# Auditory filters at low frequencies: ERB and filter shape

Spring - 2007

Acoustics - 07gr1061 —

Carlos Jurado David Robledano

Spring 2007 Aalborg University

# Preface

The report contains all relevant information related to the Master Thesis carried out by Group 1062 at the Acoustics Section of the Department of Electronic Systems at Aalborg University, from February to June 2007. It is addressed to anyone interested in auditory filtering and low frequency sound perception.

The report is structured in two parts. The first is the main report and it contains an introduction to the topic, an explanation of the method, a description of the listening experiments, a presentation of the results and the final conclusions. All information considered not essential for the report, but related to the project work has been put apart into an Appendix. In the Appendix, related theory and further practical details about the project development can be found. A CD is also attached where the the programs used during the project work are presented.

The reader can also find a Bibliography with the list of all the monographes, papers and any kind of sources that have been used to support this work. Along the report they are referenced with a number between brackets, e.g. [4]. The number in the reference corresponds with the number that can be found in the list. Figures, tables and equations are numbered according to the chapters, e.g. figure 3 in chapter 2 is written as: Figure 2.3.

The authors would like to thank all the subjects that voluntarily have participated in the experiments as well as Christian S. Pedersen for his valuable help.

David Robledano

Carlos Jurado

Group ACO-1062, Aalborg, June 7th 2007.

# Faculty of Engineering and Science

Aalborg University

#### **Department of Acoustics**

TITLE: Auditory filters at low frequencies: ERB and filter shape

**THESIS PERIOD:** February - June 2007

**GROUP:** 07gr1062

GROUP MEMBERS: Carlos Jurado David Robledano

SUPERVISOR: Christian S. Pedersen

NUMBER OF PAGES IN REPORT: 84.

NUMBER OF PAGES IN AP-PENDIX: 41.

TOTAL NUMBER OF PAGES: 125.

#### Abstract

Exposure and effects of low frequency noises demand a better understanding of auditory filtering at low frequencies. Given that below 100 Hz, auditory filters have not been systematically studied, an effort was done in this project to better describe auditory filters in this less inspected frequency region.

Therefore, filter shapes and ERBs of the auditory filters in the range 31.5-125 Hz have been analyzed and explored by means of listening experiments. The center frequencies 31.5, 50, 80 and 125 Hz were tested. Experiments where performed under binaural exposure conditions in a room conditioned for controlled low frequency sound reproduction.

To obtain the auditory filter properties, the notched noise method was applied. The measurements consisted on obtaining a number of thresholds for tones at the selected center frequencies in the presence of a notched noise masker around the tone. With these data, the auditory filters were derived by fitting them to the *rounded exponential* model.

To account for outer and middle ear attenuation, individual equal loudness contours were measured and the notched noise stimuli were accordingly compensated.

Results indicate that the filter shape, under a noise level of 54 dB, is generally symmetrical in this frequency region. It was found that the ERB continues to decrease below 100 Hz, reaching an average bandwidth of 17 Hz at 31.5 Hz. Besides, the dynamic range of the auditory filter was found to be more limited at 31.5 Hz.

# Contents

1	Introduction 7										
	1.1	Initial problem									
1.2 Scope of the project			of the project	9							
	1.3	The cr	itical band and the auditory filters	10							
		1.3.1	The critical bands	10							
		1.3.2	The power spectrum model	11							
		1.3.3	Auditory filters: bandwidth and shape	12							
		1.3.4	Equivalent rectangular bandwidth	14							
		1.3.5	Off-frequency listening	15							
	1.4	Physio	logical origin to the auditory filters	15							
		1.4.1	The <i>place</i> theory	15							
		1.4.2	The temporal theory	16							
		1.4.3	Auditory filters, loudness and basilar membrane performance	16							
	1.5	The au	ditory filters at low frequencies	17							
2	Met	hod		20							
2	<b>Met</b> 2.1	hod Metho	ds for determining the critical bandwidth and the auditory filter shape $\ldots$ .	<b>20</b> 20							
2	<b>Met</b> 2.1	hod Metho 2.1.1	ds for determining the critical bandwidth and the auditory filter shape Band limiting methods	<b>20</b> 20 21							
2	<b>Met</b> 2.1	hod Metho 2.1.1 2.1.2	ds for determining the critical bandwidth and the auditory filter shape Band limiting methods	<b>20</b> 20 21 22							
2	<b>Met</b> 2.1	hod Metho 2.1.1 2.1.2 2.1.3	ds for determining the critical bandwidth and the auditory filter shape Band limiting methods	<ul> <li>20</li> <li>20</li> <li>21</li> <li>22</li> <li>23</li> </ul>							
2	<b>Met</b> 2.1	hod Metho 2.1.1 2.1.2 2.1.3 2.1.4	ds for determining the critical bandwidth and the auditory filter shape Band limiting methods	<ul> <li>20</li> <li>20</li> <li>21</li> <li>22</li> <li>23</li> <li>24</li> </ul>							
2	<b>Met</b> 2.1	hod Metho 2.1.1 2.1.2 2.1.3 2.1.4 2.1.5	ds for determining the critical bandwidth and the auditory filter shape Band limiting methods	<ul> <li>20</li> <li>20</li> <li>21</li> <li>22</li> <li>23</li> <li>24</li> <li>24</li> </ul>							
2	<b>Met</b> 2.1	hod Methor 2.1.1 2.1.2 2.1.3 2.1.4 2.1.5 2.1.6	ds for determining the critical bandwidth and the auditory filter shape Band limiting methods	<ul> <li>20</li> <li>20</li> <li>21</li> <li>22</li> <li>23</li> <li>24</li> <li>24</li> <li>25</li> </ul>							
2	<b>Met</b> 2.1	hod Methor 2.1.1 2.1.2 2.1.3 2.1.4 2.1.5 2.1.6	ds for determining the critical bandwidth and the auditory filter shape Band limiting methods	<ul> <li>20</li> <li>20</li> <li>21</li> <li>22</li> <li>23</li> <li>24</li> <li>24</li> <li>25</li> <li>26</li> </ul>							
2	<b>Met</b> 2.1 2.2	hod Methor 2.1.1 2.1.2 2.1.3 2.1.4 2.1.5 2.1.6 Audito	ds for determining the critical bandwidth and the auditory filter shape Band limiting methods	<ul> <li>20</li> <li>21</li> <li>22</li> <li>23</li> <li>24</li> <li>24</li> <li>25</li> <li>26</li> <li>27</li> </ul>							
2	<b>Met</b> 2.1 2.2	hod Methor 2.1.1 2.1.2 2.1.3 2.1.4 2.1.5 2.1.6 Auditor 2.2.1	ds for determining the critical bandwidth and the auditory filter shape Band limiting methods	<ul> <li>20</li> <li>21</li> <li>22</li> <li>23</li> <li>24</li> <li>24</li> <li>25</li> <li>26</li> <li>27</li> <li>27</li> </ul>							
2	<b>Met</b> 2.1 2.2	hod Methor 2.1.1 2.1.2 2.1.3 2.1.4 2.1.5 2.1.6 Auditor 2.2.1 2.2.2	ds for determining the critical bandwidth and the auditory filter shape Band limiting methods	<ul> <li>20</li> <li>21</li> <li>22</li> <li>23</li> <li>24</li> <li>24</li> <li>25</li> <li>26</li> <li>27</li> <li>27</li> <li>28</li> </ul>							

		2.2.3	Auditory fi	Iter parameter determination	31
			2.2.3.1	Fitting the threshold data to the $roex(p,r)$ model	32
			2.2.3.2	Accounting for the auditory filter assumptions in the fitting pro-	32
	2.3	Measu	rement cons	iderations	33
		2.3.1	Selection of	of sound reproduction system for controlled exposure conditions	34
			2.3.1.1 I	Evaluation of headphones	36
			2.3.1.2 I	Evaluation of low frequency room system	36
		2.3.2	Accounting	g for the non-flat overall response	37
			2.3.2.1	The frequency response of the room and the reproduction system	37
			2.3.2.2	The outer and middle ear sound attenuation	38
	2.4	The no	tched noise	stimuli	39
		2.4.1	Choosing o	center frequencies to investigate	39
		2.4.2	Determinir	ng the width of the notch bands	39
		2.4.3	Notched no	bise stimuli proposed configuration	41
	2.5	Derivit	ng the audito	bry filters	43
2.5.1 The fitting procedure and the notched noise config		procedure and the notched noise configurations	43		
			2.5.1.1	Shifting the auditory filter to the max(S/N) position and modi-	15
			1		43
		2.5.2	1 Auditory fi	ilter asymmetry and filter shifts	43 46
3	Liste	2.5.2 ening ex	1 Auditory fi speriments	ilter asymmetry and filter shifts	<ul><li>43</li><li>46</li><li>49</li></ul>
3	Liste 3.1	2.5.2 ening ex Measu	Auditory fi <b>periments</b> rement set-u	ilter asymmetry and filter shifts	<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> </ul>
3	<b>Liste</b> 3.1	2.5.2 ening ex Measu 3.1.1	Auditory fi <b>periments</b> rement set-u Sound envi		<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> <li>50</li> </ul>
3	<b>Liste</b> 3.1	2.5.2 ening ex Measu 3.1.1 3.1.2	Auditory fi <b>Experiments</b> rement set-u Sound envi Control and	Idea intillig procedure	<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> <li>50</li> <li>50</li> </ul>
3	<b>Lista</b> 3.1	2.5.2 ening ex Measu 3.1.1 3.1.2 3.1.3	Auditory fi <b>Experiments</b> rement set-u Sound envi Control and Stimuli and	Idea intuiting procedure	<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> <li>50</li> <li>50</li> <li>50</li> <li>51</li> </ul>
3	<b>Lista</b> 3.1	2.5.2 ening ex Measu 3.1.1 3.1.2 3.1.3 3.1.4	Auditory fi <b>Experiments</b> rement set-u Sound envi Control and Stimuli and Subjects .	Idea intuiting procedure	<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> <li>50</li> <li>50</li> <li>51</li> <li>51</li> </ul>
3	Lista 3.1 3.2	2.5.2 ening ex Measu 3.1.1 3.1.2 3.1.3 3.1.4 Phase	Auditory fi <b>Experiments</b> rement set-u Sound envi Control and Stimuli and Subjects . 1: Threshold	ined fitting procedure	<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> <li>50</li> <li>50</li> <li>51</li> <li>51</li> <li>51</li> </ul>
3	<b>Lista</b> 3.1 3.2	2.5.2 ening ex Measu 3.1.1 3.1.2 3.1.3 3.1.4 Phase 3.2.1	Auditory fr Auditory fr cperiments rement set-u Sound envi Control an Stimuli and Subjects . 1: Threshold	ined fitting procedure	<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> <li>50</li> <li>50</li> <li>51</li> <li>51</li> <li>51</li> <li>52</li> </ul>
3	Lista 3.1 3.2	2.5.2 ening ex Measu 3.1.1 3.1.2 3.1.3 3.1.4 Phase 3.2.1 3.2.2	Auditory fr Auditory fr cperiments rement set-u Sound envi Control an Stimuli and Subjects . 1: Threshold Threshold Equal loud	Idea intillig procedure	<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> <li>50</li> <li>50</li> <li>51</li> <li>51</li> <li>51</li> <li>52</li> <li>53</li> </ul>
3	Lista 3.1 3.2	2.5.2 ening ex Measu 3.1.1 3.1.2 3.1.3 3.1.4 Phase 3.2.1 3.2.2 3.2.3	Auditory fr Auditory fr cperiments rement set-u Sound envi Control an Stimuli and Subjects . 1: Threshold Threshold Equal loud Results .	Idea intillig procedure       Image: Im	<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> <li>50</li> <li>50</li> <li>51</li> <li>51</li> <li>51</li> <li>52</li> <li>53</li> <li>54</li> </ul>
3	Lista 3.1 3.2 3.3	2.5.2 ening ex Measu 3.1.1 3.1.2 3.1.3 3.1.4 Phase 3.2.1 3.2.2 3.2.3 Phase	Auditory fr Auditory fr cperiments rement set-u Sound envi Control an Stimuli and Subjects . 1: Threshold Threshold Equal loud Results . 2: Threshold	and intermination:       The ascending method         alter asymmetry and filter shifts       The a	<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> <li>50</li> <li>50</li> <li>51</li> <li>51</li> <li>51</li> <li>52</li> <li>53</li> <li>54</li> <li>55</li> </ul>
3	Lista 3.1 3.2 3.3	2.5.2 ening ex Measu 3.1.1 3.1.2 3.1.3 3.1.4 Phase 3.2.1 3.2.2 3.2.3 Phase 3.3.1	Auditory fr Auditory fr cperiments rement set-u Sound envi Control an Stimuli and Subjects . 1: Threshold Equal loud Results . 2: Threshold	Idea intil procedule       Idea intil procedule         ilter asymmetry and filter shifts       Idea intil procedule         ip       Idea intil procedule         ip       Idea intil procedule         ironment       Idea intil production system         Idea intil generation       Idea intil procedule         Idea intil generation       Idea intil production         Idea intil production       Idea intintil production         Idea	<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> <li>50</li> <li>50</li> <li>50</li> <li>51</li> <li>51</li> <li>51</li> <li>52</li> <li>53</li> <li>54</li> <li>55</li> <li>56</li> </ul>
3	Lista 3.1 3.2 3.3	2.5.2 ening ex Measu 3.1.1 3.1.2 3.1.3 3.1.4 Phase 3.2.1 3.2.2 3.2.3 Phase 3.3.1 3.3.2	Auditory fr Auditory fr aperiments rement set-u Sound envi Control and Stimuli and Subjects . 1: Threshold Equal loud Results . 2: Threshold Chosen me Grouping s	ined inting procedure	<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> <li>50</li> <li>50</li> <li>50</li> <li>51</li> <li>51</li> <li>51</li> <li>52</li> <li>53</li> <li>54</li> <li>55</li> <li>56</li> <li>57</li> </ul>
3	Lista 3.1 3.2 3.3	2.5.2 ening ex Measu 3.1.1 3.1.2 3.1.3 3.1.4 Phase 3.2.1 3.2.2 3.2.3 Phase 3.3.1 3.3.2 3.3.3	Auditory fr Auditory fr aperiments rement set-u Sound envi Control and Stimuli and Subjects . 1: Threshold Equal loud Results . 2: Threshold Chosen me Grouping s Pilot tests a	Iter asymmetry and filter shifts       Iter asymmetry and filter shifts         IP       Iter asymmetry and filter shifts         ironment       Iter asymmetry         d sound reproduction system       Iter asymmetry         d signal generation       Iter asymmetry         d of hearing and ELC       Iter asymmetry         of hearing determination: The ascending method       Iter asymmetry         I of hearing with a notched noise masker       Iter asymmetry         and final notched noise configuration       Iter asymmetry	<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> <li>50</li> <li>50</li> <li>50</li> <li>51</li> <li>51</li> <li>51</li> <li>52</li> <li>53</li> <li>54</li> <li>55</li> <li>56</li> <li>57</li> <li>57</li> </ul>
3	Lista 3.1 3.2 3.3	2.5.2 ening ex Measu 3.1.1 3.1.2 3.1.3 3.1.4 Phase 3.2.1 3.2.2 3.2.3 Phase 3.3.1 3.3.2 3.3.3	Auditory fr Auditory fr aperiments rement set-u Sound envi Control an Stimuli and Subjects . 1: Threshold Equal loud Results . 2: Threshold Chosen me Grouping s Pilot tests a 3.3.3.1	intenting procedure	<ul> <li>43</li> <li>46</li> <li>49</li> <li>50</li> <li>50</li> <li>50</li> <li>50</li> <li>51</li> <li>51</li> <li>51</li> <li>51</li> <li>52</li> <li>53</li> <li>54</li> <li>55</li> <li>56</li> <li>57</li> <li>59</li> </ul>

			3.3.3.3	Measuring threshold using one or the other asymmetrical con- figuration	60	
			3.3.3.4	Final notched noise configuration	60	
		3.3.4	Final met	hod and grouping strategy	61	
4	Results and analysis					
	4.1	Maske	d threshold	l results as a function of notched noise masker separation	63	
		4.1.1	Overall v	ariation in threshold	66	
		4.1.2	Shape of	the threshold curves	66	
		4.1.3	Specific of	cases	66	
	4.2	Audito	ry filter sh	apes and ERB	67	
		4.2.1	Applying	the fitting procedure to the obtained data	67	
			4.2.1.1	Final modifications to the fitting procedure	67	
			4.2.1.2	Examples of fitted threshold curves and fitting performance	68	
		4.2.2	Obtained	auditory filter shapes and ERBs	69	
			4.2.2.1	Auditory filter shifts	73	
			4.2.2.2	ERB and detection efficiency	73	
	4.3	ERB as	s a functio	n of frequency	74	
		4.3.1	Difference	es in the ERB values obtained at each center frequency	75	
		4.3.2	Approxir	nate expression for the ERB in the range 31.5-125 Hz	75	
5	Con	clusions	5		77	
	5.1	Discus	sion		77	
	5.2	Conclu	nclusions			
	5.3	Further	r work .		79	
Bi	ibliog	graphy	7		84	
<b>A</b> 1	nnan	div			88	
A	phen	uix			00	
A	Ana	tomy of	the audit	ory system	89	
	A.1	The pe	ripheral au	iditory pathway	89	
		A.1.1	The outer	ear	89	
		A.1.2	The mide	lle ear	90	
		A.1.3	The inner	ear	91	
		A.1.4	Transduc	tion from mechanic waves to neural signals	92	

	A.2	The central auditory pathway	93					
		A.2.1 The auditory cortex	94					
B	Human sound perception 97							
	<b>B</b> .1	Threshold and loudness	97					
	B.2	Masking	98					
С	The low frequency room at Aalborg University 101							
	<b>C</b> .1	General description of the room	101					
	C.2	Reproduction modes	102					
		C.2.1 Pressure field mode	102					
		C.2.2 Free field mode	102					
		C.2.3 Providing a hybrid reproduction mode	104					
	C.3	Listener position in the low frequency room	105					
		C.3.1 The effect of the presence of the chair	105					
		C.3.2 The effect of the position of the chair	106					
		C.3.3 The effect of the presence of a listener	106					
		C.3.4 The effect of differences in sitting positions and listener's heights	108					
	C.4	Compensating the non-flat response of the room for the hybrid mode	108					
D	Thre	eshold determination methods and time consumption	111					
	D.1	The 2 down-1 up 2AFC method	111					
	D.2	The ascending method	113					
E	Thre	eshold and equal loudness contours results	115					
F	Generation of notched noise signals using MATLAB 11							
	F.1	Step 1: Design of the notched-noise filters	119					
		F.1.1 Determination of filter type and filter parameters	119					
		F.1.2 Testing the filter response	121					
	F.2	Step 2: Applying the ELC shaping filters	122					
	F.3	Step 3: Applying the hybrid mode compensation filter	122					
G	Liste	ening experiments: Instructions sheets	123					

# **Chapter 1**

# Introduction

This Master Thesis deals with a very specific field of psychoacoustics: the *auditory filters at low frequencies*. In this introductory chapter, the reasons for researching in such field and the principal objectives of the project are given. The main concepts involved in the phenomenon of the auditory filters and the theories and models that support these concepts are also reviewed. Finally, the specificities and difficulties that are added when working at very low frequencies are exposed.

# 1.1 Initial problem

The human being has always been exposed to noise, nevertheless the recent arise of new noise sources and presence of higher sound pressure levels have caused an increasing concern about the negative effects of this exposure. Environmental acousticians frequently agree in establishing a first classification of noise regarding the nature of the source that originates it. Then noises might be classified as:

- 1. Natural noises: This group comprises every sound that might appear naturally in our environment, such as those provoked by flow of water or the wind through tree leaves.
- 2. Human noises: All sounds derived from the normal development of human being life, e.g. steps on a corridor or conversational background noise are included here .
- 3. Machinery noises: Those sounds that come from the functioning of any kind of machines.

Noises of the third group have demonstrated to be the most annoying. Therefore, their study has become essential and the perception they provoke on humans has to be carefully determined to evaluate the annoyance they might cause.

The fundamental presence of low frequencies can be found by analyzing machinery noise spectral composition. This has led researchers to deeper investigate human perception of low frequencies. The increasing interest in low frequency perception may be observed in many examples. The approaches differ, as some of them are oriented to the study of the origins of annoyance, and some of them look more into annoyance itself. Nevertheless, probably the most investigated topic in low and infrasound frequencies is the threshold of hearing and equal loudness contours (see appendix B). In [11], a summary of the most relevant works regarding this topic is provided. These investigations, however, might not be enough to characterize noise perception in most cases, as thresholds

and contours are obtained for single tones. Real noises will likely comprise broadband noises with more than one frequency component. Later investigations, such as [2], are focusing more in the psychological effects of such noises.



Figure 1.1: Two tones at different but close low frequencies, both below the hearing threshold of a subject are presented independently. They are, by definition of hearing threshold, **inaudible**.

A particularly illustrative example of the limitations when describing hearing perception by means of the standard equal loudness contours was found in [30]. An illustrative example is the situation reflected in figures 1.1 and  $1.2^{1}$ . The case on the left side of figure 1.1 shows an inaudible single low frequency tone. On the right we can see the same case, but for a tone of slightly higher frequency. When isolated, both sounds are not audible for the human ear, as they are below the threshold of hearing for tones. On the other hand, if they are presented simultaneously (see fig 1.2), the result might be audible. This can be explained by means of the auditory filter concept: regarding loudness, auditory filters integrate the components of the sound that fall within one critical band. Therefore, individual inaudible sounds might give rise to a complex audible.

The same effect might be demonstrated for other frequency ranges as well. Nonetheless, as mentioned before, perception of low frequency sounds is becoming more an more interesting, as they are the range for a big part of annoying noises. Therefore, it seems not only convenient but necessary to identify the auditory filters also at very low frequencies. This can be used i.e. to predict threshold or annoyance of such noises given their spectrum. Such identification has not been achieved yet as far as the authors of this thesis know, both in terms of the bandwidth and shape of the auditory filter for frequencies below 100 Hz.

<sup>&</sup>lt;sup>1</sup>Figures adapted from http://hyperphysics.phy-astr.gsu.edu



Figure 1.2: The previous two inaudible tones are now presented simultaneously resulting in an **audible** complex sound due to integration performed in the hearing system and explained by the action of the so called *auditory filters*.

# **1.2** Scope of the project

The aim of this Master Thesis is to provide a better knowledge about auditory filters in low frequencies. As far as the authors know, no precise results describing the auditory filter performance have been shown yet in the range below 100 Hz. Therefore, it is not known in which manner sound is being integrated in our hearing system and we cannot predict precisely the perception of complex sounds at these frequencies. As long as the assumptions made during this work are correct, the results achieved could allow a better understanding of the performance of the hearing system in low frequencies. This would be useful, for example, in order to avoid annoyance due to low frequency noises. Sound sources could be treated in advance, because it would be known what their effects would be. Besides, models to predict threshold or loudness could be improved with a better understanding of the main properties of auditory filtering in this region.

In principle, the auditory filters are assumed to act in low frequencies in the same way they do in higher frequencies, with some particularities. Thus, interest is focused on the same properties that have been obtained for other frequency ranges: bandwidth and filter shape. Variation of these parameters with frequency is also a main consideration, as it has been suggested that the tendency of a steady decrease in bandwidth with decreasing frequency is not fully valid anymore in low frequencies (see figure1.5 and references given in it). For this reason, different center frequencies in this frequency region will be considered. This will allow to establish comparisons, both in shape and bandwidth, between the corresponding auditory filters.

The main focus is in the frequency domain, i.e obtaining the ERB and filter shape. Other representations, such as functional time domain representations are, however, out of the scope of the project. Some other related research matters, such as evolution of the auditory filters with age, or differences between normal and impaired listeners, are also beyond the scope of this project.

