,” *Computers and Mathematics with Applications,* vol. 55, pp. 680-690, 2008.

[20] M. Blanco-Velasco, “ECG signal denoising and baseline wander correction based on the empirical mode decomposition,” *Computers in Biology and Medicine,* vol. 38, pp. 1-13, 2008.

[21] X. Ning, “Chromatogram baseline estimation and denoising using sparsity (BEADS),” *Chemometrics and Intelligent Laboratory Systems,* vol. 139, pp. 156-167, 2014.

[22] P. Hall, “Spectral Analysis of Signals,” *Spectral Element Method in Structural Dynamics,* pp. 11-38, 2009.

[23] S. Hargittai, “Savitzky-Golay least-squares polynomial filters in ECG signal processing,” *Computers in Cardiology,* vol. 32, pp. 763-766, 2005.

[24] A. Savitzky, “Smoothing and Differentiation of Data by Simplified Least Squares Procedures,” *Analytical Chemistry,* vol. 36, pp. 1627-1639, 1964.

[25] J. Chen, “A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter,” *Remote Sensing of Environment,* vol. 91, pp. 332-344, 2004.

[26] C. Li, “Detection of ECG characteristic points using wavelet transforms,” *EEE Trans Biomed Eng,* vol. 42, pp. 21-28, 1995.

[27] P. Sasikala, “Robust R Peak and QRS detection in Electrocardiogram using Wavelet Transform,” *International Journal of Advanced Computer Science and Applications,* vol. 1, 2010.

[28] H. Y. Lin, “Discrete-wavelet-transform-based noise reduction and R wave detection for ECG signals,” *2013 IEEE 15th International Conference on e-Health Networking, Applications and Services, Healthcom ,* pp. 355-360, 2013.

[29] J. P. Martínez, “A wavelet-based ECG delineator: evaluation on standard databases,” *IEEE Transactions on Biomedical Engineering,* pp. 570-581, 2004.

[30] J. Singaraju, “Automatic Detection of ECG R-R Interval using Discrete Wavelet Transformation,” *International Journal on Computer Science and Engineering (IJCSE),* pp. 1599-1605, 2011.

[31] I. I. Christov, “Real time electrocardiogram QRS detection using combined adaptive threshold,” *BioMedical Engineering Online,* pp. 1-9, 2009.

[32] G. Chen, “an Ecg R-Wave Detection Algorithm Based on,” pp. 145-149, 2015.

[33] H.-Y. Lin, “Discrete-wavelet-transform-base,” *IEEE,* 2014.

[34] J. D. Petty, “(12) Patent Application Publication (10) Pub . No .: US 2002/0187020 A1,” vol. 1, 2002.

[35] A. L. Vitali, “METHOD AND DEVICE FORESTIMATING MORPHOLOGICAL FEATURES OF HEART BEATS,” 2011.

[36] S. Shantanu, “apparatus for atrial arrhythmia episode detection,” *PCT,* 2016.

[37] F. Censi, “Effect of ECG Filtering on Time Domain Analysis of the P-Wave Department of Technologies and Health , Italian Institute of Health , Rome , Italy,” pp. 1077-1080, 2008.

[38] T. B. Hardahl, “ SYSTEMAND A METHOD FOR ANALYSING ECG CURVATURE FOR LONG QT SYNDROME AND DRUG INFLUENCE,” *United States Patent,* vol. 2, pp. 12-15, 2011.

[39] L. Fabritz, “The power of P in the elderly: Small biphasic wave, big impact,” *Heart Rhythm,* vol. 13, pp. 652-653, 2013.

[40] T. Hardahl, “System and Method for Analyzing Complex Curvature of Ecg Curves,” *US Patent App.,* 2010.

[41] J. M. B. J. J. B. Laurent CLAVIER*, “P-wave parameters for atrial fibrillation risk detection. Laurent,” *IEEE Engineering in Medicine and Biology Society,* 1996.

[42] G. C. P. B. E. C. I. D. I. C. G. B. F Censi1, 1. Censi F, Calcagnini G, Bartolini P, Cervi E. Effect of ECG Filtering on Time Domain Analysis of the P-Wave Department of Technologies and Health , Italian Institute of Health , Rome , Italy,2008.

# Sitography


[1] https://www.worldatlas.com/articles/countries-with-the-largest-aging-population-in-the-world.html

[2] http://www.humanitas.it/malattie/scompenso-cardiaco

[3] https://lifeinthefastlane.com/ecg-library/basics/p-wave/

[4] https://en.ecgpedia.org/index.php?title=P_Wave_Morphology

[5] https://www.materdomini.it/malattie/fibrillazione-atriale/

[6] http://corsoecg.zeronovetre.it/engine/test.interattivo

[7] https://math.dartmouth.edu/opencalc2/cole/lecture8.pdf