# **1.3** The critical band and the auditory filters

Although the critical band concept has been reviewed along the last 60 years, it has always referred to the same phenomena beyond the concept. This is the fact that our hearing system analyzes incoming sound splitting the audible frequency range into a number of limited bands. The manner in which this analysis is performed is a complex process that suggests the existence of what can be considered 'frequency filters' in our auditory system. This has given rise to a parallel newer concept to that of the critical band: the 'auditory filters'. Nowadays this concept is preferred, and the authors of this project agree on using the term 'auditory filter' instead of critical band. The critical band term can be considered as a rough approximation of the auditory filter. However, it fails to describe in more detail how the filtering is done. It is considered then, that the auditory filter will provide a more detailed description of the filtering carried out by the auditory system.

Therefore, a short review on the origins of the auditory filter concept is given. A chronological review of the concept seems appropriate. Therefore, as starting point the critical band concept is further described. Following this, main assumptions of the masking model that was used for determining the critical bands are given. Then, the auditory filter concept is described in more detail.

## **1.3.1** The critical bands

In 1940 H. Fletcher explained an experiment that gave rise to the critical band concept. This experiment consisted on measuring the detection thresholds of a sinusoidal signal as a function of the bandwidth of a bandpass noise. The band of noise was characterized by a constant power density and its center frequency was at the frequency of the signal. He found that the threshold of the signal increased with increasing masker bandwidth, up to a width where the threshold stabilized. He named this bandwidth the critical band. Using this concept he also explained two related phenomena:

- Wegel and Lane had shown in 1924 that a sinusoidal signal is most easily masked by other sinusoid when they are the closest to the signal.
- Fletcher's own work showed that when the masker is a noise, it is the components near the the signal that dominate the masking process.

As a result Fletcher proposed a listening model in which the listener processes the sound through a bandpass filter centered at the frequency of the signal. The more powerful the masking noise or components in the noise are, and the closest they are to the signal, the more they influence (mask) the perception of the signal. He suggested that, for predicting masking, the filter shape could be simplified by a rectangle and he named the 'critical band' the width of this rectangle filter. Even though this shape was not realistic, derived results from this model presented good approximations that brought attention to Fletcher's model and to the critical band concept.

Fletcher suggested that the whole auditory system performs a frequency analysis by means of many of those (simplified) auditory filters and he assumed that the basilar membrane was responsible for this frequency analysis. He also suggested that the overall performance of the auditory

system could be seen as a frequency analyzer with overlapping filters. These filters would be centered at the same frequencies of the main components of the incoming sound, and the mentioned critical band would be their bandwidth.

Even though Fletcher proposed simplified rectangular filters, such a model was able to explain many other psychoacoustic effects and the critical band concept brought up great interest. Fletcher himself continued with his research on the topic and Zwicker went further in his investigation. He enumerated a number of phenomena reflecting the existence of a such filtering and provided methods for measuring its bandwidth. Zwicker also measured the critical bandwidth along almost the whole audible frequency range. His results showed a steady increase in the critical bandwidth with increasing frequency, except for the region below 500 Hz. In this region he observed an almost constant bandwidth of 100 Hz.

The concept of critical band exists also terms of loudness. The *critical band for loudness* has been already particularly distinguished by a number of researchers [13]. When a narrow noise band is progressively broadened in such a way that the overall energy remains the same, it has been found that the perceived loudness does not vary up to a certain bandwidth. This bandwidth has been named 'critical bandwidth for loudness' and differs from bandwidths found based on other criteria.

Another use of the critical band concept is regarding frequency discrimination. Suppose we are hearing a signal composed of to two tones of the same frequency. Our hearing system will tend to fuse the two tones and they will be perceived as a single one. However, as the tones are separated in frequency in a gradual manner, different sensations will be experienced by listeners. When the tones are sufficiently close, listeners will still perceive the signal as composed by a single tone. Increasing the separation will produce the sensation of beats. When the tones are more separated, listeners will still perceive one sound described as rough. This sensation remains until the separation reaches a point where two different tones are clearly heard. The minimal interval for which a listener is able to perceive two different tones has also been described as a critical band.

However, when Fletcher used his critical band concept for the first time, he was presenting a frequency analysis model of the ear which he deduced from specific already known phenomena. The phenomena were intrinsically related to the *masking* concept (detailed in Appendix B). The following section describes the fundamental assumptions that have allowed to determine the critical band and describe the auditory filtering in our hearing system.

## **1.3.2** The power spectrum model

The masking effect was somehow the origin of the critical band concept. Therefore, masking was a fundamental part in the first model used to calculate the critical band. The model relating masking to critical band is called *power spectrum model of masking*. It assumes that when trying to detect a signal in a noisy background, the listener uses a filter with center frequency positioned close to that of the signal. The filter attenuates a great amount of the noise and passes the signal. Only the noise components that pass through the filter will be the ones that have an effect on masking the signal.

The threshold of the signal is assumed to be dependent on the amount of noise passing through the

auditory filter. Furthermore, threshold is assumed to correspond to a specific signal to masker ratio at the output of the filter. Fletcher considered a masker noise its spectrum being approximately flat in the region of the signal. If the spectral density of the noise is  $N_0$ , then the total power passing through a critical band of width *CB* Hz, is  $N_0 \cdot CB$ . All these assumptions are known as the *power spectrum model of masking*, since the stimuli are represented by their long term power spectra and relative phases of the components and short term fluctuations of the masker are ignored. This model assumes that the noise power is directly proportional to the signal power at threshold  $P_s$ and the proportionality constant is named *K*, so the power spectrum model of masking may be expressed by:

$$P_s = K \cdot N_0 \cdot CB \tag{1.1}$$

or else,

$$P_s = K \int_{-\infty}^{\infty} N(f) W(f) df$$
(1.2)

where *K* is the proportionality constant, *f* is the frequency, N(f) represents the long term spectrum of the masker, and W(f) is the auditory filter weighting function, [21].

The value of the constant K specifies the efficiency of the detection process, revealing frequency independent effects such as practice or inattention [21]. It equals the signal to masker ratio of the filter required to achieve threshold, a smaller K indicating more efficient processing.

Initially, the shape of the filter was not considered in this model. However, the simplified model has been the basis for some experiments that have achieved satisfactory results. For example, the tendency of increasing bandwidth with frequency in the critical band has been shown with this model. This was done even though the model could not reflect anything regarding the filter shape or changes in filter shape with frequency. Later works have reflected the same tendency when considering the shape of the auditory filters [3].

However, rectangular filters seemed unrealistic and, moreover, did not explain phenomena such as asymmetrical masking patterns. Because of this, most of the effort of recent research related to auditory filtering has been centered in finding methods for determining the shape of the auditory filters.

### **1.3.3** Auditory filters: bandwidth and shape

A simplification of the effective bandwidth of the auditory filter does not seem to be a very accurate approximation and fails to explain phenomena such as asymmetrical masking patterns. After Fletcher and Zwicker, new researchers investigated parallel phenomena and suggested a new concept: the auditory filter. Both, the critical band and the auditory filter concepts, are closely related. However, the second comprises more aspects than the bandwidth. Important researchers, such as B. Moore and R. Patterson, along with other important collaborators, developed measurement methods for determining the properties of the auditory filters. The filters were no longer considered rectangular, as in Fletcher's model, but with a determined shape.

The main limitation of the critical band model was that it assumed unlimited attenuation outside the passband and flat response within the passband. Figure 1.3 depicts a schema of the critical band and auditory filter models. Four auditory filters are shown, one next to the other, assuming two different shapes:

- 1. At first the rectangular filters proposed by Fletcher are shown (dashed lines), which manifest abrupt cutoff frequencies. Considering signals in the neighborhood of the cutoff frequencies for this model, no interaction could be possible between signals that fall in different filters.
- 2. Secondly, auditory filters with a more realistic shape are shown (solid lines). It has been found that the auditory filters tend to have relatively sharp filter skirts, but a flatter response just around the center frequency [3] [12].



Frequency

Figure 1.3: Two models of auditory filter: Rectangular critical bands (dashed lines) and a more realistic representation (solid lines). The grid area shows the region where both models completely disagree.

The non-rectangular shape has been shown to be more realistic after a number of experiments. One of them, [27], consisted of determining the masking produced by a low passed noise with cutoff frequency roughly below the supposed low frequency limit of the auditory filter. The auditory filter was at the same time excited by a higher frequency tone. The masked threshold of the tone was found to be above its threshold in silence. This shows that the shape of the filter is important and is far from what Fletcher suggested. A substantial difference in the two approaches may easily be observed from the grid part of figure 1.3.

Different experiments performed by different researchers have given very similar results regarding the shape and bandwidth of the auditory filter, as well as regarding its variation with level and frequency. While this is true for middle and high frequencies, in lower frequency range, however, higher discrepancies are observed, specially concerning the shape of the filter [27].

The disagreement observed in the low frequency region is not the main surprise when observing all these works. The most interesting fact -and the one that has given rise to this project- is that no results are known to exist for really low frequencies, i. e. below 100 Hz.

#### 1.3.4 Equivalent rectangular bandwidth

Despite the findings suggesting the auditory filter shape is far from the rectangular simplification suggested by Fletcher, the importance of this model has led researchers to specify correspondences between the new values of auditory filter bandwidth and the rectangle model values. Thus, for a given auditory filter shape, its equivalence into a rectangular band is referred as the *equivalent rectangular bandwidth*, commonly expressed as ERB. The ERB is equal to the bandwidth of a perfect rectangular filter which has a transmission in its passband equal to the maximum transmission of a specific filter, transmitting the same power of white noise as the specific filter [12]. A more realistic auditory filter and its respective ERB are represented in figure 1.4.





Figure 1.4: The equivalent rectangular band compared to a more realistic approximation of an auditory filter.

Mean ERB values obtained using moderate sound levels after tests on young people with normal hearing are frequently denoted as  $ERB_N$ . In [3] an approximated formula that relates the  $ERB_N$ 's with frequency is presented:

$$ERB = 24.7(4.37F + 1), \tag{1.3}$$

where F is expressed in KHz.

This equation shows a pseudo-logarithmic increase of the critical bandwidth with frequency. This is consistent with other perceptual phenomena already demonstrated, such as the logarithmic law of frequency interval perception [6].

# 1.3.5 Off-frequency listening

Until now it has not been explained clearly how sounds are analyzed through the so called auditory filters. Auditory filters cannot be considered as an analogy to passive electronic filters. In fact, these filters are active and their performance depends on some other factors. They have been found to be level dependent [24] and they also may adapt to the frequency of the signal [20]. This last feature means that the filters tend to be centered at the main spectral components of the sound. If spectral components of the signal cover broader bandwidths than the critical bandwidth for a certain auditory filter, the response of more than one filter might prevail. However, each filter will probably still be centered close to the main components of the signal.

Nevertheless, it has been suggested [18] that the filter might not always be exactly centered at the main component of the signal. When there is a predominant partial in a complex sound above a background irregular noise, the filter is probably placed in such a way that the signal to noise ratio is best for the partial at the output of the filter. The process when a listener shifts an auditory filter to improve the S/N at the output of the filter, is referred to as *off-frequency listening*.

Off frequency listening is an important factor to take into account when finding the shape of the auditory filters. This concept, and how experiments can be designed to prevent its effect to interact with the measurements, will be further explained in following sections.

# **1.4** Physiological origin to the auditory filters

The origin of the auditory filters is not fully clear yet. Probably there is more than one organ and more than one process involved. These are the reasons why detailed description of their performance is not simple, as many factors may influence it. However, auditory filtering is closely related to some specific parts of the auditory system. Specially those parts that participate actively in the determination of frequency and loudness of the incoming sounds. Which are these organs and in which way they are related to the auditory filters is briefly explained here.

Appendix A provides a rather complete description of the hearing system. Regarding frequency analysis, the inner ear and particularly the cochlea play an essential role. A deeper look into the cochlea's performance and into the basilar membrane response may help to understand this frequency analysis. It is necessary to remind at this point that two theories are the most accepted nowadays regarding frequency analysis of incoming sound: the *place theory* and the *temporal theory*.

## 1.4.1 The *place* theory

The *place* theory proposes that frequency analysis and sound coding is made according to the area of the basilar membrane which has been excited by the incoming sound. This means that for a determined frequency, the basilar membrane, due to its mechanical properties (see appendix A), will respond with a specific pattern of vibration. This pattern will have a maximum in a single area. Bekesi in 1960 was the first among a number of researchers that confirmed the fact that the point of maximum displacement along the basilar membrane changes as the input frequency is

varied. It has also been shown that the linear distance from the point of maximum displacement to the *apex* is approximately proportional to the logarithm of the input frequency. The specific movement originated by certain frequencies will make the hair cells in the *organ of corti* attached to a corresponding area of the basilar membrane to bend. This will generate neural firings that will run through the nerves that form the auditory nerve to superior neural centers.

Features of our hearing system, such as frequency resolution or masking, have been explained by means of the place theory. Then, masking would occur when a masker generates a pattern in the basilar membrane that is close in frequency to that of the signal and sufficiently stronger in intensity. As a consequence, the maximum displacement point of the masked sound will strongly be affected by the pattern of the masker (even if it is not by its maximum). The frequency resolution will depend on the minimum frequency difference for which there will be two distinct maximums in the basilar membrane.

In terms of the auditory filters, the place theory has turned out to be of mayor importance. In fact, researchers [13] have discovered a correspondence between the ERB and a constant distance along the basilar membrane. In humans 0.9 mm correspond roughly to one ERB<sub>N</sub>.

## **1.4.2** The temporal theory

Some experiments that could not be explained by the place theory have given rise to a second theory regarding frequency coding in our hearing system. This theory, the *temporal theory*, indicates that frequency information is coded by phase locking to basilar vibration. This is carried out by specialized neurons that are able to combine their performance. This allows the hearing system to code higher frequencies than the corresponding to the minimum relaxation period of a single neuron.

It highly probable that both systems are combined. This may take place in specific frequency ranges and will lead to a better frequency resolution below around 4000 Hz, where there is not temporal coding anymore.

Normally, the critical band concept has been explained based on the older place theory. However, the temporal coding has also been briefly presented, as the main concern of this project is dealing with frequency resolution as well. Given that this kind of coding is also contributing to it, it might also influence to some extent the auditory filters.

## 1.4.3 Auditory filters, loudness and basilar membrane performance

Frequency and loudness are deeply influenced one by each other. Thus, the place theory might be applied to explain the phenomenon of the critical band for loudness. The ear seems to lump all the energy within an auditory filter together and treat it as one item of sound. Therefore, when all sound energy is concentrated within an auditory filter, loudness is proportional to the total intensity of the sound within that auditory filter. When the frequency components of the sound extend beyond an auditory filter, an additional effect occurs. In this case, more than one band is contributing to the perception of loudness and the brain appears to add the individual critical band responses to-

gether. The effect is to increase the perceived loudness of the sound even though the total acoustic intensity has not changed [13].

A psychophysical explanation is that, in the cochlea, the critical bands are determined by the place at which the peak of the standing wave occurs. Therefore, all energy within a critical band will be integrated as one overall effect at that point on the basilar membrane and will be transformed into nerve impulses as a unit. On the other hand, energy which extends beyond one critical band will cause other nerves to fire and it is these extra nerve firings which give rise to an increase in loudness [13].

# **1.5** The auditory filters at low frequencies

Up to this point, the concept of auditory filters and some of the most important related phenomena have been described. Now, a deeper look will be taken into the particularities of auditory filtering and its measurement, when dealing with low frequencies.

The determination of the effective bandwidth of the auditory filters has already been subject of many studies. However, results go only down to around 100 Hz. Besides, the lower the results go, the more they disagree between the different researches. Fletcher proposed originally a an exponentially increasing critical bandwidth from 500 Hz upwards. Below 500 Hz, he proposed a stable bandwidth of 100 Hz. Later results, however, suggest that the exponential evolution of the critical band goes much lower, although in low frequencies it seems to be shallower [12]. In figure 1.5, these results are presented. In the figure, it can be seen that no data exists at frequencies below 100 Hz, even though the great importance of this frequency region.



Figure 1.5: Results of critical bandwidth as a function of frequency after different researches,[13]

The problems in the determination of the auditory filters at such low frequencies come from different reasons. Some of them will be presented in detail in following sections, when presenting the possible methods for psychoacoustic measurements to be carried out. Only a rough explanation is given at this point.

The concept of the auditory filters is theoretically the same for all frequencies. Therefore, *assumming that the principles observed in higher frequencies might be applied in low frequencies as well* -which might not be true-, the limitations in low frequencies are practical. For example, the limitations could appear when trying to run the experiments, not being a direct consequence of a wrong theoretical approach. A division into two groups of low frequency constraints can be made:

- 1. Perceptual problems due to the extreme characteristics of the signals required for the listening tests.
- 2. Limitations due to difficulty in the practical presentation of such signals to the subject.

In the first group we find the constraints caused by the human hearing system limitations, in terms of both frequency and loudness. Regarding frequency, if the same measurement conditions applied in the tests that led to the results of figure 1.5 were directly applied below 100 Hz, in some cases even negative frequencies would be required. This is of course impossible and therefore, adjustments have to be introduced in the methods and some of their parameters revised in order to be able to provide results at very low frequencies.

As to loudness, the problems come because of the threshold of hearing at low frequencies. Two of the main problems regarding this issue are:

- 1. The threshold of hearing is relatively high and presents a steep increase towards the lower frequencies. This demands the presentation of high signal levels to the subject, which can be annoying or even impossible because of dynamic range limitations in the sound reproduction equipment.
- 2. As the equal loudness contours present a steep slope, the test signals should be compensated according to this contours, being enough above threshold of hearing. This would require, firstly to determine the threshold of hearing at very low frequencies. In fact, for some of the methods, the threshold for infra-sounds would be needed. Secondly, the total level of the signals presented has to be very high, at the risk of approaching annoying levels for the subject.

In the other group of tests constraints, limitations in presenting the required signals to the subject are found. As mentioned above, relatively high signal levels at very low frequencies are required. The generation of the signals should not be a problem, but the reproduction of those signals may present some difficulties. In principle, there are two options: either presenting the signals in a quiet sound environment with a set of headphones or using loudspeakers.

In both cases, the response of whole chain should be considered and compensated. When using headphones, compensating for the effects of the room is avoided. On the other hand, compensating for the headphone response can be even more problematic or time consuming than for loudspeakers. Moreover, the situation using a loudspeaker-room system seems to be more realistic. In this latest case, a room and a loudspeaker or a set of loudspeakers would be needed to be able to grant a sufficiently good response at very low frequencies.

There is another link in the chain that could be taken into account: the outer and middle ear effects. It is necessary to make another consideration first. It is needed to define what is precisely being measured when evaluating the auditory filters. Auditory filtering can in principle be considered either as a consequence of the response of whole hearing system, or because of the response of the cochlea, the organ that is likely causing such behavior of the system. From an audiological point of view, it seems more interesting to try to isolate the performance of the cochlea. This has been the approach in some previous studies [15], [3]. From an acoustical point of view, more worried about effects on the subjects due to external excitations, considering the whole hearing system might be the best option.

In any case, the emitter-receiver chain could be compensated, either by means of pre-processing or using post-processing of the signals. Some researchers [15], [3] have preferred the corrections applied by numerical methods after the tests were done, i. e. post-processing. However, it should be also taken into account that auditory filters are level dependent. Then, an approach where the original signals already include the compensation seems to be more appropriate. Thus, the level that reaches the cochlea is the desired one for the test signal. On the other hand, this might not be a critical factor if the results obtained show that the auditory filters tend to be less affected by level at very low frequencies [3], [24]. If this is true below 100 Hz, then probably a negligible difference would be found if the compensation is made before or after the test is done.

The main constrains when performing measurements for the determination of the auditory filters at very low frequencies have been roughly explained here. Probably, these problems have been one of the main reasons why auditory filters at center frequencies lower than 100 Hz have not been investigated until now, which is our purpose. All methods presented in next section were considered so as to be able to overcome the problems just presented. One of the methods was chosen and is the basis of the method followed in this work. The changes and adjustments that were required to achieve results at very ow frequencies are presented in the report as well. The main adjustments were applied when setting up for the measurements and when defining some of the parameters in the experiments. This process will be detailed in Chapter 3 and Chapter 4.

# **Chapter 2**

# Method

The previous chapter has given the reader an idea about the goal of this project. In the following, the method employed to achieve this goal will be described. As different approaches can be used to experimentally estimate the characteristics of the auditory filters, a whole section is dedicated to describe the most commonly accepted methods. A brief discussion explains their advantages and disadvantages.

Following this, the chosen method is described apart in detail: *the notched noise method*. This method was chosen and is the basis of the present work. Basically, the approach consists on finding some parameters that characterize the filters, and based on some filter shape assumptions (the so-called *roex* filter model), recovering the filters.

The basic experimental data required in this method are thresholds for tones in the presence of a notched noise masker in different configurations. Thresholds had to be found through listening experiments carried out on several subjects. Different considerations had to be made in this measurement. They are presented also in this chapter.

When preparing for the measurements, one of the most important aspects was the design of the notched noise stimuli. Therefore, a whole section is dedicated to describe this as well.

Finally, parameters that can characterize the auditory filters are calculated with these measured thresholds. With these parameters, the auditory filters can be derived. The last part of the chapter details this procedure.

# 2.1 Methods for determining the critical bandwidth and the auditory filter shape

The critical bands and auditory filters are one of the topics that have been most studied in psychoacoustics. Researchers have used sometimes totally different approaches and many of them have been considered as possible basis for this study. Some factors have been taken into account to choose the most adequate method. First, the method should allow to evaluate the filter shape; second, the method should be suitable for measuring at very low frequencies; third, if possible the method should be reliable and established or accepted in most of the reviewed literature. All methods considered in this study are described in the following.

## 2.1.1 Band limiting methods

This is the classic method following an experiment performed by Fletcher [13]. The threshold of a sinusoidal signal was measured as a function of bandwidth of a bandpass noise masker. The noise was centered at the signal frequency and the noise power density was kept constant. This means, the overall noise power increased as bandwidth increased. Figure 2.1 shows an example of the results taken from [13]. As the noise bandwidth increases the threshold of the signal increases



Figure 2.1: Threshold of a 2000 Hz sinusoidal signal as function of bandwidth of a noise masker centered at the signal frequency (taken from [13])

until a certain bandwidth were it flattens off. From this point, further increases in bandwidth did not alter the signal threshold significantly.

From the power spectrum model of masking assumptions, increases in the noise bandwidth will result in more noise passing trough the auditory filter, as long as the noise bandwidth is less than the filter bandwidth. When the noise bandwidth exceeds the filter bandwidth, further increases in the noise bandwidth do not increase the noise passing through the filter.

The bandwidth at which the signal threshold ceased to increase was called *critical bandwidth* (Fletcher, 1940). This corresponds to the breakpoint in the threshold curve shown in figure 2.1.

For ease in the analysis of the results, the filter shapes were assumed to be rectangular. For this kind of filter, all components in the passband are passed equally and all components outside are removed.

From the power spectrum model, when the noise just masks the tone, the power of the tone,  $P_s$ , divided by the power of the noise inside the critical band will be equal to a constant *K*. Assuming a white noise with power density  $N_0$ , for an hypothetic rectangular auditory filter of bandwidth *CB* Hz, the total power falling into a critical band *CB* will be  $N_0 \times W$ . Thus, according to equation

1.2, the critical bandwidth could be estimated indirectly by measuring the threshold of the tone in a broadband noise as:

$$CB = \frac{P_s}{K \times N_o} \tag{2.1}$$

The value of K can be seen as a measure of detection efficiency. Most recent critical band measurements show that it varies with center frequency, being markedly higher (less detection efficiency) at low center frequencies [24].

**Discussion:** The results of critical bandwidth estimates using this method are not in complete agreement with measures using the same method, or with measures using other methods, specifically in low frequencies. In the mid and high frequencies different experiments have given reasonable similar estimates of the critical band values as when using this method [13].

However, there are also contradicting results using this method, where no break point in the threshold curve has been found [28]. This suggests the ear is capable of integration over bandwidths much greater than the critical bandwidth.

If the filter is assumed to be symmetrical, this method could in principle be used to determine the filter shape [12]. However, once the noise bandwidth is as wide as the filter passband, further increases in the noise bandwidth produce negligible increases in the total noise power passing through the filter. Therefore, the method is considered not sensitive enough to determine the filter shape in a useful range [12]. Besides, the power spectrum model assumptions may not hold for the narrowest noise bands where energy fluctuations inherent in narrowband noises can have a significant effect on the signal's detectability [19].

### 2.1.2 Two-tone masking methods and psychophysical tuning curves

**Two-tone masking methods** are based on the measurement of the threshold of a tone or narrowband noise signal in the presence of two masking tones placed each at either side of the signal. As the separation of the tones is increased the threshold of the signal drops. The tones can be placed either symmetrically or asymmetrically around the signal. In principle, this method can directly estimate the filter shape [31].

**Psychophysical tuning curves** are obtained in a similar fashion but they use only one second tone and are restricted to low signal levels. The curves aim to represent the output of neurons of similar center frequencies. The results are shown as the power of the second tone required to mask the signal as a function of frequency. If the assumption was valid, the curve can be seen as an inverted auditory filter shape [12]. Figure 2.2 shows an example of the resulting curves.

**Discussion** Some practical problems arise when using this methods. If the signal is a sinusoid, when the masking tones are close to the signal they will beat with the signal. This may provide an additional cue to the subject and the resulting filter will seem to have a notch at its center frequency [12]. Narrowband noises can be used to avoid the beats.



Figure 2.2: Psychophysical tuning curves for a 2000 Hz sinusoid in the presence of a sinusoidal and a narrowband noise masker (circles and squares, respectively), at different levels ((a) 30 dB SL, (b) 20 dB SL (c) 10 dB SL), [12].

Off frequency listening can affect both of the methods [12]. In the two-tone masking experiments, when the masking tones are narrowly spaced, it is not clear whether the subject is using a filter centered at the frequency of the signal or not, because all of the filters in the region of the signal will have similar signal to masker ratios.

Besides, the presence of combination tones produced by interaction of the maskers and the signal can also introduce irregularities in the threshold functions.

If precautions are taken care off to reduce this problems (i.e choosing narrowband signals and appropriate maskers for combination tones), the filters obtained with the two-tone masker method are very similar to the ones obtained with other methods [12].

#### 2.1.3 Rippled-noise method

In this method the filter shapes can be derived by the use of masking experiments with rippled noise. This noise has a long term spectrum that varies sinusoidally on a linear frequency scale. It is produced by adding a white noise to a copy of itself that has been delayed T seconds. The delayed noise can be added in phase or out of phase. The in phase adding produces peaks at 0 Hz and at every multiple 1/T Hz, while the out of phase adding produces deeps at those locations. The spectrum of the rippled noise is:

$$N(f) = N_0 (1 \pm M \cos 2\pi fT), \qquad (2.2)$$

where  $N_0$  is the original noise spectrum and M is the modulation depth determined by the attenua-

tion of the delayed noise [12]).

Based on threshold determination of a pulse sinusoid in the presence of this noise masker according to delay time, T, and polarity of the rippled noise, the filter shape has been determined [5]. This was done using the general masking model and a trigonometric Fourier series application.

**Discussion** The filter shapes that have been determined using this procedure are similar to the ones obtained with other methods. However, there can be practical difficulties in the derivation of the filter shapes when the ripple densities are high due to threshold differences being comparable to measurement errors [12].

Furthermore, even in low ripple densities, small measurement errors can give noticeable irregularities in the filter shape. The dynamic range of the measurements is considerably less than in other methods and it requires more subject time in general [12].

### 2.1.4 Loudness experiments

As said before, the concept of critical band is not unique and can be applied to loudness perception. The loudness of a complex sound of fixed energy and bandwidth W has been found to be independent of W as long as W is less than certain value. After this value the loudness begins to increase with bandwidth. The value where the breakpoint begins has been called CB<sub>L</sub>, or critical band for loudness [31], [13]. The values of CB<sub>L</sub> have been found to be a little greater than the ERB<sub>N</sub> [13].

A possible measurement might consist simply of providing the subjects with a band increasing noise of constant energy. The  $CB_L$  would be the bandwidth for which the loudness begins to increase, as it would reveal that more than one filter contains energy.

**Discussion** The exact value of  $CB_L$  is hard to determine [13]. Besides, the resulting loudness of a complex tone will be independent of bandwidth at low sensation levels (10-20 dB SL). Thus, high sound pressure levels would be required to measure the  $CB_L$  at low frequencies. However, this is considered an interesting alternative approach for the determination of a critical bandwidth.

### 2.1.5 Perceptual frequency selectivity experiments

When two tones of fixed amplitude are presented together different perceptual effects happen when they are close in frequency. This effects are similar as those evidenced in the two-tone masking experiments and psychophysical tuning curves. The difference is that here no masking is required and the tones are fixed in amplitude. Very close in frequency the tones beat together and they are perceived as a fused tone which sounds "rough". As frequency separation increases the fused tone gives way to two separate tones which still sound rough. With a further increase in frequency separation the two tones are perceived no longer as separate rough but a "smooth" separate sensation starts and persists within the range of the listener's hearing [6].

The point where the two tones are distinguishable can be seen as a point where two different

peaks on the basilar membrane emerge from a single maximum. If the separation of the peaks is sufficient they would be perceived as smooth separate tones [6].

The critical bandwidth, in this terms, can be obtained from the frequency difference between the tones where the listener's perception changes from "rough" and separate to "smooth" and separate. Therefore, the subjective perception of smoothness can be used as a measure of the critical bandwidth.

**Discussion** Not much reference material describing in detail the application of this method has been found. Because of its nature, it would only provide a measure of the critical bandwidth and it would not give an estimate of the auditory filter shapes. Most accepted and commonly found estimates of the critical bandwidth are based on other methods. The subjective sensations of "roughness" and "smoothness" may be considered weak when compared to a threshold determination in the presence of a masker. However, this method seems to be of simple application and its use was not discarded in case alternative measures of the critical band were required or other methods presented too many difficulties at low frequencies.

## 2.1.6 Notched noise method

This method allows the determination of the auditory filter shape and was designed to prevent off frequency listening. Figure 2.3 shows an illustration of the method, as proposed by Patterson, (1976) [18]. A signal fixed in frequency is located symmetrically at the center of a noise masker with a bandstop or notch centered at the signal frequency.



Figure 2.3: Notched noise method schematic illustration

The deviation from each of the noise edges to the signal frequency is denoted by  $\Delta f$ . The measurement consists of determining the signal threshold for different notch widths, while maintaining the level of the noise masker constant. Since the signal is symmetrically placed at the center of the notch, the method cannot reveal any filter asymmetries.

As the width of the notch is increased, less and less noise leaks through the filter skirts and the threshold is reduced. The variation in threshold with notch width can be seen as a measure of the area of the noise leaking through the filter skirts.

Then, assuming that threshold corresponds to a constant signal to masker ratio, the filter function can be obtained by differentiating the threshold function respect to  $\Delta f$ , given that the integral of a function between certain limits corresponds to the area under that function. This is the basic idea that has been used to determine the filter shapes using this method.

#### 2.1.6.1 Asymmetric notch noise method

In its original form [18], the notched noise method did not allow the measurement of filter asymmetries. In order to determine filter asymmetries, the method has been extended so that the signal is not placed symmetrically at the notch center [20]. In this manner, differences in the slopes of the upper and lower filter skirts can be determined, (see figure 2.4). In this case, the differences between the signal frequency and the lower and upper edges of the noise bands are  $\Delta f_L$  and  $\Delta f_U$ , respectively.



Figure 2.4: Asymmetric notched noise method schematic illustration

**Discussion** This method allows a measurement of the critical bandwidth (or the ERB) and the calculation of the auditory filter shapes. It does not present the practical difficulties of two-tone masking experiments, although the filter shape calculation is not direct. It requires less time and has a better signal to noise ratio than the rippled noise method [12].

Its variations allow the measurement of filter asymmetries and to account for certain degree of off frequency listening. There is generally good agreement between different studies estimating the ERB with this method [14]. Besides, the notched noise method has been used to obtain filter shapes at low frequencies (down to approx. 100 Hz [3], [24]). Therefore, for all this reasons, it is the primarily chosen method for the critical bandwidth determination at low frequencies.

In the next section, a more detailed description of this method is given. Proposed modifications and further description of the used technique are described in the next chapter.

# 2.2 Auditory filter calculation and the notched noise method

The basic idea behind the notched noise method is that the threshold as a function of  $\Delta f$  represents how the area of the two-band noise masker leaks under the auditory filter skirts (for example see figures 2.3 and 2.4).

Thus, taking the derivative of the threshold function respect to  $\Delta f$  would allow to determine the filter shape. The relationship between threshold, noise spectrum and filter shape comes from the power spectrum model and is given by eq. 1.2. By differentiating both sides of eq. 1.2 respect to  $\Delta f$  and choosing a masker spectrum that simplifies the integral, the weighting function of the auditory filter can be determined.

For example, if it is assumed that the noise has flat spectrum,  $N_0$ , and the distances from the tone frequency to the lower and upper edges of the notch are  $\Delta f_L$  and  $\Delta f_U$ , respectively (see figure 2.4), the masking equation 1.2 will take the form [20]:

$$P_{s} = KN_{0} \int_{-\infty}^{f_{0} - \Delta f_{L}} W(f) df + KN_{0} \int_{f_{0} + \Delta f_{U}}^{\infty} W(f) df$$
(2.3)

In this manner, the filter function is left on its own under the integral. In order to solve this equation and determine the critical bandwidth and filter shape in a way that describes the auditory filter behavior as close as possible to the expected real situation, different assumptions regarding the auditory filter properties have to be made. They are presented next.

#### 2.2.1 Auditory filter assumptions

The following points describe the auditory filter assumptions that are often made.

**A) Filter symmetry assumptions:** Filter shapes obtained with other methods [5], [17], were found to be approximately symmetrical leading to suggest the use of the symmetric notched noise method to determine filter shapes [18].

Under these conditions  $\Delta f_L = \Delta f_U = \Delta f$ , and eq. 2.3 will take the form:

$$P_s = 2KN_0 \int_{-\infty}^{f_0 - \Delta f} W(f) \tag{2.4}$$

The symmetric filter assumption has been found to be valid only at moderate levels and middle frequencies [18], [3]. At higher levels the lower skirt becomes broader, increasing the filter bandwidth generally [9], ending up with a markedly asymmetrical shape [12].

At low frequencies the auditory filter shape has been found to be asymmetrical in general [15], [3]. Therefore, in this project, the measurement technique will take into account the possible asymmetry in the auditory filter shapes at low frequencies. To obtain the filter shape, the modeled filter transfer function will then have to describe the filter skirts individually for each side of the filter.

**B)** Filter position assumptions: Early studies for the determination of the auditory filter shape assumed the filter was centered at the frequency of the signal (see i.e [17] and [18]). More recent studies, however, do not support this hypothesis and now it seems clear that the filter shifts to maximize the signal to noise ratio at its output. This is called the "max(S/N)" assumption [20],[3]. The amount of shift is expected to be closely related to the underlaying asymmetry of the filter and the relative position of the noise maskers.

The shift of the filter position at low frequencies will be studied from the obtained data and considered for a more proper derivation and modeling of the filter shape.

**C)** Filter bandwidth assumptions: The variation of the ERB (see section 1.3.4) with frequency has been estimated in different studies [?], [14],[13]. If it is assumed that the filter is not fixed in frequency and the shift is not negligible, a change in filter bandwidth can be expected [3]. Therefore, the change in bandwidth with filter shifting has been included in the fitting procedure described in [3]. However, since no model of the change in bandwidth of the auditory filter at very low frequencies has been found to exist, the change in bandwidth with filter shift will not be considered in the fitting process. Nevertheless, it is expected to use the ERB data as a function of center frequency to estimate the change in bandwidth of the auditory filter at very low frequencies.

In order to use the threshold data to obtain the filter shapes, a filter function has to be assumed and its parameters have to be fitted to the data. The relationship between the threshold data and the filter transfer function can be obtained from an expression such as equation 2.3. The mathematical representation of the auditory filter shape as a function is discussed next.

### 2.2.2 Auditory filter mathematical representation

To be successful in the goal of reconstructing the auditory filters, an appropriate mathematical representation has to be used fulfilling different criteria. The criteria may be empirical, practical or theoretical. In general, a simple mathematical expression is desired, with a small number of free parameters.

There has been agreement between different experimenters [20],[3], [24], to model the auditory filters in a functional form called the *rounded exponential filter* or *roex* model. Although there are other newer representations, such as time domain *gammatone* or *gammachirp* filters [7], focus is put on the frequency domain and such representations are considered out of the scope of the project.

The main reasons for choosing the *roex* functional form and properties of this type of filter model are presented next.

#### 2.2.2.1 The rounded-exponential filter model

From the results of different studies obtaining thresholds as a function of separation of notched noise bands or two-tone maskers, an approximately linear relation between threshold and masker separation can be inferred. Figure 2.5 depicts these findings, where the results have been corrected compensating for different noise levels and maskers used [20].

From figure 2.5 it can be seen that the logarithm of the signal power at threshold (given in dB) is



Figure 2.5: Threshold as a function of masker separation for different studies using symmetric notched noise and two-tone maskers, [20]

approximately a linear function of masker separation, indicating a good first approximation of the filter shape could be a *negative exponential function* [20].

Besides, the fact that the threshold functions for two-tone maskers and notched noise maskers look similar, may provide additional support to the negative exponential function. This is because the two-tone maskers and notched noise maskers will indicate how the filter function and area under the function varies, respectively. Given that the area represents the integral of the function and exponential functions have the same exponential function as integral (and derivative), this is strong support towards the exponential shape.

One practical advantage of an exponential function as the auditory filter shape descriptor would be the easiness in the prediction of threshold. The threshold can be obtained by integrating the filter function, and an exponential function would greatly simplify the calculations.

Another argument towards an exponential function is theoretical. The power spectrum model of masking ignores the detection mechanism that follows the filter [20]. If the detection mechanism operates on stimulus energy, it has been shown that the filter shape would not be the derivative of the threshold curve, but rather the square root of the threshold curve times the derivative of the threshold curve [19]. Therefore, given that the threshold and its derivative are approximately exponential, the power spectrum model of masking approximates the energy detection model in most cases [20].

Notice that the data deviates consistently at both ends of the curve from the linear relationship between threshold level and masker separation (see figure 2.5). The deviations suggest the filter may have a flatter top, instead of just being composed of a pair of exponentials set back-to-back [20].

A filter with a flatter top would benefit more from off frequency listening, making the auditory filter a more reasonable filter. When shifting, such a filter would attenuate less the signal and re-

duce the amount of noise (related to the filter skirts) to a greater extent, due to the higher slope present at the skirts.

Therefore, an exponential function with a rounded top,*a rounded exponential function (roex)*, seems appropriate for the description of the auditory filter.

A two parameter function with these characteristics, a roex(p,r) filter, has been described as [21]:

$$W(g) = (1-r)(1+pg)e^{-pg} + r$$
(2.5)

where  $g = |f - f_c|/f_0$  is the normalized distance from the center of the filter (*f* is the frequency at one edge of the noise band,  $f_c$  is the filter center frequency and  $f_0$  is the tone frequency), *p* is the exponential parameter and *r* is a constant.

The exponential parameter p will control the slope of the filter skirts, i.e a low value will be related to a broader filter and viceversa. The value of p may be allowed to differ for each side of the filter ( $p_l$  and  $p_u$  for lower and upper side, respectively), in order to reflect any filter asymmetry.

The parameter *r* flattens the filter at frequencies remote from the center frequency, thereby placing a dynamic range limitation on the filter. The value of *r* is normally assumed to be the same for both sides of the filter [21], [3]. The equivalent rectangular bandwidth of this filter will be approximately equal to  $2f_c/p_u + 2f_c/p_l$  [3].

Figure 2.6 shows an example symmetrical *roex* filter where the effect of varying the values of p and r can be observed. Figure 2.7 shows an example asymmetrical *roex* filter where p has been allowed to vary for each side of the filter.



Figure 2.6: Example of symmetric *roex* filter where the parameters p and r have been varied

The next section introduces the *roex* filter definition in the general masking equation, allowing to see how the threshold data will be related to the filter parameters.



Figure 2.7: Example of asymmetric *roex* filter for parameters  $p_l = 16$ ,  $p_r = 7$  and r = 0.15 for both sides

#### 2.2.3 Auditory filter parameter determination

The following derivations will assume the auditory filter can be fairly represented by a rounded exponential function and that certain degree of asymmetry in the filter will be present.

Introducing the normalized frequency distance, g, into the general masking model (see eq. 2.3), leads to [21]:

$$P_{s}(\Delta f/f_{0}) = KN_{0}f_{0}\int_{-\infty}^{-\Delta f/f_{0}} W(g)dg + KN_{0}f_{0}\int_{\Delta f/f_{0}}^{\infty} W(g)dg$$
(2.6)

For asymmetric notched noises the frequency distance,  $\Delta f$ , will take different values,  $\Delta f_l$  and  $\Delta f_u$ , for the lower and upper side of the filter, respectively. The indefinite integral of the *roex*(*p*,*r*) filter is [21]:

$$\int W(g) = -(1-r)p^{-1}(2+pg)e^{-pg} + rg$$
(2.7)

Because the tail of the *roex* filter ends up being flat, the limits of integration have to be bounded. A limit of normalized frequency distance of g = 0.8 has been suggested in [21], based on their fitting procedure being satisfactory enough and insensitive beyond this point. This limit might be different for low frequencies though.

Introducing the filter integral into the masking eq. 2.6 and using the recommended integration limit, will give:

$$P_{s}(\Delta f/f_{0}) = KN_{0}f_{0}([-(1-r)p_{l}^{-1}(2+p_{l}g)e^{-p_{l}g}+rg]|_{-0.8}^{-\Delta f/f_{0}}) + KN_{0}f_{0}([-(1-r)p_{u}^{-1}(2+p_{u}g)e^{-p_{u}g}+rg]|_{\Delta f/f_{0}}^{0.8}),$$
(2.8)

where the filter has been allowed to be asymmetric by introducing different exponential parameters,  $p_l$  and  $p_u$ , for the lower and upper sides of the filter respectively. For asymmetric notched noises, the value of  $\Delta f$  and therefore g, will be different for each side of the integral.

#### 2.2.3.1 Fitting the threshold data to the *roex*(*p*,*r*) model

In order to obtain the filter parameters,  $p_l$ ,  $p_u$  and r, that best fit the data, a minimization procedure can be used to fit these parameters to the threshold data, so that deviations are minimized [3].

For each center frequency the notched noise method is used to define one filter. The equation used to fit the data is an equation of the form of eq. 2.8 [21]. The limits of integration have to be set according to the different notched noise configurations. The configurations will produce different separations between the noise bands, in such a way that a range of thresholds can be measured and compared to the fitted data.

Starting values of  $p_l$ ,  $p_u$  and r can be assumed and a minimization procedure, such as the least-squares method, can be used to minimize the mean square deviation between the data and fitted values [3].

It is recommended to fit the threshold data measured in decibels to 10 times the logarithm of an equation such as eq. 2.8 [12].

The value of the constant K, representing the efficiency of the detection process, will depend on the subject, type of masker and center frequency [3]. When expressed in decibels, K is the additive constant that adjusts the predicted thresholds so that their average coincides with the average of the threshold data.

To summarize, the fitting can be achieved by [12]:

- Calculating a threshold curve for particular values of  $p_l$ ,  $p_u$  and r (for a given center frequency, the curve is obtained considering all masker separations).
- Calculating the mean squared deviation between the data points and the predicted threshold function.
- Varying  $p_l$ ,  $p_u$  and r so as to minimize the mean square deviation.

#### 2.2.3.2 Accounting for the auditory filter assumptions in the fitting procedure

Different studies for the critical band and auditory filter shape determination have made a number of simplifying assumptions. For example in [18] and [24] the filters were assumed to be symmetrical (even at high levels in [24]) and the filter position was assumed to be centered at the signal frequency. In [15] and [20] symmetry was not assumed and the filter position was assumed to be at "max(S/N)", but the filter bandwidth was kept fixed as the filters shifted.

In [3], the fitting procedure was modified to account for most auditory filter assumptions, avoiding simplifications and allowing more accurate derivations. The following points will discuss how to account for the auditory filter assumptions (outlined in section 2.2.1) in the fitting procedure.

Allowing for the determination of filter asymmetry: In [3] it is strongly recommended against deriving filter asymmetry from symmetric notched noise data. Therefore, to obtain the filter shapes the data will be fitted to the masking model with an equation such as eq.2.8. Given that the notch
will be asymmetrically positioned, the value of the normalized frequency, g, will be different for the lower and upper part of the filter.

Allowing for shifts in the auditory filter position: When fitting the data, the "max(S/N)" auditory filter centering assumption will be considered. In order to understand how this position can be estimated, an illustration of an hypothetical shift is shown in figure 2.8.



Figure 2.8: Illustration of auditory filter shift to maximize the S/N at its output

If the filter shifts a distance  $\Delta c$  off frequency, the signal will be attenuated by a factor  $W(-\Delta c)$ , and the edges of the noise bands will be  $g_l + \Delta c$  below and  $g_r + \Delta c$  above, relative to the filter center frequency. As the upper noise band is positioned further away from the tone (of frequency  $f_0$ ), the auditory filter shifts off frequency to the right side of the tone. For this conditions, the threshold equation will become [20]:

$$P_s = \frac{KN_0f_0 \cdot I_L(g_L + \Delta c) + KN_0f_0 \cdot I_U(g_U - \Delta c)}{W(-\Delta c)},$$
(2.9)

where the left and right integrals (see equation 2.8) have been expressed as  $I_L$  and  $I_R$ , respectively. In order to find the max(S/N) for a given notched noise configuration, equation 2.9 can be minimized respect to  $\Delta c$  [20]. For each particular notched noise configuration, the value of  $\Delta c$  for which equation 2.9 is minimum will be then the max(S/N) position where the auditory filter can be assumed to be centered.

**Considering changes in bandwidth with filter shifting:** As mentioned in section 2.2.1, there have been found no results yet that can lead to an estimation of the changes in bandwidth as the filter shifts off frequency at very low frequencies. Therefore, fitting procedure will in principle not consider this changes. Nevertheless, an estimation of the changes in auditory filter bandwidth will be considered once some results in between each center frequency can be compared.

# 2.3 Measurement considerations

The main measurements required in this work consisted of a number of *threshold of hearing for notched noise masked tones* determinations. The design of these measurements is a critical stage

in this project. This section explains what has been taken into account for the measurements and gives an overall overview of the system designed to obtain the auditory filters. It includes specifications about the chosen stimuli, sound field and equipment.

In principle, part of the measurement system can be considered as a *black box* where the tonenotched noise stimuli is the input and the output are the individual thresholds obtained for each configuration. In the diagram shown in figure 2.9, what is actually happening inside that *box* can be found.

The chosen method to find the auditory filters was the notched noise method and it has already been carefully detailed in the previous chapter. Therefore, the stimuli provided to the subjects had to be single tones with a notched white noise around them for different frequencies and configurations. These original stimuli had to be specially designed for one reason: it was desired to reach the inner ear of each subjec with approximately the same non-individual notched noise stimuli. It is supposed that the auditory filters have their origin from at least the inner ear [3]. Therefore, individual compensations had to be applied to the original notched noise stimuli. Thus, the measurement had to account for:

- 1. The response of the sound reproduction system and the room. This can be done by measuring this response and applying its inverse to the original stimuli. In appendix C all measurements and considerations regarding this issue can be found.
- The response to the outer and middle ear. It has been proposed that the equal loudness contours mostly represent the inverse of this response so they can be applied to shape the stimuli [3]. Then, what finally arrives to the cochlea will be more likely the desired notched noise stimuli. To individually design the stimuli, measurements of the threshold of hearing and equal loudness contours (*ELC*) were performed. These listening experiments are reported in the next chapter.

Assuming that the right signals are arriving to the subjects' cochlea, performing a threshold of hearing measurement for tones in presence of a white notched noise masker would provide threshold values that would allow to find parameters that characterise the auditory filters. With these parameters, the reconstruction of the auditory filter is possible.

Below, a further explanation about the just described process is given.

#### 2.3.1 Selection of sound reproduction system for controlled exposure conditions

In order to obtain valid results, certain criteria were established on the exposure conditions. The first criteria was to achieve a controlled sound field in the overall bandwidth of the signals to reproduce. The overall bandwidth of the notched noise signals consisted of all frequencies up to 275 Hz. The highest demand on the required sound pressure levels was expected at the very low frequencies, with estimated levels between 120 dB and 130 dB at 8 Hz. It was also necessary to maintain the same exposure conditions throughout the whole experiment.

The second criteria, considered less restrictive, is more directly related to human sound perception. Normally we hear with two ears, this is the most natural manner in which we are exposed



Figure 2.9: Diagram of the process involved in finding the parameters to derive the auditory filters

to sounds. However, most of the literature found on critical band measurements describe setups with headphones, (where only one cup is used, see i.e [24], [15]). It is not known to what extent our ability to discriminate sounds benefits from binaural hearing (in terms of auditory filtering). Neither it is an objective to determine which are the benefits of binaural hearing in this project, but it is not desired to disregard them if any. Therefore, if it was possible and criteria 1 was fulfilled, a sound reproduction system that allowed to achieve exposure conditions as "real" as possible would be preferred.

Based on the available equipment at the Acoustics Laboratory of Aalborg university, two options were considered: Headphone reproduction or the use of a specially designed room for lowfrequency sound reproduction. The two options were evaluated in terms of the established criteria. This is described in what is following.

#### 2.3.1.1 Evaluation of headphones

Headphones are often used in psychoacoustic experiments because of practical advantages over other reproduction systems. They do not require special setups and are easy to carry. Moreover they avoid having any influence of the room on the measurements.

For the evaluation of this reproduction alternative the response of different headphones was examined. Responses were found in the Phd. Thesis [4], page 20. Here the responses of different headphones on 40 subjects are plotted together from 100 Hz. The responses vary considerably between headphone types and none of them can be considered flat even in low frequencies. Besides, the response of the headphone varies depending on the subject, the variation is observed in the overall frequency range and can reach up to approximately 10 dB around 100 Hz. Therefore, to achieve a flat response, individual calibration would be required. This involves performing a measurement of the headphone response and designing an inverse filter for each subject. Moreover the fitting of the headphones on the subject ear is not going to be exactly the same every time the subjects wears them.

#### 2.3.1.2 Evaluation of low frequency room system

The advantages of doing the measurements in a room are basically two: avoiding the headphone fitting problem and achieve a better emulation of the normal binaural hearing conditions. The problems would come from difficulties in having a controlled sound field at such low frequencies. Achieving a flat or at least a compensable response in the range of interest and being able to provide high enough levels is not an easy task.

The low frequency room at Aalborg University offered an optimal starting point to possibly achieve the desired experimental conditions. The mentioned advantages as well as the fact that binaural exposure has been less explored in terms of auditory filtering encouraged us to choose this option.

However, the previous set-up of the room was not completely suitable to perform the required experiments. Therefore, some modifications had to be made. A full description of the room and these modifications can be found in Appendix C.

# 2.3.2 Accounting for the non-flat overall response

The threshold equations derived from the power spectrum model of masking have been simplified in the sense that the noise spectrum ( $N_0$ ) has been considered as a constant and has been taken out of the integral (see i.e equation 2.6). This, of course, assumes firstly that any spectrum coloration of the noise signals due to the non-flat response of the sound reproduction system has been taken into account (or that there is no coloration at all). Secondly, it assumes that the unequal efficiency of the sound transmission through the outer and middle ear depending on frequency has also been taken into account [3].

In fact, the response of the chain *reproduction system-room-outer and middle ear* is not flat at all. In the following sections, how this response was compensated is presented.

#### **2.3.2.1** The frequency response of the room and the reproduction system

If the frequency response of the sound reproduction system is flat, to solve the threshold equation the noise spectrum can be in principle taken out of the integral (see i.e equation 2.6) and an analytical integral can be evaluated. Even though, this is in most of the cases unfortunately not true, it has been the basis of earlier methods in the estimation of the auditory filter [18],[21].

One way this problem has been addressed, is to solve the threshold equation numerically, applying a correction for the frequency response of the sound reproduction system used (usually headphones) [15], [24],[3]. The effect of applying or not the correction can affect considerably the filter shape, and has been found to be larger at low center frequencies [3]. This method of adjusting the noise shape to match the sound reproduction system response is only done after auditory filtering, in order to fulfill the power spectrum model equation.

In [4] the responses of different headphones on many subjects have been plotted together, indicating the variability in the responses and suggesting that in order to make any corrections or to achieve a flat response when using headphones, individual equalization should be considered.

However, most of the studies where the headphone correction has been applied do not state clearly that individual responses have been measured or that individual equalization has been applied, leading to believe that common correction curves have been used.

Individual equalization seems to be the most proper alternative if headphones are used. However, since it requires to measure the response on each subject, other alternatives are to be evaluated for a more convenient sound reproduction.

Another issue is at which point the correction for the sound reproduction system is applied. It seems reasonable to apply the correction before auditory filtering, allowing to achieve in principle a flat response of the noise, and therefore solve the power spectrum model equation with an analytical integral. This option was also preferred because in this manner the level reached at the cochlea is supposedly the one that has previously been determined (as the compensating filter and the response of the room cancel each other). In appendix C, the design of this compensating filter applied to all the stimuli is detailed. The responses of the room are depicted as well. Appendix F describes how the compensating filter was obtained using MATLAB.

The following section describes how to compensate for the next element in the chain between the original notched noise stimuli and the cochlea.

#### 2.3.2.2 The outer and middle ear sound attenuation

Another consideration is how to deal with the frequency dependent attenuation of the outer and middle ear.

The basis of the auditory filters is presumably in the cochlea and/or in some higher stage of the auditory system [3]. Under this terms, the frequency dependent attenuation of the outer and middle ear can be considered as a fixed filter previous to auditory filtering. Because of this filtering, the notched noise signals will be modified and they will no longer be flat when reaching the cochlea.

To account for this fixed attenuation, previous studies have applied a similar procedure as when accounting for the non-flat frequency response of the sound reproduction system. This consists of estimating the attenuation of this previous filter and numerically evaluate the integral according to this estimation.

In order to estimate and account for the fixed attenuation of the outer and middle ear some researchers have used threshold or equal loudness correction curves based on particular assumptions [3], [22], [15], [24].

Threshold corrections curves are based on the assumption that the transducers in the cochlea in young normally hearing subjects are equally sensitive at all audible frequencies and the variation of absolute threshold with frequency reflects the frequency dependent attenuation of the outer and middle ear [23].

However, part of the variation in absolute threshold in the low frequency region (below 1 KHz) may arise in other ways [3]. Absolute threshold appears to change more rapidly with frequency than what would be predicted from the transfer function of the outer and middle ear [32]. A possible explanation for this is that internal noise in the cochlea is greater at lower frequencies than at higher frequencies [16]. This would rise absolute thresholds at low frequencies, but would have little effect on the perception of low frequencies at high levels [3]. This could explain why equal loudness contours at high levels are flatter than threshold curves, as they would be almost unaffected by the internal noise.

Based on this considerations, it has been assumed that equal loudness curves at sufficient levels would reflect the transfer characteristics of the outer and middle ear [15], [3]. Therefore, ELC corrections (usually the 100 phon curve) have been used in the fitting procedure to account for this fixed attenuation [3], [15], [22].

The threshold or ELC corrections, however, have been applied following standard curves and individual corrections have not been found in the cited literature. Because of individual variability, assuming the described hypothesis hold, the corrections can only be considered as rough approximations of the individual outer and middle ear transfer characteristics.

For this project it was therefore considered that individual ELC corrections would more accurately describe the fixed attenuation of the outer and middle ear of subjects and, therefore, allow

more precise estimations of the auditory filter characteristics.

Regarding at which point this correction had to be made, again it seemed more appropriate to apply it before providing the subjects with the stimuli, i.e. it should be included in the stimuli. This is considered more precise, taking into account that auditory filters are level dependent, as they seem to be in middle and high frequencies [3]. Then, if the correction is made in the stimuli, the level arriving the cochlea would be the desired one. The curves taken as references for the individual compensation are shown in Appendix E. The points shown as blue circles are the absolute thresholds of hearing obtained as described in section 3.2. Appendix F describes how these corrections were applied to the original stimuli using MATLAB. After this step, a set of individualized simuli was obtained for the measurements.

# 2.4 The notched noise stimuli

This section describes the notched noise stimuli proposed configuration. The center frequencies to investigate and main characteristics of the notched noise stimuli are given here. However, the final properties of the stimuli were set after performing pilot psychoacoustic tests and corresponding adjustments were done. Figures shown in this section show symmetric auditory filters just for illustration purposes, however no symmetric filter assumption is made.

#### 2.4.1 Choosing center frequencies to investigate

This project is low frequency-oriented, therefore the emphasis in the critical bandwidth measurements is at the very low frequencies, the lower end approaching the infrasonic area. Few studies focus on the critical band determination at very low frequencies, and what is called the "low frequency region" usually spans from 1 KHz down to 100 Hz [15], [24], [22]. Measurement results below 100 Hz have not been found in the examined literature.

In order to provide results in this uninspected area (below 100 Hz), and at the same time, be able to compare results with previous studies, it is proposed to measure the critical band in a region that spans between 32 Hz and 125 Hz. The proposed center frequencies for the critical band measurement and filter shape determination are: 32 Hz, 50 Hz, 80 Hz and 125 Hz. These center frequencies were chosen so as to allow a sufficient separation between frequencies, with relatively equal distances between them, and at the same time span the overall region of interest.

# 2.4.2 Determining the width of the notch bands

Another important feature in the notched noise method configuration is the width of the noise bands. The considerations taken when determining an appropriate width of the notched noise bands are described in the following.

One consideration is based on assumptions regarding the auditory filter bandwidth. In the notched noise method, the noise bands are thought as always being covering the filter up to the end of the filter skirts. Because of this, it is important to ensure that at extreme cases, no part of the auditory filter would lay outside the region of coverage of the noise bands. This is illustrated in figure 2.10.



Figure 2.10: Illustration of improper noise bandwidth selection

Consider the two noise bands of equal bandwidth  $B_N$ . The worse case will be when they are together symmetrically over the signal frequency ( $\Delta f = 0$ ). As figure 2.10 depicts, if the noise bands are not wide enough, part of the filter skirts will lay outside the noise bands. This will lead to a wrong determination of the real underlaying relationship between threshold and amount of noise power leaking through the filter.

Figure 2.11 illustrates how an appropriate noise bandwidth would cover the entire filter bandwidth in the worse case configuration. In order to choose an appropriate width for the noise bands



Figure 2.11: Illustration of proper noise bandwidth selection

it is therefore necessary to make assumptions about the auditory filter bandwidth. Assumptions can be based on previous results at low frequencies [15], [24], [22]. Although there is some degree of variation in the results, the a approximate values seem to be similar. For the lowest center frequencies where results are found in the literature, 100 Hz and 200 Hz, the *ERB* are approximately equal to 40 Hz and 50 Hz, respectively. This means that at this low center frequencies, the *ERB* is about 25% to 40% wide, relative to the center frequency.

The noise bandwidths used in most experiments may also provide a clue about what is a proper value for the noise bandwidth. The bandwidth of each of the noise bands used by different authors was found to vary from  $0.4f_0$  to  $0.8f_0$ , where  $f_0$  is the tone center frequency. Based on the filter width consideration and what previous studies have done, each of the noise bandwidths was cho-

sen to be equal to  $0.4f_0$ . Considering that at low frequencies the *ERB* can be around  $40\% f_0$  wide, in the worse case configuration the two noise bands would be  $80\% f_0$  wide placed symmetrically over the signal (i.e see figure 2.11), doubling the estimated *ERB* width. The noise bands can be considered wide enough, assuming the auditory filters will be relatively sharp [15].

Choosing the bandwidth of each noise band equal to  $0.4f_0$  will determine a maximum normalized frequency separation of 0.6 in the low frequency side, where the noise bandwidth can be kept the same. Further separations would need a reduction in the bandwidth of the lower side noise band. As a starting point, 0.6 will be the maximum normalized frequency separation considered, and the bandwidth of each noise band will be kept fixed for every normalized separation.

## 2.4.3 Notched noise stimuli proposed configuration

The proposed notched noise stimuli configuration to be used in the psychoacoustic experiment is described here. From what was discussed in the previous two sections, the required center frequencies and appropriate noise bandwidth were determined. The center frequencies ( $f_0$ ) were 32 Hz, 50 Hz, 80 Hz and 125 Hz and each noise bandwidth was chosen to be  $0.4f_0$  wide. Because filter symmetry was not assumed, the notch will be placed asymmetrically respect to the center frequency  $f_0$ .

For every center frequency, the pairs of values of normalized deviations ( $\Delta$ ) for the lower and upper noise bands were: 0.0 and 0.2, 0.1 and 0.3, 0.2 and 0.4, 0.3 and 0.5, 0.4 and 0.6, 0.2 and 0.0, 0.3 and 0.1, 0.4 and 0.2, 0.5 and 0.3, 0.6 and 0.4. The normalized deviations were chosen in such a way that for each deviation there was going to be a "mirror image" one respect to the signal frequency.

The minimum spacing between both of the noise bands would be  $0.2f_0$  and the maximum  $f_0$ . The deviation steps were chosen to be  $0.1f_0$ . There was always a tradeoff between the deviation resolution and measurement time. The chosen resolution had lead to satisfactory results in other studies [24], [15],[22].

An illustration of the notched noise configuration for a given center frequency  $f_0$  is shown in figure 2.12.

Then, as a starting point, for each center frequency there were 10 notched noise stimuli, each of them requiring a threshold determination.

The notched noise stimuli were generated using using MATLAB (Matlab Release 13, V. 6.5) [29]. To obtain the notched noise stimuli to be presented to the subjects in the listening experiment 3 steps were necessary. The first step was to design the filters that shape the noise stimuli according to the design shown in figure 2.12. This filters are called here the "notched noise filters". The second step was to shape the notched noise stimuli obtained after step 1 according to the ELC curves that were measured for each subject. The third and final step was to apply the hybrid mode compensation filter to the resulting signals obtained after step 2. A detailed description of each step is given in Appendix F.



Figure 2.12: Notched noise configuration for a particular center frequency  $f_0$ 

# **2.5** Deriving the auditory filters

The masked threshold data is the basic information needed to reconstruct the auditory filters. However, the process of deriving the actual filter shapes is not direct, but requires many calculations. Moreover, considerations have to be done in order to adjust appropriately the algorithms. Some of these aspects have already been discussed in section 2.2 in general terms. In the following sections, some practical aspects of the fitting procedure related to the proposed notched noise configurations are given. The use of ideal examples helped to define aspects that could be used in advantage for the threshold measurements. These examples also helped to determine the procedure to be applied with the actual data. The discussion will be centered on three phenomena:

- The fitting performance when using asymmetrical configurations, either in one, the other, or both sides of the notched noise
- The fitting procedure to allow for asymmetry and off frequency shifted filters.
- The interaction of these last two effects.

#### **2.5.1** The fitting procedure and the notched noise configurations

In order to prepare for data analysis and see how the notched noise configurations can be used for obtaining the filter shapes, some example threshold data was used for applying the fitting procedure before experimental data was available. Symmetrical and asymmetrical *roex* filters were simulated to generate ideal threshold data.

At first, the fitting was applied assuming that the auditory filter was positioned at the notch center. The fitting procedure converged satisfactorily. Secondly, the auditory filter was assumed to shift to the max(S/N) position. The fitting procedure was modified to account for the shifting and converged to the threshold curve again.

Programs were created in MATLAB to generate simulated data and apply the fitting procedure. These programs can be found in the attached CD. The *curve fitting tool* in MATLAB was used as integral part of the data analysis system [29]<sup>1</sup>.

Ideal threshold data was obtained from simulated *roex* filters, where the parameters of the filter (r, p and K) were arbitrarily set, and, in the case of asymmetrical filter, p was different for each side of the filter.

Equation 2.8 was applied and threshold points were determined at different normalized frequencies. Custom equations were created in the *curve fitting tool* as models to fit the data. The fact that the normalized distance in the asymmetric configurations follow the rule that  $g_r = g_l + 0.2$  or  $g_l = g_r + 0.2$  was exploited in the threshold equations, as the integration limits will be set for both noise bands at the same time in the normalized frequency axis.

The symmetric threshold equation was tested and parameters were accurately determined using

<sup>&</sup>lt;sup>1</sup>Given that the equations used to fit the data have the slope parameters of the filter on exponential terms, this is a non linear minimization problem. The curve fitting toolbox in MATLAB uses for this case a "trust region reflective Newton method" [29].

the curve fitting toolbox. For the asymmetric case, a *roex* filter with parameters  $p_l = 5$ ;  $p_r = 8$ ; r = 0.025, and K = 3 was simulated. Threshold was determined by integrating the filter function under the limits set by the proposed notched noise configurations (see section 2.4.3) and extended to further separations.

Both asymmetrical cases, one where the upper band is further away from the tone and the mirror image condition, were tested. Figure 2.13 shows the filter function and the thresholds achieved with the two asymmetrical cases. As can be seen, due to the underlying asymmetry of the auditory



Figure 2.13: Simulated filter response and calculated thresholds achieved in both asymmetrical conditions

filter function, thresholds achieved by placing the bands in exact mirror image conditions respect to the center of the notch are clearly different.

The curve fitting tool was used to estimate the filter parameters in each case, using either  $g_u = g_l + 0.2$  or  $g_l = g_u + 0.2$  to set the integration limits when either the upper or the lower band were 0.2 normalized frequencies away from the notch center.

Figure 2.14 shows the fitted curves and the data points for each case. As can be seen, in both cases the estimations are very accurate, and the parameters obtained are the same ones and correct in both cases.



Figure 2.14: Fitting example for threshold data obtained for both asymmetrical configurations

This means that, in principle, the filter shape can be estimated with only one asymmetrical configuration (even though threshold data will be different). This is possible because when estimating threshold in one asymmetrical configuration, information about both sides of the filter is being used. Therefore, if the integration limits are correctly set, threshold can be predicted without the use of the mirror image asymmetrical condition.

The fitting procedure shown here considers more threshold values than what can be measured in reasonable time. Nevertheless, the fitting was tested for the 5 discrete points (for each asymmetrical condition) that will be available from the measurements and the estimations worked equally well.

The results may indicate that it would be enough to measure threshold under just one set of asymmetrical configurations without the mirror image condition for each case. However, given that at that point only assumptions could be made about auditory filter properties in this frequency region, it was also thought that measuring both conditions could provide a more complete description of the auditory filter.

The fitting procedure so far had been tested with ideal threshold data and assuming the auditory filter is positioned at the notch center. However, this unfortunately does not give the best fit when dealing with real threshold data [3]. The following will discuss how the fitting can account for shifting in the auditory filter to the max(S/N) positions.

#### 2.5.1.1 Shifting the auditory filter to the max(S/N) position and modified fitting procedure

For the same given parameters ( $p_l = 5$ ;  $p_r = 8$ ; r = 0.025, and K = 3), the max(S/N) position for each asymmetrical configuration was determined. This was done by applying equation 2.9 and finding the shift  $\Delta c$  for which threshold was minimum.

Shifts were not allowed to exceed 0.2 normalized frequencies following recommendations given in [3]. It was assumed that given that the shifts are small, the filter shape will practically not vary when shifting [3].

For this case, threshold data will be the thresholds achieved after shifting the filter to the max(S/N) positions.

The fitting procedure in this case consisted of the following steps:

- 1. From starting values (experimenters' best guess) of  $p_l, p_r, r$  and K calculate the max(S/N) filter shifts.
- 2. Use the filter shifts to calculate threshold values using the parameters coming from point 1). Equation 2.9 is applied here.
- 3. Vary the filter parameters to minimize deviations between estimated threshold in point 2) and the "measured threshold". Measured threshold in this case consists of the threshold achieved after shifting the original filter to the max(S/N) positions.
- 4. Go back to point 1) and replace starting values of the parameters with the values achieved after point 3)

The routine is repeated until deviations are minimized. Using this procedure, the initial parameters of the unshifted filter were exactly recovered. This means that in principle, the max(S/N) shifting can also obtain the filter parameters by using either of the asymmetrical configurations.

**Summary** The fitting procedure with and without the max(S/N) assumption was working satisfactorily. The tests were performed on ideal threshold data, obtained from a determined filter shape, under the assumptions set by the power spectrum model of masking.

Given that information about both sides of the filter is used when achieving a threshold (either in symmetrical or asymmetrical configurations), the fitting can in principle find the filter parameters with only one set of asymmetrical configurations. The mirror image conditions, although yielding a different threshold, will lead to the same resulting filter parameters.

Because only assumptions about auditory filter properties could be made in this frequency region, it was thought that measuring both asymmetrical conditions will provide a more complete description of the auditory filter so this possibility was not discarded at this point.

However, the possibility of reducing measurement time by only measuring one asymmetrical configuration was going to be explored in pilot tests performed by the authors of this thesis. A symmetrical configuration is not needed, and would only be relevant if the auditory filter could safely be assumed to be symmetrical.

#### 2.5.2 Auditory filter asymmetry and filter shifts

The relationship between auditory filter asymmetry and shifting in the auditory filter to the max(S/N) position is analyzed here and some examples are presented.

Assuming an asymmetrical auditory filter ( $p_L = 5$ ;  $p_R = 8$ ; r = 0.025; K = 3), the max(S/N) positions were found by applying equation 2.9 and finding the shifts that yield minimum values of eq. 2.9. This was done considering every separation between noise bands. Figure 2.15 shows

a schematic illustration of auditory filter shifts for the cases where the lower and the upper band are closer to the tone ((a) and (b), respectively). The slopes of the filter on each side are drawn according to the assigned parameters. The filter shifts  $\Delta c$  to the right and  $\Delta c'$  to the left when  $g_u > g_l$ and  $g_l > g_u$ , respectively. Figure 2.16 shows the max(S/N) shifts obtained for the asymmetrical



Figure 2.15: Schematic representation of auditory filter shifting according to max(S/N) assumption

configurations (see section 2.4.3) in the cases where the upper noise band is further from the notch center (a) and when the lower band is further (b), as a function of normalized separation between bands. In order to appreciate more the effect of separation between bands, two extra separations were added to the proposed configurations. As can be seen, the shifts are greater for case (a). Furthermore, the shifts tend to increase with increasing separation between bands for case (a) and tend to decrease with separation for case (b).

This can be seen as a direct consequence of the auditory filter asymmetry [20]. Looking back at figure 2.15, the amount of area of the noise band closer to the tone removed from the filter as it shifts can be compared to how much the tone signal is attenuated. This can give an indication of the signal to noise ratio at the output of the auditory filter.

For case (a), given that the slope of the filter is less steep on the left side, as the filter shifts and attenuates the tone, a greater amount of noise is removed from the filter when compared to the mirror image case. This allows to benefit more from the shift for case (a) (assuming  $p_l < p_u$ ), and as a consequence, greater shifts are observed.

The variation in the amount of shift of the filter position as a function of separation between noise bands can also be explained as consequence of filter asymmetry. As the noise bands are further from the tone, the part of the filter skirt that leaks noise becomes less and less steep since the r parameter starts to take effect.



Figure 2.16: Filter shifts to maximum signal to noise positions as a function of separation between noise bands for the cases where the upper and lower band are 0.2 normalized frequencies further from the notch center

Since the slope of the left side of the filter is less steep than the right one, the effect of the r parameter is more pronounced on the left side than the right side. For case (a), the amount of area removed from the filter as it shifts and attenuates the tone is greater when compared to the same configuration, but when r has less effect. Thus, shifts will tend to increase with increasin separation between bands for this case. Besides, the difference in the amount of noise removed from the nearer side compared to the amount added on the other side will increase as well.

For the mirror image case, (b), as separation between bands increases, the r parameter also allows to take more noise out from the right part of the filter. However, the filter seems to benefit from shifting up to some limiting point. As the r parameter affects both sides, the amount of noise added to the left side of the filter (less steep side) starts to grow faster with frequency separation compared to the right side case. As a consequence, the filter no longer benefits from the shift, since any shift would immediately add more masking noise on the left side than what is removed from the right side of the auditory filter.

# **Chapter 3**

# **Listening experiments**

In the previous chapters, the most important considerations that should be taken into account in order to obtain the auditory filter shape and bandwidth have been discussed. The methods that could be applied and their respective pros and cons have also been discussed. Finally, the chosen method to be followed in this project work has been presented. Moreover, the whole process needed to obtain the auditory filters has been detailed.

In the fitting procedure, there are some parameters needed that require specific measurements in order to obtain the auditory filter shapes. The following chapter describes how these measurements were designed and performed. The design and set up for the measurement is explained. Also, some partial results are shown.

To provide the reader a more systematic description of the experiment, a practical division is made. This division is derived from the design of the experiment itself. It was planned to account for personal specificities in hearing, in order to extract general results. This meant providing specific stimuli so that what actually arrived to the inner ear was subjectively the same for each subject.

Therefore, the stimuli (the tone-notched noise stimuli already defined) were *custom-made*. This stimuli customization required previous measurements describing the hearing features of each subject. The required measurements comprised what is called here the *phase 1* of the experiment. The main test will be referred as *phase 2*.

Both measurement phases had many things in common, mainly the physical setup. Therefore, this setup will be explained first.

Thereafter, phase 1 will be described in detail, in terms of both procedure and results. The results were used as base for designing the stimuli used in the second phase.

Finally, the second phase of the experiment is described. The design of the customized stimuli has been already described. Then, the procedure and primary results are shown.

# 3.1 Measurement set-up

Many of the elements involved in the set-up were the same for both experiment phases. Therefore, in this section only a brief review of the set-up is given, as most of it has been already described in detail in previous sections.

# 3.1.1 Sound environment

The most important elements in the sound reproduction chain are the loudspeaker system and room environment. These two elements are gathered together achieving a high degree of control in the low frequency room at Aalborg University (see Appendix C for a detailed description). The use of this room allowed a controlled sound field in the frequencies of interest and to achieve a sufficient level for the reproduction of low frequencies. It also made possible to emulate the normal binaural hearing conditions. The selection of the room as sound exposure environment influenced the whole design of the measurements and sound stimuli. For this reason, further details about the room are presented in Appendix C. In this section, just a summary of what is described there is presented.

The room has inner dimensions of  $2.72m \ge 2.70 \le 2.40 \le 2$ 

The physical properties of the room and the specific signal processing allow different modes of reproduction. In the measurements, the *Hybrid* reproduction mode was used. This mode is a combination of the two other modes, the free field and the pressure field mode (see Appendix C). In the hybrid mode, the maximum possible sound pressure level that can be achieved in the room lies slightly below 130 dB at very low frequencies. At higher frequencies, the maximum reachable level is somewhat lower. However, this was not a constrain for the measurements, as only at the lowest frequencies the demand of level was higher.

The positioning of the subjects in the low frequency room was carefully selected (see Appendix C.3) and the frequency response of the room at that point was compensated in the stimuli.

# 3.1.2 Control and sound reproduction system

The low frequency room is controlled by a computer based interface in the control room next to it. The implementation uses one channel for each loudspeaker, yielding 40 different channels. On each channel a signal is passed through its respective amplifier and A/D converter. Each signal is convolved with a specific FIR filter and the resulting 40 filtered signals are amplified and fed to the loudspeakers. The FIR filters are designed to obtain the desired wave in case of the free field mode (or in the hybrid mode at the higher frequency region, above 30 Hz).

The implemented system works with a 48 KHz sampling frequency, but the control program manages files of 1 kHz. Therefore, up-sampling is required after filtering the 1 KHz input files.For more details see Appendix **??**,[26] and [25].

# 3.1.3 Stimuli and signal generation

All signals provided to the subjects were handled by the control PC placed in the control room next to the low frequency room. The signals might be generated in the PC or loaded from previously generated files that can be stored in the PC and used for playback in the room.

In phase 1 of the experiment, a threshold and an ELC determination were carried out. Thus, all signals were single tones played back at different levels.

In the second phase, the stimuli consisted on two components: a custom designed notched noise signal, and a tone placed in-between at a specific center frequency. Thus, the stimuli had to be carefully prepared. For this reason, the part of the stimuli that consisted of the notched noise, was previously designed using MATLAB (see Appendix F). The individual notched noise files were then transferred to the control computer, to be played back in the room. At the same time, while the notched noise signals were continuously played back in the room, a threshold determination measurement was performed. The threshold determination used practically the same procedure as the one used in phase 1.

# 3.1.4 Subjects

Primarily 14 subjects participated in the tests. This number comprised 6 male subjects and 8 female subjects, all aged between 23 and 28 years. A standard audiometry was performed on every subject and all showed normal hearing. Therefore, all 14 subjects were considered for the 1st phase of the experiment. However, due to the special conditions required for the second phase, 3 of the subjects had to be discarded from the final experiment.

# **3.2** Phase 1: Threshold of hearing and ELC

In order to account for variations in the perception of loudness along frequency, the stimuli was shaped respect to an equal loudness contour. As an approximation, standard contours could have been used. This is the approach that has been used in some previous works [15], [3]. However, because of individual variations, the application of individually acquired equal loudness contours was considered better so as to be able to achieve more reliable results.

Once this individual analysis was decided, an approximate phon level of the ELC determination had to be chosen. This level, the sound pressure level of the reference tone for the ELC determination, would be approximately the level at which the auditory filters were determined. Because of the level-dependance of the auditory filters, the final results could depend on this phon level. The basic criterion was to do the critical band determination at a level *well above hearing threshold*, in such a way that most components of the test stimuli could be clearly audible.

On the other hand, there are dynamic range limitations in the low frequency room sound reproduction system. The maximum possible levels in the room have to be considered before setting an approximate phon level, since the threshold of hearing at very low frequencies is at high levels. As described in Appendix C, the maximum reachable level in the room is limited to around 130 dB at 8 Hz. This is the minimum frequency in the ELC measurement and at which the highest level would be required.

Considering these two constraints (hearing threshold level and maximum reachable level), a compromised solution was applied: From inspection of the ELC standard [8], it was determined that staying close to the 60 phon curve would allow to be sufficiently above threshold in the required frequency range. To ensure that the levels of an ELC determination at around 60 phon fulfilled this criteria, it was decided to measure absolute threshold of hearing for each subject. Once this was done, an approximate 60 phon ELC determination was performed for each subject as well. In this way, individualized stimuli could be obtained. At the same time, the threshold measurement could be used to ensure that the level criteria was accomplished for each subject. Special attention was paid to the lowest frequencies, were this criterion could be more difficult to achieve because of the stretching between ELC's in this frequency range.

The analyzed frequency range was the range between 8 and 250 Hz, as that was the region were the notched noise stimuli contained practically all energy. In this manner, the stimuli could be weighted respect to the measured ELC. In this first phase, all stimuli used were single tones centered at 9 frequencies: 8, 16, 31.5, 50, 63, 80, 125, 200 and 250 Hz.

#### **3.2.1** Threshold of hearing determination: The ascending method

Many different methods can be applied to determine the threshold of hearing. One important feature to take into account when deciding for a method is its time consumption. This feature can set important limitations in a very time restricted project. In the case of the 2 phase experiment performed in this project, the total length of the experiment (including the 2 phases) was rather long. This lead to the use of a time efficient method if possible. The methods of limits are a family of methods that have demonstrated [10] to be very efficient in time. The method used in this work is based on one of these methods: the ascending method, which is particularly fast.

The method begins with the presentation of a tone at a level that is below the hearing threshold. The following presentations are at higher levels, increasing in 5 dB steps as long as the subject answers that he/she does not hear the sound. When the subject hears the sound, the first reversal in the answers is observed and a descending jump of -7.5 dB is produced in the stimulus. If the subject does not hear the sound the level is increased again in 5 dB steps until a new reversal in the answers. The level is then decreased by 7.5 dB again and it will be again successively increased and decreased until at least two reversals have been observed at the same level out of three ascents. The threshold of hearing is calculated as the mean SPL of the three ascents.

In order to provide a starting level that is below the threshold of hearing, a familiarization phase is required. This familiarization consists on descents in steps of 10 dB from a level well above the expected threshold value (around 30 dB more). After a reversal in the answers is observed, the level is still pulled down another 5 dB, looking for consistency in the previous answer. Then, the level is increased in 10 dB steps until a new turn over is reached. The starting level is set to 12.5 dB below the level of the last turnover.

## 3.2.2 Equal loudness contour determination: Maximum likelihood method

To determine points in an equal loudness contour, comparisons must be made with a reference tone. In this project, the reference point was a tone of 80 Hz at 78 dB SPL. The level was chosen to be as close as possible to the 60 phon contour and, at the same time, without risk of exceeding the maximum SPL for the test room. This reference SPL value (obtained after inspecting results from [11]) was chosen to possibly guarantee the already mentioned criterion of being well above threshold. The reference frequency was set after primary tests. The frequency of 80 Hz was chosen because it was found that it was easier to compare with the other frequencies under analysis. The 80 Hz tone seemed neither too low to compare with the 250 Hz tone, nor too high to compare with the lowest frequencies.

Then, comparisons were made between the loudness of the fixed 80 Hz reference tone at 78 dB SPL and the tones at the other 8 frequencies (at different levels). These frequencies were used to interpolate an equal loudness contour for every subject. The task was to find the level of the variable tone that makes it sound as loud as the reference tone. However, this is not that simple and in practical terms a range of uncertainty appears. Therefore, to obtain an equal loudness point, statistical analysis of the responses can be applied in the method.

The method used was based in the so-called *maximum likelihood estimate*. This is a mathematical term for a technique by which a set of parameters are adjusted to maximize a probability density function. The resulting parameter values are called the maximum likelihood estimates of the true values. In this case, the parameters are the mean  $\mu$  and standard deviation  $\sigma$  and their estimates are respectively  $\hat{\mu}$  and  $\hat{\sigma}$ .

The procedure to find each point in the ELC is as follows: A reference tone and variable tone are presented (in a random order, 2 seconds long each) separated by an interval of one second. The subjects indicates which of the tones was perceived as loudest and  $\hat{mu}$  and  $\hat{\sigma}$  are calculated following the maximum likelihood rule. After a new tone is presented at a pre calculated level, a new answer is given and parameters are re-calculated. This procedure is repeated until the estimation is considered to be accurate enough.

The maximum likelihood method requires, however two first input data to start the calculations: one test tone level that is louder than the reference tone level and one that is lower. Then, the experiment begins in a special way in which the first tone presented is an expected value to be achieved. In this case, the values were the ones shown in table 3.1. The next test tone level is set 10 dB lower or higher depending on whether the first level seemed louder or more quiet to the subject, respectively. Normally the second answer will be the opposite to the first one and the algorithm will be ready to go on. Otherwise, different levels are presented until two values are confirmed to be above and below the starting point.

The levels for the test tones are chosen in the uncertainty region so that they provide the maximum amount of information. They were  $\hat{\mu}$ ,  $\hat{\mu} - \hat{\sigma}$ ,  $\hat{\mu} - 2\hat{\sigma}$ ,  $\hat{\mu} + \hat{\sigma}$  and  $\hat{\mu} + 2\hat{\sigma}$ . These values were given the same probability, but already given levels were discarded. After the presentation of these five levels the algorithm finished, calculating the last set of parameters. The measurement was validated only if  $\sigma$  was smaller than 5 dB.

#### 3.2.3 Results

In this section a summary of the results obtained for both measurements, the threshold of hearing and equal loudness contour (approximately 60 phon) is presented. In Appendix E, the curves obtained for each subject are presented. In this section, only the most important aspects of the results are presented.

In global terms, the first phase of the listening experiment was a previous test needed to design and configure the main experiment. From this experiment, data was gathered for each subject so that phase two of the experiment became an individualized test. The main information that was needed to extract were the two results of phase 1:

- An equal loudness contour at approximately the loudness at which the stimuli will be presented. Individual results will be used to shape these stimuli, for each subject.
- A verification that for every subject the level of presentation of the stimuli (approximately 60 phon) is well above threshold of hearing.

Regarding the first of these goals, the obtained ELC are going to be used objectively, as they are considered as a tool for the main experiment of auditory filter determination. However, it should be pointed out that there are limitations in the conditions in which the ELC were obtained. The estimated time for whole experiment (the two phases) was considered too long and therefore the design of methods as short as possible was decided. In this case, this meant that no repetitions of the ELC measurement could be done for each subject. This limitation is present both, in ELC and threshold of hearing determination. However, considering the whole experiment, it is still thought that a good tradeoff was achieved in terms of total measurement time, accuracy and achieving all the proposed measurement goals.



Figure 3.1: Threshold of hearing and equal loudness contour for subject 12. The loudness range in low frequency seems wide enough for the experiment to be run well above threshold.

As to the second objective, unfortunately it was not possible to fulfill the condition of being clearly above threshold for every subject. Figure 3.1 shows one of the subjects for whom the condition

was accomplished the best. The difference between the threshold curve and the ELC was considered large enough to provide a safe working range. This was true for this subject and for most of the rest of the subjects (see Appendix E).

However, for 3 of the subjects a tight stretching in loudness range was observed at very low frequencies, limiting the working range. These were subject number 5, 8 and 9. The case of subject 8 is particularly extreme, and it was decided to discard this subject for the next experimental phase. His response can be seen in figure 3.1.

Two other subjects showed an unusual behavior that did not allow to use them as subjects for the second phase. In this case, the reason was the opposite. Their ELC was obtained and very high levels at low frequencies were observed. Thus, there could have been a substantial working range for the experiment. Unfortunately, for the lowest frequencies, the level required by the subjects exceeded the dynamic range of the room.

Therefore, at the end of phase one, 11 subjects were found appropriate for the second phase of the experiment. This amount comprises 5 male subjects (two of them were the experimenters) and 6 female, all aged between 23 and 28 years.



Figure 3.2: Threshold of hearing and equal loudness contour for subject 8. The extreme stretching observed for this subject in terms of loudness in low frequency led to discard this subject for the second phase of the experiment.

# **3.3** Phase 2: Threshold of hearing with a notched noise masker

The second phase of the listening experiment is the main test. It consists of a threshold determination of single tones in the presence of a masker. The masker consists of a notched noise around the frequency under analysis.

The experiment was carried out in the same room and with the same sound reproduction system that was used in the first phase. The chosen position in the chamber did not change either and the same subjects were used, except the 3 that had to be discarded.

#### CHAPTER 3. LISTENING EXPERIMENTS

The measurement was based on the notched noise method procedure, but with some particularities derived from the extreme frequency range. Because of this, special features of hearing were considered, such as the possible influence in the results of stretching in the loudness contours at low frequencies. For this reason, the stimuli were carefully individualized and therefore more complicated to derive.

However, apart from these particularities, the measurement will still be basically a threshold determination and usual procedures might be applied. Therefore, it was decided to take advantage of one of the already designed methods. Nevertheless, the first task was to find which of the possible methods was the most appropriate for our purposes.

In principle, accuracy was a requirement. While testing the fitting algorithms (that allow to reconstruct the filters from obtained threshold data), it was observed that errors or inaccuracies would imply a worse fitting when trying to incorporate the results to the *roex* model.

On the other hand, the chosen method had to be time efficient. From the starting notched noise proposed configurations (see section 2.4.3), 10 threshold values were needed in order to find the filter shape at one single frequency. Eleven subjects were going to be tested, in such a way that average results could be obtained. Besides, it was intended to test 4 center frequencies: three of them in the *unexplored* region (32, 50 and 80 Hz) and one more that would allow to make comparisons with results from other researchers (125 Hz). A measurement with individual subjects at different frequencies was also desired, in order to be able to observe individual changes along frequency as well. Considering this, for each subject 40 threshold determinations would be needed.

Moreover, repetitions in each threshold determination were needed. In the examined literature, it was found that in all cases repetitions of the threshold determination are done to ensure accuracy [15], [22], [24], [21]. As said before, accuracy was a primary requirement. Therefore, 2 threshold determinations for each subject were considered as minimum for the experiment. This meant that at least 80 threshold determinations would be needed for each subject. All this led to a previous analysis on time consumption of the methods.

#### 3.3.1 Chosen method

The most common methods for threshold determination that were found in the related literature are the 2 down-1 up or 3 down-1 up methods. The two alternative forced choice (2AFC) decision procedure was incorporated in these methods [24]. On the other hand, other traditional methods have proved to achieve good performance as well. Among them the, the ascending method is considered to be a fast method [10]. Moreover, this method was the one used in the first phase of the experiment and was already available. Therefore, two methods were considered for phase two of the listening experiment: the ascending method (used in phase one) and the implementation of a short version of the 2 down-1 up approach used in [24].

In Appendix D, a discussion on the time consumption of these two methods is provided. In that section, the reader can find more practical details about the methods. After this analysis, it was decided to choose the ascending method. Its description has already been given in section 3.2.1.

# **3.3.2** Grouping subjects

From the analysis given in Appendix D, it was determined that the original experiment plan was going to take too long, whichever method was selected. A tempting solution could be to reduce the number of repetitions for each frequency and subject. However, this could lead to very biased results as single significant inaccuracies in a threshold determination can deeply influence the final filter obtained. Therefore, it was preferred to reduce the number of center frequencies tested for each subjects and use a grouping strategy. This means using a group of subjects for certain center frequencies and other groups for other center frequencies, instead of using all subjects for all center frequencies.

In this way, accurate enough data could be obtained to derive the auditory filters with less time constrains. If the grouping is accordingly designed, it can still be possible to have a sufficient number of derived filters on each center frequency. This would allow to find average values and even to compare averaged results between different center frequencies.

## **3.3.3** Pilot tests and final notched noise configuration

Pilot tests were performed where the two authors of this thesis participated as subjects. The tests were performed in order to test the setup and see how preliminary results were looking. In this manner, adjustments in the experiment could be done before running the main experiment with the voluntary subjects. It is thought that the adjustments performed allowed a more efficient data recollection, considering the time constrains.

Threshold was measured 2 times for center frequencies of 32, 50, and 80 Hz for each of the two asymmetrical conditions and for each of the 2 experimenters. The normalized deviation from the tone frequency to the closest band was 0, 0.1, 0.2, 0.3, 0.4 and 0.5, the furthest band being +0.2 further from the tone.

Results are shown in figures 3.3, 3.4 and 3.5. The figures show the averaged threshold obtained from the two measurements as a function of masker separation, for the two asymmetrical cases. The threshold values were found to vary in some cases by more than 5 dB between repetitions. However, the most common difference was within 2 dB.

The fact that variations in threshold values were found to exceed in some cases 5 dB for the same conditions, is considered a limitation of the method. Ideally, with no time constrains, many repetitions could be done in the main experiment and the results that deviate by more than i.e 2 dB could be discarded. However, the main tests would become too long if more repetitions are added, and tiredness of subjects could affect the results. Considering the time limitations and voluntary nature of the subjects, two repetitions are considered for the main test.

It was planned, however, to perform another set of measurements by the 2 authors of this thesis including more repetitions in the threshold determination. The results will be analyzed for comparison purposes. In these results, the center frequency of 125 Hz will be included so as to be able to compare results with other studies as well.

The next sectionss describe the main observations that were done with the preliminary data. Adjustments done to the main threshold determination measurement are described as well.



Figure 3.3: Measured threshold as function of masker separation for a 32 Hz tone and the 2 asymmetric configurations, (a) Carlos, (b) David



Figure 3.4: Measured threshold as function of masker separation for a 50 Hz tone and the 2 asymmetric configurations, (a) Carlos, (b) David



Figure 3.5: Measured threshold as function of masker separation for a 80 Hz tone and the 2 asymmetric configurations, (a) Carlos, (b) David

#### 3.3.3.1 Overall variation in threshold

Looking back at figures 3.3, 3.4 and 3.5, threshold seems to generally decrease with masker separation. The drop in threshold level seems to be, in general, less accentuated for the first masker separations.

The overall variation in threshold with masker separation was found to increase with center frequency. For example the difference between the maximum and minimum threshold found for each center frequency increases from approximately 12 dB at 32 Hz up to 24 dB at 80 Hz center frequency. From the auditory filter simulations and threshold determination programs, it was found that a low overall variation in threshold can be associated with the two or at least one slope parameter being relatively small (i.e. *p* close to 4). On the other hand, higher overall variations in threshold are associated with larger *p* values, assuming  $r \ll 1$ .

#### 3.3.3.2 Threshold variation for small masker separations

Ideal threshold obtained from simulated auditory filters indicated that in no case threshold would increase with increasing masker separations, if assumptions made are correct. Nevertheless, for small masker separations and shallow filter skirts, threshold drop can easily be only within 2 dB or even less. Under these conditions, in principle, any raise in threshold level with masker separation should be considered to be a result of limitations in the threshold determination measurement.

Given that the changes in threshold level for the smallest masker separations were in general small (typically around 2 dB), for these cases any measurement inaccuracies would more easily produce an apparent increase in threshold level with masker separation. Therefore, to avoid to some extent the influence of measurement inaccuracies, it was decided that for the main experiment threshold should be measured starting from the second masker separations (where  $g_u + g_l = 0.4$ ).

#### 3.3.3.3 Measuring threshold using one or the other asymmetrical configuration

The results shown in figures 3.3, 3.4 and 3.5 show that there will be in general consistent differences when measuring threshold for one or the other asymmetrical configuration. For some cases, one configuration tends to give higher thresholds than the other and for other cases the opposite is true. Ideal threshold data indicates that when the lower auditory filter skirt is shallower than the upper one, threshold will be higher when the right side band is further from the tone compared to the mirror image case. The opposite is true if the upper filter skirt is shallower than the lower one.

Considering this, preliminary results indicate that the lower auditory filter skirt is shallower than the upper one at 50 Hz for subjects (a) and (b) and 80 Hz center frequency for subject (a). The opposite is the tendency for 80 Hz center frequency subject (b) and 32 Hz center frequency subjects (a) and (b).

When applying the fitting procedure at this point, it became clear that some refinement in the fitting method was needed when dealing with real data. Firstly, it is not yet clear how some contradicting data points will be treated, particularly when assuming that a partial grow in threshold with masker separation is considered to be a consequence of inaccuracies in the measurement method.

Secondly, appropriate boundaries in the possible values the filter parameters may assume are needed. Mathematically some fits can be very accurate, but the parameters sometimes reach values with no real physical significance. Therefore, before making any quantitative description of the filter behavior with this preliminary results, some adjustments in the fitting method are planned.

However, even though no detailed description of the auditory filter behavior wass available at this point, it was believed that the obtained preliminary threshold results indicated that the auditory filters are clearly asymmetric. Furthermore, it was thought that the degree of asymmetry could be consistently measured using either of the notched noise asymmetrical configurations.

#### 3.3.3.4 Final notched noise configuration

It was desired that in the main experiment threshold was going to be determined at least for two center frequencies for each subject. This would help to estimate the evolution of the auditory filter with frequency. At the same time, this brought up the constraint of subject time. Therefore, because one asymmetrical configuration was considered sufficient for determining the filter properties, it was planned to perform the threshold determination with only one asymmetrical configuration in the main test. This would provide enough data to estimate the filter shape properties, and at the same time, estimate individual variations of these properties with frequency within reasonable time.

Because focus was put on one asymmetrical configuration, it was decided to add an extra masker separation to the final notched noise configuration. It was chosen to add an even more separated configuration (where  $g_l + g_u = 1.4$ ), given that more separate configurations might provide more information about the *r* parameter, related to the dynamic range limitation of the auditory filter.

Thus, the final notched noise configuration for the main test consisted of one asymmetrical configuration where the normalized separation between the tone and the closest noise band was: 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6, the furthest band being +0.2 further from the tone. It was chosen to have

the upper band further from the tone ( $g_u = g_l + 0.2$ ) so as to avoid decreasing the bandwidth of the lower band as no components of the noise can fall below zero Hz.

# 3.3.4 Final method and grouping strategy

Before describing the final measurement plan, the results from the pilot test are summarized: It was decided that 6 notched noise configurations were appropriate for the threshold measurements and at least 2 repetitions were needed. It also seemed that a training session was required, as the task was found to be more difficult due to the characteristics of the stimuli. This would slightly increase the total time of the experiment. Moreover, the experiment seemed to produce tiredness, if not annoyance, therefore time per session should be reduced as much as possible.

The ascending method was therefore the chosen method. This method has been described in section 3.2.1. The particularity in this case is that the masking noise was being reproduced continuosly during the experiment. The noise level was individually shaped but always keeping the same level at 80 Hz. This reference level was 54 dB (the maximum not causing distortion for any subject in any configuration of the noise). In the pilot tests it was observed that an appropriate first presentation level was around 35 dB above absolute threshold for each frequency.

After setting these levels each threshold determination took less than 70 seconds on average. This means that around 420 seconds were required to measure one center frequency (6 notched noise configurations). If 2 repetitions and 2 center frequencies are considered per subject, the experiment would last itself around 28 minutes. Considering one break in between of 10 minutes, 10 minutes more for giving the instructions and 5 minutes for the training session, the test would take approximately 53 minutes in total, which seemed reasonable.

On the other hand, as only two frequencies were considered per subject, it was then necessary to perform the already mentioned grouping strategy. The next table shows how subjects were distributed among frequencies. Note that the last two subjects (the experimenters) performed the experiment for every center frequency considered, including 125 Hz. This frequency was no longer considered for the rest of the subjects. The aim of the project was to achieve results below 100 Hz, and the 125 Hz frequency was tested by the experimenters to be able to compare results with previous researches. Moreover, additional measurements were performed only by the experimenters in the same center frequencies. A total of 4 thresholds for each center frequency and notched noise configuration were measured. In this way, the influence of the number of repetitions on the results could be observed.

Subjects	31.5 Hz	50 Hz	80 Hz	125 Hz
Subject 1	×	×		—
Subject 2	×	×		—
Subject 3	×	×		—
Subject 4	×	—	×	—
Subject 5	×	—	×	—
Subject 6	×	—	×	—
Subject 7	—	×	×	—
Subject 8	—	×	×	—
Subject 9	—	×	×	—
Subject 10	×	×	×	×
Subject 11	×	×	×	×
Total number of subjects	8	8	8	2
Number of repeated subjects	5	5	5	2

Table 3.1: Grouping strategy to analyze average data in the four center frequencies.

# Chapter 4

# **Results and analysis**

In this chapter the masked threshold experimental results are given and the obtained threshold data is analyzed. The section is divided in three parts. In the first part the threshold results are shown and obtained data is discussed. The second part describes the filter shapes and ERB obtained for each subject after the fitting procedure. In the third part, the filter shape and ERB data is analyzed in terms of the grouping and statistical tests are performed on the grouped data. An expression that approximates the variation of the ERB in the frequency range 31.5-125 Hz is suggested as well.

# 4.1 Masked threshold results as a function of notched noise masker separation

The masked thresholds obtained in the listening experiment for each voluntary subject are shown in figures 4.1 to 4.5. The figures show threshold obtained for each of the 9 voluntary subjects, where each person was tested for 2 center frequencies and the measurement was repeated 2 times.



Figure 4.1: Measured threshold as function of masker separation for 31.5 Hz and 50 Hz tones, (left) Subject 1, (right) Subject 2

All figures show the averaged threshold obtained after performing the 2 repetitions of the measurement. Differences between threshold values obtained in the first and second repetition were in general small. However, in some cases differences exceeded 5 dB. For the cases where an apparent increase in threshold with masker separation was found, the lower value of threshold was kept



Figure 4.2: Measured threshold as function of masker separation, (left) Subject 3, 31.5 and 50 Hz tones, (right) Subject 4, 31.5 and 80 Hz tones



Figure 4.3: Measured threshold as function of masker separation for 31.5 Hz and 80 Hz tones, (left) Subject 5, (right) Subject 6



Figure 4.4: Measured threshold as function of masker separation for 50 Hz and 80 Hz tones, (left) Subject 7, (right) Subject 8

instead of the average. This was not necessary in the majority of cases.

Measurements were done with more repetitions to see if the estimation of threshold as a function of masker separation could be improved. In this case, for practical reasons, the two authors of this thesis participated as subjects (subjects 10 and 11). Figures 4.6 and 4.7 show threshold



Figure 4.5: Measured threshold as function of masker separation for 50 and 80 Hz tone, Subject 9

obtained after averaging 4 repetitions of the measurement. In these cases, all center frequencies were tested on each subject and the 125 Hz center frequency was included to be able to compare results with other studies.



Figure 4.6: Measured threshold as function of masker separation for subject 10, (left) 31.5 and 50 Hz tones, (right) 80 and 125 Hz tones



Figure 4.7: Measured threshold as function of masker separation for subject 11, (left) 31.5 and 50 Hz tones, (right) 80 and 125 Hz tones

As can be seen from figures 4.6 and 4.7, the general trends of the obtained curves with more averages look very much alike to the results from subjects 1 to 9 shown in figures 4.1 to 4.5.

Nevertheless, threshold curves shown in figures 4.6 and 4.7 seem more smooth in general when compared to the curves shown in figures 4.1 to 4.5.

The next sections describe the general observations performed on the obtained threshold data.

#### 4.1.1 Overall variation in threshold

Considering all subjects, the overall variation in threshold increased slightly with center frequency. The mean difference between the maximum and minimum threshold obtained for the center frequencies of 32, 50 and 80 Hz was 11 dB, 13 dB and 14 dB, respectively. A T-test indicated the differences were not significant.

Threshold values were, however, always higher for the lowest center frequency tested on each subject when compared to the higher one. For each subject, overall threshold levels of the 32 Hz tone where always higher than for the 50 Hz tone. The same happened when comparing threshold values for the 50 Hz tone to the ones obtained for the 80 Hz tone, individually. For the two cases where the 125 Hz center frequency was tested, overall threshold values were the lowest when compared with the values achieved by the same subject for the other 3 center frequencies. This indicates the detection efficiency decreases steadily with frequency.

#### 4.1.2 Shape of the threshold curves

The shape of the threshold curves varies depending on the subject and center frequency. Most of the curves, however, present some similar features. Threshold decreases with masker separation and in general the degree of decrease in threshold is not so accentuated in the first masker separations.

The main differences are, in general, at the end of the threshold curve. In the last masker separations, for some subjects and center frequencies, threshold seems to have reached a point where it flattens off or no longer decreases as strongly as in the previous separations. On the other hand, for other cases the threshold curve does not seem to have reached that point where it "flattens off". Compare for example threshold for Subject 9 at 32 Hz and 50 Hz (figure 4.5), for Subject 10 at 31.5 Hz and 80 Hz (figure 4.6) and subject 11 at 125 and 80 Hz (figure 4.7).

The ability to reach a point where threshold curves seemed to flatten off depended on the subject and center frequency. This point may provide direct information about the effective bandwidth of the auditory filters, since beyond this limit the masking noise bands do not seem to influence threshold levels anymore. Threshold levels where the curves flattened off where found to be still much higher than absolute threshold for the particular subjects. Therefore, it is thought that reaching this point in the threshold curves will provide useful information about the dynamic range parameter of the auditory filters (*r* parameter).

#### 4.1.3 Specific cases

Some specific results are discussed here. Focus is put on results that deviate somehow from the general cases or from the expected results.

For example the threshold values obtained for Subject 6 at 80 Hz deviate from the general trends observed for the rest of the subjects (see figure 4.3, (b)). Here threshold starts at a lower value and steadily increases to a maximum at a normalized separation of 0.8 and then abruptly decreases to a value similar to the starting point for the last masker separations.

When running the experiment, for this subject a session had to be repeated because inconsistent answers were observed. Instructions were repeated, indicating to answer when clearly distinguishing the tone from the noise. Since it is not clear the real nature of the values obtained for this subject at this frequency, and data deviates from the "normal" subjects, this particular measurement is discarded from further analysis.

Another somehow "special" case is observed in the threshold curves obtained from Subject 8, figure 4.4, (b). Here threshold does not seem to clearly decrease until reaching normalized masker separations greater than 1. The measurement will not be discarded in principle, but will be treated with care in the further analysis.

# 4.2 Auditory filter shapes and ERB

In this section the main results of the project work are given. All effort has been put in order to obtain the auditory filter shapes and ERBs in this, until now, uninspected frequency region. The section is divided in two parts. The first one details the fitting procedure applied to the data and provides some fitting examples. In the second part, the auditory filter shapes and ERB obtained for each subject are given. The detection efficiency is also discussed in this part of the section.

#### 4.2.1 Applying the fitting procedure to the obtained data

The fitting procedure has been described in detail in section . However, some modifications were needed when dealing with real threshold data. Before showing fitting examples and discussing how the fitting worked in general with the data, the final modifications performed to the fitting procedure are described.

#### 4.2.1.1 Final modifications to the fitting procedure

From what was found when trying to fit the real preliminary threshold data, the fitting procedure had to be modified to some extent. One modification was to restrict the possible values the parameters could reach. When applying the fitting without restricting the parameter values, sometimes values with no physical significance where reached. The p parameters, representing the filter slopes, where restricted to be positive from a lower boundary of 0.05 to an upper limit of 60. The lower limit of p represents practically a flat filter and the upper limit is much sharper that what has been found at much higher frequencies [12].

The value of the dynamic range parameter, r, was allowed to vary from a very small positive number  $(10^{-5})$  to 0.95. The lower limit represents almost no effect, meaning the threshold curve could provide practically no information about the parameter (threshold probably did not "flatten off"). The upper limit can not exceed 1 given that the maximum gain of the filter is set to 0 dB, so

a limit of 0.95 means the auditory filter reaches almost immediately its limiting dynamic range.

The *K* parameter, representing the detection efficiency, was allowed to vary from a small positive number  $(10^{-5})$  to 90. The values were considered adequate after fitting example threshold data. Given that the function of *K* is to shift the threshold curve either up or down to match the measured threshold curve, it can take in principle any positive values that allow it to freely shift the required amount of dBs for any particular case.

Another modification needed was to set a higher integration limit in the threshold equation applied to fit the data (see equation 2.8). The previous limit of 0.8 had been set because of a recommendation given in [21]. However, the value was recommended because in this particular study their fitting was insensitive beyond this point. Given that the integration limits of equation 2.8 are unbounded for the lower and upper side, in principle any high enough value could be used. In practical terms a high enough value has to be set where the fitting works equally well for the data set. Following this procedure, a value of 2 was found to be appropriate.

Because of practical reasons, the filter shifts to the maximum S/N position are not currently implemented in this modified fitting procedure. Given that shifts are considered to be in general small, they don't significantly affect the filter shape and therefore the filter parameters [3]. However, it was planned to calculate the filter shifts from the obtained parameters after the fitting, in order to study how the auditory filter may benefit from shifting in this low frequency region.

Fitting examples and a description of the fitting performance in general and for particular cases are given next.

#### 4.2.1.2 Examples of fitted threshold curves and fitting performance

Figures 4.8 to 4.10 show examples of fitted threshold curves. The goodness of the fit can be expressed in terms of the  $r^2$  value, which gives an indication of how accurate can be the threshold prediction. A value close to 0 indicates threshold data can not be predicted by the model and close 1 indicates the model achieves an accurate prediction of threshold.



Figure 4.8: Example of fitted threshold curves, (left) Subject 4, 31.5 Hz; (right) Subject 3, 50 Hz

In general the model fitted the threshold data with  $r^2$  values greater than 0.91. The only 2 cases where the model could not fit accurately the data are the 2 measurements performed on subject 8 (see figure 4.4, (b)). The obtained values were already discussed in section 4.1.3. Since the model


Figure 4.9: Example of fitted threshold curves, (left) Subject 9, 80 Hz; (right) Subject 6, 31.5 Hz Hz



Figure 4.10: Example of fitted threshold curves, (left) Subject 11, 31.5 Hz; (right) Subject 10, 125 Hz

cannot predict the threshold values and the threshold curves for this subject deviate from the "normal" curves, the 2 measurements performed on this subject are discarded from further analysis.

In general, it was found that the accuracy in threshold estimation improved to some extent for the cases of subject 10 and 11, where threshold was obtained from averaging more repetitions of the measurement. As described in section 4.1.1, threshold values obtained in this manner where slightly smoother than in the other cases, making it easier for the model to match the measured curve.

The next section shows the filter shapes, ERBs and detection efficiencies achieved for each subject after fitting the model to the threshold data.

#### 4.2.2 Obtained auditory filter shapes and ERBs

Figures 4.11 to 4.22 show the auditory filters obtained for each subject and associated center frequencies. In the figures, the p and r parameter values are given as well.



Figure 4.11: Auditory filters for Subject 1, (left) 31.5 Hz, (right) 50 Hz



Figure 4.12: Auditory filters for Subject 2, (left) 31.5 Hz, (right) 50 Hz



Figure 4.13: Auditory filters for Subject 3, (left) 31.5 Hz, (right) 50 Hz



Figure 4.14: Auditory filters for Subject 4, (left) 31.5 Hz, (right) 80 Hz



Figure 4.15: Auditory filters for Subject 5, (left) 31.5 Hz, (right) 80 Hz



Figure 4.16: Auditory filter for Subject 6, 31.5 Hz center frequency



Figure 4.17: Auditory filters for Subject 7, (left) 50 Hz, (right) 80 Hz



Figure 4.18: Auditory filters for Subject 9, (left) 50 Hz, (right) 80 Hz



Figure 4.19: Auditory filters for Subject 10, (left) 31.5 Hz, (right) 50 Hz



Figure 4.20: Auditory filters for Subject 10, (left) 80 Hz, (right) 125 Hz



Figure 4.21: Auditory filters for Subject 11, (left) 31.5 Hz, (right) 50 Hz



Figure 4.22: Auditory filters for Subject 11, (left) 80 Hz, (right) 125 Hz

The results shown in figures 4.11 to 4.22 indicate the filter shapes will strongly vary depending on subject and center frequency. However, some general features can be inferred. For example, even though the p values were allowed to vary in an extensive range, it turned out that in 17 out of 23 cases the filters ended up being practically symmetrical.

The dynamic range parameter, r, was well determined practically in all cases where the underlying threshold curve had reached the point where it flattens off. For the other cases, it achieved its minimum possible value ( $10^{-5}$ ), meaning that in the range the threshold curve was measured, the dynamic range limitation of the auditory filter probably had still no effect on the data. For any further measurements in this frequency range, it is therefore recommended to include even more separated configurations.

Another feature is how the filter slope parameters at 31.5 Hz defined in general sharper filters than at other center frequencies. At the same time, the dynamic range limitation of the filters at 31.5 Hz was better determined than at other frequencies (threshold flattened sooner), and the filters seem to have a lower dynamic range in general compared to the other center frequencies.

#### 4.2.2.1 Auditory filter shifts

Filter shifts were calculated with the obtained parameters. For the majority of cases, where the auditory filters are practically symmetrical, the shifts are rather small (less than 6% of center frequency). For the fewer asymmetrical cases, the shifts can easily vary from 0% to 20%, depending on the notch noise configuration, as was seen in section 2.5.2.

#### 4.2.2.2 ERB and detection efficiency

The ERB was calculated for every case from equation  $ERB = 2f_0\{(1-r)p_l^{-1}[1-(1+p_l)exp(-p_l)] + r\} + 2f_0\{(1-r)p_r^{-1}[1-(1+p_r)exp(-p_r)] + r\}$  instead of the approximation  $f_0\{2/p_r+2/p_l\}$  [3], given that for low *p* values the approximation deviates from the exact equation.

ERBs and detection efficiency (*K*) values obtained for every subject are shown in table 4.1. The detection efficiency values are given in dB as  $20log_{10}(K)$ . High detection efficiencies (small *K* values) will lead to small or negative dB values and the opposite is true for poorer detection efficiencies.

#### CHAPTER 4. RESULTS AND ANALYSIS

31.5	50	80	125
$26.0_{\{12.2\}}$	$58.1_{\{-3.5\}}$		
$13.6_{\{24.1\}}$	$42.6_{\{-0.2\}}$		
$13.9_{\{8.9\}}$	$55.0_{\{-4.5\}}$		
$12.6_{\{18.9\}}$		$43.1_{\{-14.9\}}$	
$37.1_{\{-0.4\}}$		$54.0_{\{-6.6\}}$	
$12.9_{\{26.9\}}$			
	$40.9_{\{3.6\}}$	$80.6_{\{-22.8\}}$	
	$21.4_{\{8.5\}}$	$58.5_{\{-14.8\}}$	
$9.9_{\{23.5\}}$	$32.8_{\{7.6\}}$	$54.3_{\{-16.4\}}$	$49.6_{\{-24.1\}}$
$12.5_{\{13.8\}}$	$41.6_{\{-9.2\}}$	$49.8_{\{-19.9\}}$	$50.5_{\{-23.3\}}$
	$\begin{array}{c} 31.5\\ \hline 26.0_{\{12.2\}}\\ \hline 13.6_{\{24.1\}}\\ \hline 13.9_{\{8.9\}}\\ \hline 12.6_{\{18.9\}}\\ \hline 37.1_{\{-0.4\}}\\ \hline 12.9_{\{26.9\}}\\ \hline \\ \hline \\ \hline \\ \hline \\ 9.9_{\{23.5\}}\\ \hline 12.5_{\{13.8\}}\\ \end{array}$	$\begin{array}{c ccccc} 31.5 & 50 \\ \hline 26.0_{\{12.2\}} & 58.1_{\{-3.5\}} \\ \hline 13.6_{\{24.1\}} & 42.6_{\{-0.2\}} \\ \hline 13.9_{\{8.9\}} & 55.0_{\{-4.5\}} \\ \hline 12.6_{\{18.9\}} & - \\ \hline 37.1_{\{-0.4\}} & - \\ \hline 12.9_{\{26.9\}} & - \\ \hline - & 40.9_{\{3.6\}} \\ \hline - & - \\ \hline 21.4_{\{8.5\}} \\ 9.9_{\{23.5\}} & 32.8_{\{7.6\}} \\ \hline 12.5_{\{13.8\}} & 41.6_{\{-9.2\}} \end{array}$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$

Table 4.1: Equivalent Rectangular Bandwidths and detection efficiency  $(\text{ERB}_{\{K\}})$  values obtained for every subject and associated center frequencies

From table 4.1, it can be seen that the detection efficiency for every subject decreases as center frequency decreases. This corresponds well with the fact that subjects needed higher SPLs to achieve thresholds when tested for the lower center frequency cases.

Even though there is variation in the ERB values depending on subject and center frequency, there seem to be some general trends in the data. The ERBs at 31.5 Hz appear to be substantially lower than for the other center frequencies. This indicates the effective bandwidth of the auditory filters does not reach its limiting point at frequencies around 100 Hz, but continues to decrease as frequencies approach the infrasonic range.

In the next section the grouped ERB data is analyzed in more detail.

## 4.3 ERB as a function of frequency

For each center frequency the mean ERB value and standard deviation were calculated. The results are shown in figure 4.23, where the standard deviation of each group of values are the vertical lines crossing each mean (circle) value. Table 4.2 shows the mean ERB values obtained for every center frequency.

Note that, because 3 cases had to be discarded from analysis, there are 8 ERB values at 31.5 Hz, 7 at 50 Hz and 6 at 80 Hz center frequency. The 2 values at 125 Hz were included in the graph and table for comparison purposes.

Center frequency (Hz)	31.5	50	80	125
Mean ERB (Hz)	17.3	41.8	56.7	50.0

Table 4.2: Mean Equivalent Rectangular Bandwidth values obtained at each center frequency



Figure 4.23: Equivalent rectangular obtained at each center frequency (mean and standard deviation)

#### 4.3.1 Differences in the ERB values obtained at each center frequency

The data were tested to see if differences between the mean ERB values obtained at 31.5, 50 and 80 Hz were statistically significant. In order to do this, T- tests were performed on the data. The T-test compares the means of different groups of elements, testing if the differences are high enough to be considered statistically different from 0. The null hypothesis is that the means are equal and the alternative hypothesis is that they are different. A p value is found to support or reject the null hypothesis.

The 31.5 Hz group was compared against the 50 Hz group and the 50 Hz group against the 80 Hz group. Because there are 8 ERB values at 31.5 Hz, 7 at 50 Hz and 6 at 80 Hz and the groups to be compared must have the same number of elements, a rule was applied to discard one element in each case. The element that deviates more from the mean was discarded. Then, when comparing 7 ERB values at 31.5 Hz with the 7 values at 50 Hz, a p value of 1.4% was found. This is strong evidence against the null hypothesis, meaning the differences between the ERBs obtained at 31.5 Hz and at 50 Hz are statistically significant. When comparing 6 ERB values at 50 Hz with the 6 values at 80 Hz a p value of 10.8% was found. This indicates a substantial degree of difference between the ERBs at 50 Hz compared to the ones at 80 Hz, although not at such a high significance level as in the 31.5-50 Hz case.

#### 4.3.2 Approximate expression for the ERB in the range 31.5-125 Hz

The data was fitted in order to find an approximated expression that describes the variation of the ERB with frequency in the range 31.5-125 Hz. For the two subjects measured at 125 Hz (subjects 10 and 11), the ERB value seems to stabilize when compared to the ERB at 80 Hz for the same subjects. Even though data at 125 Hz was obtained for just two subjects, the fact that for this subjects the fits where the most accurate and thresholds where measured with more repetitions is considered supporting evidence that the ERB may stabilize between 80 Hz and 125 Hz. Therefore, the ERB value at 125 Hz was set at the same value of the average obtained at 80 Hz. In this manner, an expression describing ERB changes between 31.5 Hz and 125 Hz would reflect data obtained

after averaging more subjects in the range 31.5-80 Hz, and at the same time support the hypothesis that the ERB stabilizes between 80 Hz and 125 Hz.

In a linear frequency scale a cubic polynomial was found to fit the data satisfactorily whereas in a logarithmic frequency scale a quadratic polynomial provided a good fit as well. The expressions are given in equations 4.1 and 4.2.

$$ERB(f) = 11.07 \cdot 10^{-5} (f)^3 - 0.03486 (f)^2 + 3.602f - 65.03$$
(4.1)

$$ERB(f_L) = -151.5(f_L)^2 + 611.2f_L - 558.7,$$
(4.2)

where f is the frequency in Hz and  $f_L$  is the logarithm of the frequency,  $f_L = log_{10}(f)$ ,  $(f = 10^{f_L})$ .

Because of the fewer parameters and common use of logarithmic scale, expression 2 is preferred. Figure 4.24 shows the fitted curve using expression 4.2 and the mean ERBs for each center frequency.



Figure 4.24: Approximated curve representing the variation of the ERB in the frequency range 31.5-125 Hz

When applying the expression recommended in [3], ERB = 24.7(4.27F + 1), where *F* is in KHz, the ERB values deviate in general from the mean ERB values obtained in this work, the ERB in 31.5 Hz being 65% overestimated. Therefore, expression 4.2 is recommended when trying to estimate the ERB in this frequency range.

# **Chapter 5**

# Conclusions

This chapter is divided in three parts. First the applied procedures for the determination of the auditory filter properties are discussed. The pros and cons of the method and approaches with their possible influence in the results are addressed. Following this, main conclusions extracted form the results are suggested. Finally, proposals for further work are given.

## 5.1 Discussion

Exposure and effects of low frequency noises demand a better understanding of auditory filtering at low frequencies. Given that below 100 Hz, auditory filters have not been systematically studied, consistent descriptions of auditory filtering under 100 Hz have not been found. Because of the importance of this frequency region when analyzing spectral components of machinery noise, an effort was done in this project to better describe auditory filters in this less inspected frequency region.

Therefore, main characteristics of the auditory filters in the range 31.5-125 Hz, such as filter shape and ERB, have been analyzed and explored by means of listening experiments. The center frequencies 31.5, 50, 80 and 125 Hz were tested. Experiments were performed under binaural exposure conditions in a room conditioned for controlled low frequency sound reproduction.

To obtain the auditory filter properties, a number of threshold measurements have been performed using the notched noise method with threshold determination based on the ascending method. In the applied method, main assumptions are based in the power spectrum model of masking, indicating only magnitude responses are considered and relative phases of the signals are not taken into account.

One observation regarding the threshold determination, is that the model is sensitive to small variations in threshold level. This is even more critical at low frequencies, given that overall threshold variations where found to be much more restricted than what has been observed from other studies focussed at higher center frequencies. Therefore, the threshold determination method should be as accurate and reliable as possible. In the case of this project, in a tradeoff between time restrictions and accurary, the the ascending method was chosen. It was found that if no repetitions are included, the method might not be the best in terms of accuracy. However, when more repetitions were performed, averaged results showed more consistent trends.

#### CHAPTER 5. CONCLUSIONS

To account for outer and middle ear attenuation, individual equal loudness contours where measured and notched noise stimuli where accordingly compensated. A threshold determination was also performed in order to ensure the ELC was obtained well above threshold. However, it has to be taken into account that it is not fully clear how to estimate outer and middle ear attenuation. The compensation applied is based on previous works which have used standard equal loudness contours under the assumption that such responses are due to outer and middle ear attenuation [3],[24],[15]. Thus, compensating for this response means considering the auditory system from the cochlea up to higher stages. In this project, the origin of auditory filtering was considered to be in the cochlea and it was assumed equal loudness contours would reflect the previous fixed attenuation of the outer and middle ear. However, some questions arise regarding the applied correction:

- The first question is where it is more suitable to apply the correction. The compensation may be applied mathematically after the measurements, in order to fulfill the power spectrum model of masking. On the other hand, it is also possible to apply the compensation a priori, in the stimuli. The second option was preferred because in this manner, under the assumptions made, the signal arriving the cochlea would be the desired one. Besides, it was considered more accurate to apply the compensation individually.
- Secondly, each stimuli consisted of two different kind of sounds, a noise and a tone. However, the ELC correction applied considers the relative loudness perception only for pure tones. Then, a possible limitation is that the relative loudness perception of the notched noise stimuly is not being considered and may be a better approximation. On the other hand, in order to control the level at which the tones reach the cochlea, in principle each tone should be compensated for each subject as well. However, in practical terms this only affects the *K* parameter, which adjusts the threshold curves up and down in overall level. The resulting ERB and the filter shape, would remain however, unaffected.

To determine auditory filter asymmetries, the notched noise method was applied with one asymmetrical configuration. Previous researchers have used two sets of asymmetrical configurations that consisted of mirror image conditions respect to the tone and a symmetrical configuration [15], [3]. From ideal threshold data, it was determined that to measure under only one set of asymmetrical configurations would provide all information needed regarding the filter parameters, according to the assumptions made. This allowed a substantial reduction of time in the experiments, given that mirror image conditions and symmetrical configurations could be avoided. However, it is considered that the use of only one asymmetrical configuration to derive the filter shapes has to be more systematically proven with more subject data using both asymmetrical configurations and after corresponding comparisons.

Results indicate that the filter shape, under a noise level of 54 dB, is generally symmetrical in the frequency region 31.5-125 Hz. It was found that the ERB continues to decrease below 100 Hz, reaching an average bandwidth of 17 Hz at 31.5 Hz. Besides, the dynamic range of the auditory filter was found to be more limited at 31.5 Hz. However, the dynamic range limit of the auditory filters could not be fully determined in many cases, indicating masker separation was not wide enough. For further measurements at low frequencies, it is therefore recommended to include even more separated notched noise configurations.

Comparisons with other works indicate previous descriptions of the changes in ERB with frequency are not accurate when applied below 100 Hz and compared to the obtained data. A wrong estimation of the ERB may have several consequences on perceptual models such as loudness models. Therefore, a new expression describing the observed behavior of the ERB in the frequency range 31.5-125 Hz is suggested.

## 5.2 Conclusions

From the obtained results and according to what has been discussed, some general conclusions have been extracted. The notched noise method had been previously applied for a wide range of frequencies showing consistent results. However, no results had been found below 100 Hz. In this Master Thesis, this was the chosen method to explore auditory filtering in the frequency range 31.5-125 Hz. According to the obtained results, it is considered that the notched noise method can still be applied in this frequency region, leading to coherent results.

From the obtained results, it can be concluded that the bandwidth of the auditory filters does not reach its limiting point at frequencies around 100 Hz, but continues to decrease as frequencies approach the infrasonic range. Considering results from previous works, where the limit for the determination of auditory filter properties was found to be at 100 Hz, the results obtained in this project aim to continue and complement these previous studies.

It was also found that in general terms, at a noise level of 54 dB, the auditory filters are practically symmetrical in the 31.5-125 Hz frequency range. This agrees with results from other studies focussed at higher frequencies, where it has been found that for low and moderate noise levels the auditory filters tend to be symmetrical [15], [3].

The dynamic range of the auditory filters was found to be more limited at the lowest frequency measured, 31.5 Hz. Besides, at this frequency, relatively sharper filters where found in general. It is recommended, however, to include even wider masker separations than the ones used in this work, to determine dynamic range limitations in this frequency range.

The results indicate that approximated expressions describing the ERB changes with frequency, although useful in other frequency ranges, fail to describe what has been obtained in the range 31.5-125 Hz. Deviations are found in general, the average ERB at 31.5 Hz (17 Hz) being overestimated. Therefore, an expression describing the observed ERB behavior is suggested. The general ERB behavior in other frequency ranges, except below 100 Hz, has been shown in figure 4.24. Equation 1.3 has been used to appoximate these results. In figure 5.1, a description of the ERB in low frequencies can be observed. The frequency range 20-80 Hz has been approximated by the new proposed expression (equation 4.2), while the rest of the frequency range is described by equation 1.3.

## 5.3 Further work

Throughout the development of this project, many decisions had to be made regarding important aspects, such as scope of the project and applied methodologies. However, at some of these decision points different approaches could have been taken as well. Main aspects that appeared to be interesting to investigate to the authors are presented next.



Figure 5.1: The ERB as a function of frequency. The expression suggested by the authors for frequencies below 80 Hz is plotted together with the the expression 1.3 used in other frequency ranges.

- This work has considered the notched noise method for binaural exposure conditions. It seems interesting now to explore the situation for each ear separatedly, in order to establish comparisons. In this manner, conclusions about the importance of the central auditory system regarding auditory filtering might be done. Regarding this, a simple experiment is proposed. It consists of determining if the filters can be also observed when the tone is presented in one ear and the masker is input to the other one. If a similar behavior is observed in terms of auditory filtering, this would mean that central processing is at least as important as cochlea performance in auditory filtering.
- Other methods for obtaining the ELC and thresholds could be considered. Especially for the main experiment, as the task was considered rather difficult, a 2AFC based method may allow better determinations. In the case of the ascending method, averaging more repetitions is required to obtain more reliable results.
- In order to estimate individual outer and middle ear attenuations based on loudness comparisons, other signals but the tone could be used. Given that the signals applied and corrected for this attenuation are the notched noise stimuli, a method that determines loudness comparisons using the notched noise signals may be more appropriate and could be developed.
- The use of only one set of asymmetrical configurations was found to be enough to determine auditory filter properties with ideal threshold data. However, the reliability of this approach still needs to be tested with more subject data under mirror image conditions and corresponding comparisons. Therefore, it is proposed to measure subjects under both asymmetrical configurations and compare results in order to see if the resulting filter parameters are approximately equal. This is thought to be an important feature of the notched noise method, given that a substantial reduction of measurement time can be achieved if this approach is applied. The time saved could be directly invested in the threshold determination method, improving its accuracy.
- The widest masker separation used in the notched noise configurations was 1.4 normalized

frequencies ( $g_l = 0.6$ ;  $g_u = 0.8$ ). However, this was found to be not wide enough in many cases to determine the point where the threshold curve flattens off and the dynamic range limit of the auditory filter can be determined. Therefore, it is proposed that for any further measurements at this frequency region, wider masker separations are used in the threshold determination.

• During this work, only one reference level of the noise was considered. However, it is suggested to repeat the measurements including more levels. In this manner, tendencies as the observed in higher frequencies (for example, filters more asymmetrical at higher masker levels) can also be evaluated in the frequency range below 100 Hz.

#### CHAPTER 5. CONCLUSIONS

Bibliography

#### CHAPTER 5. CONCLUSIONS

# **Bibliography**

- [1] P. Buser and M. Imbert. Audition. The MIT Press, 1st edition, 1992.
- [2] Jin Sup Eom, Ae ran Kwon, Jeong In Hwang, Jae Gap Suh, Moon Jae Jo, Sung Soo Jung, and Jin-Hun Sohn. Effect of low frequency noise on psychological responses. 12th International meeting on low frequency noise and vibrations and its control, 2006.
- [3] B. R. Glasberg and Brian C. J. Moore. Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, 1990.
- [4] D. Hammershøi. *Binaural technique a method of true 3D sound reproduction*. PhD thesis, Aalborg University, 1995.
- [5] T. Houtgast. Auditory filter characteristics derived from direct-masking data and pulsationthreshold data with a rippled noise masker. *Journal of Acoustical Society of America*, 1977.
- [6] D. M. Howard and J. Angus. Acoustics and Psychoacoustics. Focal Press, 1st edition, 1996.
- [7] T. Irino and R.D. Patterson. Dynamic compresive gammachirp auditory filterbank for perceptual signal processing. *Acoustics, Speech and Signal Processing*.
- [8] ISO. Normal equal loudness level contours, 2nd edition, 2003.
- [9] R. Lufti and R. D. Patterson. On the growth of masking asymmetry with stimulus intensity. *Journal of Acoustical Society of America*, 1984.
- [10] M. Lydolf. *The threshold of hearing contours of equal loudness*. PhD thesis, Aalborg University, 1999.
- [11] H. Møller and C. S. Pedersen. Hearing at low and infrasonic frequencies. *Noise and health* 2004, 6.
- [12] B. C. J. Moore. Frequency selectivity in hearing. Academic Press.
- [13] B. C. J. Moore. *Psychology of Hearing*. Elsevier Academic Press, 5th edition, 2004.
- [14] B. C. J. Moore and B. R. Glasberg. Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *Journal of Acoustical Society of America*, 1983.
- [15] B. C. J. Moore, R. W. Peters, and B. R. Glasberg. Auditory filter shapes at low center frequencies. *Journal of Acoustical Society of America*, 1990.
- [16] V. Nedzelnitsky. Sound pressures in the basal turn of the cat cochlea. *Journal of Acoustical Society of America*.

- [17] R. D. Patterson. Auditory filter shape. Journal of Acoustical Society of America, 1974.
- [18] R. D. Patterson. Auditory filter shapes derived with noise stimuli. *Journal of Acoustical Society of America*, 1976.
- [19] R. D. Patterson and G. B. Henning. Stimulus variability and auditory filter shape. *Journal of Acoustical Society of America*, 1977.
- [20] R. D. Patterson and I. Nimmo-Smith. Off-frequency listening and auditory-filter asymmetry. *Journal of Acoustical Society of America*, 1980.
- [21] R. D. Patterson, I. Nimmo-Smith, R. Weber, and D. L. Milroy. The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. *Journal of Acoustical Society of America*, 1982.
- [22] R. W. Peters and B. C. J. Moore. Auditory filter shapes at low center frequencies in young and elderly hearing-impaired subjects. *Journal of Acoustical Society of America*, 1992.
- [23] J. O. Pickles. An introduction to the psychology of hearing. Academic press.
- [24] S. Rosen and D. Stock. Auditory filter bandwidths as a function of level at low frequencies (125hz-1khz). *Journal of Acoustical Society of America*, 1992.
- [25] A. Santillán, M. Lydolf, and H. Møller. Low frequency test chamber with loudspeaker arrays for human exposure to simulated free-field conditions. *10th International Meeting*, Low Frequency Noise and Vibration and its Control.
- [26] A. Santillán, C. S. Pedersen, and M. Lydolf. Experimental implementation of a low-frequency global sound equalization method based on free field propagation. *Applied Acoustics*, 2006.
- [27] A. P. Sek. Auditory filtering at low frequencies. Archives of acoustics, vol 25, part 3.
- [28] M. F. Spiegel. Thresholds for tones in maskers of various bandwidths and for signals of various bandwidths as a function of signal frequency. *Journal of Acoustical Society of America*, 1981.
- [29] Inc. The MathWorks. *Matlab Help 6.5*.
- [30] T. Watanabe and S. Yamada. Study on perception of complex low frequency tones. *Journal* of low frequency noise, vibration and active noise.
- [31] E. Zwicker and H. Fastl. *Psychoacoustics-Facts and Models*. Springer-Verlag, 2nd edition, 1999.
- [32] J. J. Zwislocki. *Auditory Sound Transmission: An Autobiographical Perspective*. Lawrence Erlbaum Associates, 1rst edition, 2002.

Appendix

# Appendix A

# Anatomy of the auditory system

This project concerns an specific aspect of our hearing system performance: the critical band. The critical band in this report is considered to be closely related to the cochlea performance. But to better understand the project and some of the considerations that have been done we are going to present an overview on the human auditory system and the human sound perception.

In order to give a more practical explanation about the hearing system, the natural direction of hearing process is followed, from when a sound wave arrives to the outer part of the ear to when it evokes an auditory event. To analyze it in a more systematically way a initial division can be made. Thus, we may consider two main parts in the auditory path of the human being, in which they can be found many organs involved. These main parts are [1]:

- 1. The peripheral auditory pathway.
- 2. The central auditory pathway.

Both parts are very important for the understanding of human hearing, we will focus first on the peripheral auditory pathway, in which the cochlea is included. However a brief description of the central auditory pathway will be given too, especially because we will consider binaural hearing and that means internal auditory information processing should be taken into account.

## A.1 The peripheral auditory pathway

The main purpose of peripheral auditory pathway is to detect and, to some extent, to analyze the sound. Three sub-parts can be considered within the peripheral auditory pathway. These three sub-parts of the ear are shown below in figure A.1 and are called *outer ear*, *middle ear* and *inner ear*.

#### A.1.1 The outer ear

In the peripheral auditory pathway, the outer ear receives the sound and sends it to the *ear drum* or *timpanic membrane* through the ear canal. This is also called the *Acoustic Path*. Its most external element is the *pinna*, in charge of the sound collection. The pinna is connected by the *concha* 



Figure A.1: The outer, middle and inner ear and their different organs.

to the *auditory canal*. In this part of the peripheral auditory pathway we can point out two main effects. First, an effect of acoustic filtering by the pinna and the concha which will help in sound localization ability. Second, the auditory canal will reinforce the frequency region around 4 kHz due to a resonance effect. The outer ear-middle ear boundary is the *timpanic membrane*. The tympanic membrane is a light, thin elastic structure consisting of three different layers. It converts acoustic pressure variations from the outside into mechanical vibrations to the middle ear.

#### A.1.2 The middle ear

Also called *timpanic cavity*, it transforms the pressure variations acting at the tympanic membrane into the internal mechanical vibrations of the bone structure formed by the *hammer*, the *anvil* and the *stirrup*. Then, these vibrations will be converted into compressional waves in a fluid at the inner ear. So it can be be considered as the bridge between the vibration of the ear drum at the end of the outer ear and the vibration of the oval window at the inner ear. Two functions of the middle ear must be emphasized:

- 1. The amplification effects due to mechanical reinforcement of the incoming vibration. It appears thanks to:
  - Lever effect of hammer and anvil.
  - Area difference between the tympanic membrane the tympanic membrane and stirrup footplate, which is attached to the oval window.

As a result of these effects the pressure at the oval window is around 34 times that at the tympanic membrane. In fact this is going to compensate the transmission losses due to the change of propagation media, from air at the outer ear to the inner ear liquids.

2. The protection against too high levels of the incoming sound. This feature is carried out by two muscles in the middle ear: the tensor tympani and the stupids muscle. These muscles contract automatically in response to sounds with levels greater than approximately 75 dB.

This have the effect of increasing the impedance in the middle ear by stiffening the ossicular chain. Thus, the transmission efficiency is reduced attenuating low levels signals up to 12 to 14 dB (but only for frequency signals below 1000 Hz).

#### A.1.3 The inner ear

Is made up of a series of intercommunicating cavities that together form the bony labyrinth: the cochlea, the vestibule and the semicircular canals. The cochlea is the most important organ of the three, specially in regarding the present study, because the basis on which the critical band concept lies is the behavior of the basilar membrane. The cochlea's main function is to transform the energy of the compressional wave within the inner ear fluid into nerve impulses which can be transmitted to the brain.

The cochlea consists of a tube coiled into a spiral. At one end we find the already mentioned oval window as well as the round window. This is the *base*. The opposite extreme is the *apex*. The tube is divided in three cavities by two membranes: the *Reissner's membrane* and the *Basilar membrane*. The outer channels are the scala vestibuli and the scala tympani and are filled with an incompressible fluid known as perilymph -through which the sound vibrations are transmitted-, and the inner canal is the scala media. The scala vestibuli terminates at the oval window and the the scala tympani at the round window. The figure A.2 is a cross section to the cochlea and shows its actual aspect and its internal structure.



Figure A.2: The snail-like structure of the cochlea and its internal cavities and membranes.

The cochlea is playing the most important role in the analysis of sound by means of a frequencyspace analysis carried out by the basilar membrane. In shape, the basilar membrane is both narrow and thin at the base end of the cochlea , becoming both wider and thicker along its length to the apex. Thanks to this particular shape, input pure tones at different frequency will produce a maximum basilar membrane movement at different positions or *place* along its length. This is the basis of the *place* analysis of sound by the hearing system. the figure A.3 presents an schema of this analysis relating position of the maximum of movement at the maximum membrane with frequency that elicits this response.



Figure A.3: Tonotopical analysis of frequencies along the basilar membrane.

#### A.1.4 Transduction from mechanic waves to neural signals

The transformation of compressional waves into nerve impulses represents the starting point of the central auditory pathway and it is produced as follows: The receptor organ which generates nerve impulses in response to vibration of the basilar membrane is called organ of corti. This organ lies on the surface of basilar fiber and basilar membrane and the sensory receptor in this organ are two types of hair cells.

- 1. Internal hair cells: Single row of cells roughly 3500 hair cells and of about 12 micrometer in diameter each one.
- 2. External hair cells: Three or four rows of cells. There are about 25000 and have a diameter of 8 micrometer each.

Inner and outer hair cells have very different functions. The inner cell seems to have major role in afferent information transport, while outer cells are involved mainly in efferent information transport. Most of the afferent neurons, which carry information from the cochlea to the higher level of auditory system are connected to inner hair cells, each inner cell is contacted by about 20 neurons.

The mechanism involved in the process of converting vibrations into neural firing makes use of these two kind of cells and can be described as follows: The hairs of the outer hair cells are fixed to the reticular lamina. Upward movement of the basilar fibers hit the reticular lamina through the rods of corti. The motion of the reticular lamina makes the hair moves back and forth against the tectorial membrane which is not the case for the inner cell, since there stereocilia are not necessarily attached to the tectorial membrane and a fluid move back and forth over the hairs and bend them. Bending the hair in one direction depolarizes the hair cell and bending the hair in the opposite direction hyperpolarizes them. This process excites the bases and sides of hair cells which synapse with the network of cochlear nerve ending creating action potentials in the neurons of the auditory nerve.

## A.2 The central auditory pathway

So far we have explained what is also known as the acoustical path: The outer ear receives the sound and sends it to the middle ear. The middle ear serves to transform the energy of a sound wave into the internal vibrations of the bone structure and ultimately transforms these vibrations into a compressional wave in the inner ear. The inner converts then the energy of a compressional wave within the inner ear fluid into nerve impulses which can be transmitted to the brain. At that point begins what is called the central auditory pathway.

The figures A.4 shows an schema of the central auditory pathway which represents the main stages of the sound information processing in the brain. These figures distinguish between the monaural and binaural auditory central pathways. We are not going to look in detail on how exactly are processed these informations, but at least we are going to mention the main process and the brain localization of the auditory information processing centers after leaving the cochlea. These are:



Figure A.4: Left figure: Binaural central auditory pathway. Right figure: Monaural central auditory pathway.

- 1. Cochlea
- 2. Cochlear Nucleus (CN): With two main parts, the contralateral anteroventral cochlear nucleus (AVCN) and the dorsal cochlear nucleus (DCN).
- 3. Medial nucleus of the trapezoid body (MNTB).
- 4. Superior Olivary Complex (SOC): With two main parts, the medial superior olive (MSO) and the lateral superior olive (LSO).
- 5. Dorsal nucleus of the lateral lemniscus (DNLL)
- 6. Inferior Colliculus (IC).

- 7. Medial geniculate nucleus (MGN).
- 8. Auditory Cortex (AC).

The same is represented in a more simplified way and showing its actual localization in the human brain in figure A.5.



Figure A.5: Primary Auditory Pathway: from cochlear nerve to the auditory cortex

#### A.2.1 The auditory cortex

The auditory cortex is the most important center of sound processing in the brain The location of the *auditory cortex* is a deep groove in the brain called fissure of Sylvius. Th auditory cortex is divided into primary auditory cortex and several other surrounding fields. Figure A.6 shows the location of the auditory cortex in the brain. It is in this part of the cortex where the most part of sound interpretation take place.



Figure A.6: The Auditory Cortex: View of the location of the auditory cortex in the brain

# **Appendix B**

# Human sound perception

This section does not intend to be a profound description of the human ability of hearing, but only a first approximation to the basic principles of sound cognition. Then, below there is a basic explanation about how do human beings hear in terms of perception and not of the physical signal only. Particularly, this project concerns an specific aspect of our hearing system performance: the critical band. This concept is closely related to other features of the human sound perception. Thus, to better understand the project and some of the considerations done, the most relevant of these features are going to be presented in this appendix: the loudness perception and the masking phenomenon.

## **B.1** Threshold and loudness

The first thing it can be analyzed is what human beings can hear and what they cannot. What human beings perceive are changes in the pressure. So first question is what kind of changes do they perceive as sound. This leads to the consideration of two different domains, the frequency of the sound wave and the pressure level of sound wave. Because not every frequency of sound is audible by human auditory system. And even if the frequency is in audible range for humans not every sound wave can be perceived.

The standard frequency range considered as audible range is traditionally the range between 20 Hz and 20 kHz. However this range may change much along different individuals together with aging and in fact people is normally capable of perceiving frequencies below 20 Hz, although some features of the sound might be lost.

On the other side the loudness perception is influenced also by other effects, as the kind of the sonorous stimulus (e.g. pure tones or complex sounds and its bandwidth), the frequency of this stimulus and the possible simultaneous presence of other sound waves.

The first consideration about sound loudness perception is the pseudo-logarithmic response of human auditory system to sound pressure changes. Because of that, sound loudness is often represented by decibels scales. But this non-linear response is not constant along the frequency. To represent that response, measurements were made on a very large population of young subjects [8]. The results are known as the curves of human threshold of hearing and equal loudness contours





Figure B.1: Hearing threshold and equal loudness contours (adapted from ISO 226: 2003). The human auditory system response better to middle frequency stimulus.

In these curves it is observed that the ear identifies middle frequency sounds as louder than low or high frequencies when equal sound pressure is applied. So it is in this range where human ear is sharpener or more efficient. To perceive a similar sensation of loudness much higher pressure is required in low and high frequencies. Furthermore, this changes in loudness perception along frequency are different along wave pressure itself: when the level of the sound is higher the curves become flatter.

## **B.2** Masking

Another important effect on loudness perception is that not every sound that is above the threshold of hearing (the lower curve in figure B.1) can be perceived. If a sound is given to a subject, while a louder is also present, then it is possible that the subject does not perceive the lower one. This process seems intuitive, but on the psychoacoustic and cognitive levels it becomes very complex. The term for this process is masking, it is probably the most researched phenomenon in audition and depends on different parameters. These parameters are mainly the frequency of masker and masked sounds, their spectral structures and their difference in sound level. The figure B.2 represents the masking curve produced by a single pure tone over other tones when it is reproduced at different levels.

Definitions of masking differ according to what field it is being related. Masking as defined by the American Standards Association (ASA) is the amount (or the process) by which the threshold of audibility for one sound is raised by the presence of another (masking) sound . For example, a loud car stereo could mask the car's engine noise. The term was originally borrowed from studies of vision, meaning the failure to recognize the presence of one stimulus in the presence of another at a level normally adequate to elicit the first perception.



Figure B.2: Simultaneous masking pattern of a single masking tone and for differents masker levels. For each level, a tone under the corresponding curve, would result inaudible.

# **Appendix C**

# The low frequency room at Aalborg University

In this section a general description of the available facilities for the controlled reproduction of low frequencies is presented. The setup basically consists of a low frequency room and respective control room with a control PC. This setup had been implemented in the acoustic laboratory at Aalborg University for the use in other low frequency related projects before the start of this project [26],[25].

A brief overview of the properties of the low frequency room and of the required processing to achieve a desired response is given in this appendix. References are also given for further details about the functioning of the low frequency facilities.

The main objective in the low-frequency room setup was to be able to reproduce very low frequencies at relatively high sound pressure levels and, at the same time, have a controlled sound field throughout the entire bandwidth of the signals. The different steps that were necessary to reach this objective are described in what is following as well.

## C.1 General description of the room

The room has been designed to reproduce low frequencies, including infrasonic sounds. It was reconstructed from an existing infra-sound chamber. It has inner dimensions of 2.72m x 2.70 m x 2.40 m. Figure C.1 shows a diagram of the low frequency room.

There are 20 loudspeakers on each side of the two walls parallel to the x direction. The test chamber has been designed with double concrete walls to obtain a high sound insulation. To reproduce very low frequencies it had to be designed as airtight as possible.

More details about its construction and materials can be found in [26] and [25].



Figure C.1: Schematic representation of the low frequency room (taken from [26])

# C.2 Reproduction modes

The reproduction modes of the low-frequency room consist of a *pressure field mode* and a *free field mode*.

#### C.2.1 Pressure field mode

This reproduction mode is valid when the wavelength of the sound is much larger than the room dimensions. For the low-frequency room a pressure field reproduction can be achieved up to a maximum frequency of about 25 Hz. In this mode, the amplitude of the sound field will be practically the same at any position, when all loudspeakers are fed with the same signal. The lowest frequency that can be reproduced depends on the airtightness in the room. Experimentally it has been found to be 0.07 Hz [25].

The maximum sound pressure level achieved with this mode depends on the allowed harmonic distortion of the system. The criteria followed is that the second harmonic has to be attenuated 30 dB respect to the fundamental, third harmonic 40 dB and higher harmonics 50 dB lower than the fundamental [25]. According to this, there will be a limitation on the maximum displacement of the loudspeaker's membrane. With these criteria the maximum sound pressure level achieved in the room lies around 130 dB [25].

At frequencies lower than about 25 Hz, the response is practically flat. However, at higher frequencies the excitation of the room modes becomes evident and peaks and deeps appear in the frequency response curves. Special signal processing is required to achieve a flat response in higher frequencies. The description of a free field reproduction mode that implements flat response in higher frequencies is given next.

#### C.2.2 Free field mode

In order to achieve a flat response in higher frequencies, a sound equalization system has been previously designed [26], [25]. The idea is to generate a traveling plane wave with the 20 loud-

speakers in one wall and actively absorb the sound when it reaches the other wall using the other 20 loudspeakers.

The implementation uses one channel for each loudspeaker, yielding 40 different channels. The input signal at each channel is convolved with a specific FIR filter and the resulting 40 filtered signals are amplified and fed to the loudspeakers. The FIR filters are obtained from a least-squares approximation which minimizes the error between the desired signal (a plane propagating wave) and a measurement of the sound field for each channel at various predetermined positions [25]. Figure C.2 shows a diagram of the implemented system.



Figure C.2: Free field reproduction system diagram

The implemented system works with a 48 KHz sampling frequency, therefore upsampling is required after performing the filtering of 1 KHz input files.

The main limitation of the free field reproduction mode comes from the fact that in order to have the same sound pressure amplitude at different frequencies, the displacement of the membrane of the loudspeakers has to be inversely proportional to the frequency [25]. The maximum possible SPL will therefore decrease with frequency, as more displacement is required and harmonic distortion increases with increasing displacement. For example, the maximum SPL at 10 Hz will be approximately 106 dB [25].

The free field mode provides a listening region where the response is flat. This region consists

of all the volume except points within 0.7 m from each of the two walls with the loudspeakers.

#### C.2.3 Providing a hybrid reproduction mode

Before the start of the project work, the low frequency room had been previously set up and used for either free field or pressure field reproduction modes. However, because of the high sound pressure levels required at low frequencies in this project, it was necessary to have the low frequencies reproduced in the pressure field mode. At the same time, the higher frequencies were satisfactory reproduced only in the free field reproduction mode.



Figure C.3: Flow diagram of signal path for hybrid reproduction mode in the low frequency room. The low-pass and high-pass crossover filters are the boxes labeled as *LP* and *HP*, respectively

No mode by itself would have allowed an acceptable reproduction of the noise signals. This is the reason why a hybrid reproduction mode was necessary to implement, reproducing both, the low frequencies in pressure mode, and high frequencies in free field mode, at the same time. This was done by means of a pair of crossover filters, which filtered out the lowest frequencies so that they do not go through the special filtering required for free field reproduction mode and then they are added back to the high frequencies after this filtering. A flow diagram of the system with the crossover implementation is shown in figure C.3. The room frequency response with this mode implemented is shown in figure C.4.


Figure C.4: Room magnitude response in hybrid reproduction mode

As can be seen, the lowest frequencies up to approximately 10 Hz are reproduced fully in pressure mode and from approximately 30 Hz in pure free field mode. In between, the crossover filter increasingly attenuates the pressure field reproduction to finally meet the free field reproduction mode at around 30 Hz. There is some small attenuation of the free field response in points close to the point where the reproduction modes meet (from approximately 30 Hz to 40 Hz).

Based on the room frequency response, new maximum possible values of reproduction were estimated. The implementation of the hybrid reproduction mode allowed to stay close to the 60 phon curve for the ELC measurements.

### C.3 Listener position in the low frequency room

In the free field reproduction mode a determined listening area where the response is flat is obtained. With the new reproduction mode, however, it was not clear what listening position or region was more adequate to use. This section illustrates how an adequate listening position was determined.

#### C.3.1 The effect of the presence of the chair

With the purpose of having a controlled sound field, the effect of the chair in the low-frequency room was measured. Figure C.5 shows the room magnitude response with and without the chair inside the room.

As can be seen in figure C.5, there is in general little effect of the chair. The main differences appear in the higher frequencies as expected, because of diffraction of the sound waves. For this measurement the chair was positioned in the standard position selected for the measurements of threshold, ELC and masked thresholds. The room response was measured at different positions of the chair, to see if there was an optimal or more convenient position to place the listeners. This is described next.



Figure C.5: Room magnitude response with and without chair positioned inside.

#### C.3.2 The effect of the position of the chair

In order to find an optimal position of the chair for the measurements, its position was moved and the room responses evaluated. In each case, the microphone was positioned where the listener's head would approximately be. The change in position consisted in considering 3 different distances to the front wall. The distance to the lateral walls was kept the same. Figure C.6 shows a schematic representation of the measured positions. The room magnitude responses are shown in figure C.7.



Figure C.6: Schematic representation of test positions of the chair and microphone in the low frequency room (distances in cm, H is the room height and h the microphone height).

Figure C.7 shows that there will be almost no difference in placing the chair in position 1 or 2, but if the chair is positioned more close to the front speakers (position 3), there will be substantially more attenuation at the crossover point. Note that position 3 is within the "safe" free field mode listening region, meaning that following this criterion would not have been adequate.

#### C.3.3 The effect of the presence of a listener

The effect of the listener in the sound field was considered also. Two measurements were performed with a person inside the room. In one measurement the microphone was positioned close



Figure C.7: Room magnitude response for different positions of the chair.

to the listener's head and in the other more far over the head. Figure C.8 shows the results.



Figure C.8: Room magnitude response with a listener inside for two different microphone positions.

The main differences appear approximately from 200 Hz. The fact that the response increases between 200 and 300 Hz when the microphone is close may indicate reflections from the listener's head and body. The response in the far measurement position resembles the measurement without chair (see figure C.5), indicating the microphone is in a position where the free field playback is not disturbed by the chair and body.

It is thought that the ideal microphone position would be one just outside the pinna, but reflections from the listener's head would contaminate the measurement of the incoming sound. Following the same reasoning, it may be argued that the best estimation of the sound field is the measurement without the listener, at a position where the ears would be. Besides, the fact that different listeners may alter differently the sound field would complicate any compensation for the room frequency response. These are the reasons why no compensation for the presence of the listener was intended.

### C.3.4 The effect of differences in sitting positions and listener's heights

The response of the room was measured with the chair in a fixed position (position 1 see figure C.6,  $P_1$ ) with the microphone at positions where the listener's head would be. The chosen points aim to represent the possibility of different sitting positions and listener's heights. The 3 tested positions are 15 cm left from position 1, 15 cm to the right of position 1 and 20 cm from position 1 to the front wall. The results are plotted together with position 1 in figure C.9.



Figure C.9: Room magnitude response at different points around position 1.

The results show that there will be almost no variation around position 1 considering the possibility of different sitting positions. The curve that most deviates from the others is 20 cm in front, showing the same effect of more attenuation at the crossover point as one gets closer to the front wall shown in figure C.7. However, the deviation is not so pronounced, and considering the fact that the most comfortable positions will be the ones where the listeners have their back on the chair, this position is considered more unlikely to be used.

The effect of different heights was also tested with 3 similar points around a point 20 cm higher than position 1. The obtained results show a similar pattern as in figure C.9, with practically the same magnitude response as the measured at point 1.

# C.4 Compensating the non-flat response of the room for the hybrid mode

The final set up used in the room (hybrid mode) resulted in an overall frequency response that was not flat (see figure C.5). However, an overall flat response was desired.

To account for the non-flat room frequency response, a compensation filter was designed. The target function of the filter is the inverse magnitude response of the room, so as to achieve a flat output when reproducing the signals.

The compensation filter approximates the target magnitude response recursively using a Yule-Walker method based on a least-squares procedure. The design produces an IIR minimum phase approximation of the target response. The program used to obtain the filter coefficients, *LFR\_compensation\_filters*.

was written in MATLAB and can be found in the attached CD.

The low frequency room response was measured with an MLSSA system and the measurement has already been shown in section C.2.3. Given that the room impulse response was sampled at 8000 samples/sec, it was necessary to downsample before running the Yule-Walker method to approximate the target the magnitude response. This was necessary because the filter design method tries to estimate the magnitude response in the overall bandwidth of the signal, loosing accuracy in the estimation at very low frequencies.

Because the signals to be used in the control PC for the low frequency room require to be sampled at 1000 Hz, the downsample factor was chosen to be 8. Given that the highest frequency of the noise signals to be reproduced do not exceed approximately 350 Hz, the target response was flattened for frequencies higher than this (up to 500 Hz which is the maximum frequency, fs/2). Because the lowest frequency to be reproduced will be approximately 8 Hz, for the lowest frequencies the target response was also flattened. Figure C.10 shows the target and designed responses.



Figure C.10: Target and designed compensation filter responses

Applying the designed filter to the *MLS* signal used for the previous measurements, the compensated room magnitude response was measured in the room and is shown in figure C.11.



Figure C.11: Compensated low-frequency room magnitude response

As can be seen, the response is practically flat until approximately 360 Hz. The main deviations are less than 2 dB and are confined to the frequency region from 20 Hz to 30 Hz.

### APPENDIX C. THE LOW FREQUENCY ROOM AT AALBORG UNIVERSITY

# **Appendix D**

# Threshold determination methods and time consumption

The most common methods for threshold determination that were found in the related literature are the 2 down-1 up or 3 down-1 up methods. The two alternative forced choice (2AFC) decision procedure was incorporated in these methods [24]. On the other hand, other traditional methods have proved to achieve good performance as well. Among them the, the ascending method is considered to be a fast method [10]. Moreover, this method was the one used in the first phase of the experiment and was already available. Therefore, two methods were considered for phase two of the listening experiment: the ascending method (used in phase one) and the implementation of a short version of the 2 down-1 up approach used in [24].

### D.1 The 2 down-1 up 2AFC method

This method is a descending method in which the subject is asked to answer twice for the same presentation level and only if both answers are positive the level is pulled down. If one of the answers is different, the level of the signal is raised up (2 down-1 up). The manner in which the subject can provide an answer is called 2 alternative forced choice. In the 2AFC the subject is presented twice a stimulus and he or she is forced to state where the signal was present. The procedure consists of 4 level presentations as familiarization procedure with 10 dB steps. After the first two reversals, 5 dB steps in the reduction are applied. Again, after 2 reversals more, reductions in 3 dB steps are finally applied.

The calculation of threshold can be made with the last 4 reversal levels. The minimum proposed stimuli duration was 1.5 seconds for each alternative with a 0.5 second silence in-between. One of the alternatives would contain only the masker and the other would comprise 0.7 seconds of only masker and 0.8 seconds with both masker and tone. It was considered that shorter durations in the stimuli might not be enough, given that the integration time in the ear for the lowest components of some of the stimuli could be violated. With relatively short stimuli, the long term spectrum of the noise would not be as planned. Besides, it is not sure if the tones could have had sufficient number of cycles, and this might have an effect on their perception. With the proposed durations, and assuming the subjects may take around 4.5 seconds to provide an answer, for each tone level around 8 seconds could be used.

An ideal subject (perfectly consistent), would then require 16 presentations in an optimal case.



Figure D.1: Temporal structure of the 2AFC stimulus proposed for the listening experiment.

This would mean a total of 128 seconds (see figure D.2). If, for example, 2 inconsistencies are considered as an average performance, more presentations would be needed. Considering this, an average case would take around 20 presentations. In this case, the total time would be around 160 seconds per threshold determination.



Figure D.2: Total time consumption of the 2down-1 up 2AFC threshold determination method for an ideal case -shortest possible-.

If 10 notched noise configurations and 2 repetitions are needed per center frequency, more than 53 minutes of running experiment would be reached. Considering 2 breaks of 10 minutes -the minimum we should consider-, the experiment would take 1 hour and 15 minutes per subject for just one center frequency. Therefore, the option of considering all 4 center frequencies for each subject should be discarded if using this method.

Another variation that could help to reduce the time consumption would be presenting continuously the masking noise during the whole experiment. In this manner, each stimuli length could be reduced as the problem regarding integration time for the noise would be solved. However, this option would never reduce the total time to less than about 45 minutes per center frequency.

### **D.2** The ascending method

The ascending method was already used in phase 1 of the experiment. Consequently, time consumption was approximately known for this method. Thus, just few minor changes were considered, such as changing the familiarization starting level. As an average, subjects used around 70 seconds per threshold determination during phase 1. An average of 80 seconds is considered for the calculations, as the task could be found more difficult in this case.



Figure D.3: Structure of the stimuli for the ascending method. The noise is played back continuously as the tones are reproduced for two seconds with a pause in between.

The structure of the signals can be seen in figure D.3 and the procedure and duration of one determination for a *consistent* subject is shown in figure D.4. A consistent subject is one who provides only correct answers with respect to his own threshold. The total duration seems to be relatively short, but it should noticed that the timing is just an approximation and the example is one of the best possible cases. It should be also noticed that in the case of the ascending method, the subject is expected to press the button during the time the sound is being played back. This makes the next level to be presented shortly after the previous is finished, making the procedure faster.



Figure D.4: Example of a threshold determination with the implemented ascending method for an ideal case with a totally consistent subject -shortest possibility-.

#### APPENDIX D. THRESHOLD DETERMINATION METHODS AND TIME CONSUMPTION

As 10 threshold determinations (for each of the 10 notched noise configurations) and two repetitions are needed, almost 27 minutes would be required for each frequency. With one break in between the experiment would take around 40 minutes. It would still not be possible with this method to investigate the 4 center frequencies individually. If this was the case, the experiment would take too long, considering subjects were volunteers. Nevertheless, 2 center frequencies could be considered, which still offers the possibility of establishing comparisons between different frequencies.

To conclude this time consumption comparison, it should be noticed that when doing the same number of measurements, the 2 down-1 up 2AFC method would take almost double of time compared to the ascending method.

# **Appendix E**

# Threshold and equal loudness contours results

In order to account for the variations on perception of loudness along frequency, the signal stimuli were shaped with respect to the equal loudness contour of each subject for a level loud enough to be clearly audible but not above the limit of the room. The ELC chosen as reference was that corresponding to 78 dB at 80 Hz (around 60 phons curve).

The frequency range of interest was that between 8 and 250 KHz. For these frequencies both the mentioned ELC and the threshold of hearing were obtained. Comparing both we can observe the safety margin in dB from audible to maximun level admissible for each subject. However only the ELC was directly used further in the project. For the threshold of hearing determination the ascending method was used, and for finding the equal loudness contour the maximum likelihood method was used.

In this appendix they are presented the results obtained for both the threshold of hearing and equal loudness contour for 60 phons. They are shown all the curves obtained, subject by subject, including that (subject 8) who was discarded for the rest of the test. Note that two other subjects were analised but the required level for this subjects to match the 8 Hz tone with the reference (80), was so high that it was not possible with the set up available. For this reason results could not be achieved and they are not in this appendix.



Figure E.1: Thresholds of hearing and reference equal loudness contours. Subjects 1 (left) and 2 (right).



Figure E.2: Thresholds of hearing and reference equal loudness contours. Subjects 3 (left) and 4 (right).



Figure E.3: Thresholds of hearing and reference equal loudness contours. Subjects 5 (left) and 6 (right).



Figure E.4: Thresholds of hearing and reference equal loudness contours. Subjects 7(left) and 8 (right).



Figure E.5: Thresholds of hearing and reference equal loudness contours. Subjects 9 and 10 (right).



Figure E.6: Thresholds of hearing and reference equal loudness contours. Subjects 11 (left) and 12 (right).

### APPENDIX E. THRESHOLD AND EQUAL LOUDNESS CONTOURS RESULTS

# **Appendix F**

# **Generation of notched noise signals using MATLAB**

The notched noise stimuli were generated using MATLAB [29]. To obtain the notched noise stimuli to be presented to the subjects in the listening experiment 3 steps were necessary. The first step was to design the filters with the chosen configurations for the noise bands around the tones. These filters are called here the "notched noise filters". The second step was to shape the notched noise stimuli obtained after step 1 according to the ELC curves that were measured for each subject. The third and final step was to apply the hybrid mode compensation filter to the resulting signals obtained after step 2.

### F.1 Step 1: Design of the notched-noise filters

The "Filter Design & Analysis Tool (FDA)" was used for designing the appropriate filters and obtaining the filter coefficients. The filter coefficients were input to a program used for generating and filtering the noise signals. This program is called *"signals.m"* and can be found in the attached CD.

In most cases three filters were needed to obtain the pre-designed notched noise configurations. At first, a stopband filter was designed and used to filter a white noise signal. Following this, the filtered noise signal was highpassed and then lowpassed (the order is not important since the same result is achieved in any case). The cutoff frequencies for every notch configuration are the ones indicated in figure 2.12. Figure F.1 shows a schematic representation of the required filtering to obtain the notch noise configurations.

### F.1.1 Determination of filter type and filter parameters

The "Filter Design & Analysis Tool" allows to enter parameters such as filter type, filter order, stopband frequencies, stopband attenuation and sampling frequency. All this information is used to calculate the filter's response.

Based on inspection of the magnitude response of the filters, it was established that a Chebyshev type II IIR filter gave responses that were satisfactory according to what was desired. The





filter order was chosen to be 200 in order to achieve transition bands as narrow as possible. The stopband attenuation was set at -80 dB. The filter achieves a completely flat passband and ripple is only observed in the stopband. An example of a designed response is shown in figure

F.2.



Figure F.2: Example stopband filter using a 200 order Chebyshev IIR filter with 75 Hz and 150 Hz stopband frequencies

The response shown in figure F.2 is the one that is calculated in the FDA tool. The calculated filter performs well in low frequencies and has a very sharp transition, achieving practically a square box response.

Using the same procedure the lowpass and highpass filters were designed. The filters where also 200 order Chebyshev type II IIR, and the cutoff frequencies for each center frequency were chosen according to the desired notched noise configuration (see figure 2.12).

Once the filters were designed, white noise signals were generated and successively filtered with the 3 filters according to what is shown in figure F.1. Wavefiles were stored and then reopened for testing the response. To avoid clipping, the data was normalized by the maximum value of the noise before storing the wavefiles. FFT analysis was performed on the resulting files for testing the response. This is described next.

#### **F.1.2** Testing the filter response

Once the filters were designed they were used to filter a white noise signal generated in MATLAB. Because of the fluctuating nature of the noise, linear average of FFT's performed on windowed segments of the noise was used in order to display the resulting magnitude responses. To attenuate spectral leakage, Hanning windows were used.

The FFT size was set to N = 1024, and given that  $f_s = 1000$  Hz this gives a resolution of  $f_s/N \approx 1$  Hz for observing the results. The linear average was performed by averaging several 50 % overlapping blocks of 1024 point FFT magnitude responses. Figure F.3 and F.4 show example responses for cases of narrow notch (generally present at the low center frequencies) and wider notch (generally present at the higher center frequencies). The gains were manually set to zero dB for illustration purposes (given that in MATLAB the calculated gains will have arbitrary magnitudes depending on the FFT size). The program that performs this calculations is called *signals\_overlap.m* and can be found in the attached CD.



Figure F.3: Average long term spectrum of filtered white noise signal, 20 Hz wide notch



Figure F.4: Average long term spectrum of filtered white noise signal, 160 Hz wide notch

As can be seen, the notches are very sharp. According to the threshold data shown in [3] and [15], the signal level at threshold does not vary more than 45 dB. This considers the difference between the threshold SPL at the narrowest notch separations and at the widest notch separations, being the starting threshold SPL values approximately equal to the noise level. Thus, the signal to noise ratio achieved inside the notch, that is where the tones will be placed, can be considered sufficiently high.

### F.2 Step 2: Applying the ELC shaping filters

The objective of the ELC measurement was to apply a filter that shapes the notched noise signals so that they have the same shape as the obtained ELC curves. Given that individual ELC curves were obtained, individual stimuli will be obtained after applying this filters.

In order to obtain a smooth continuous curve from the 8 discrete measured frequencies, a cubic spline interpolation was performed [29]. Given that in the main experiment the noise levels will be approximately the same as the levels of the ELC measurement, relative gain values respect to the reference level (80 Hz at 78 dB) were obtained rather than absolute gain values.

The gain value at each frequency was obtained by subtracting the reference SPL level from the SPL level obtained at each frequency where equal loudness was determined. The gain at the reference frequency (80 Hz) was set to zero dB. Following this, the interpolation was performed to obtain a smooth continuous gain curve.

To obtain a filter that follows the desired gain curve, a Yule-Walker method to approximate a desired magnitude response was used [29]. The desired magnitude response was obtained after the gain values in dB were transformed into absolute magnitude values. The method used is in principle the same used for obtaining the hybrid mode compensation filter (see section 2.3.2).

Because the required SPLs at frequencies lower than 8 Hz will be in most cases exceeding the dynamic range of the low-frequency room, the interpolated curve was manually flattened for frequencies below 8 Hz, before running the filter design method.

Figures in E show the ELC curve interpolations, where the measured values are shown in circles.

### F.3 Step 3: Applying the hybrid mode compensation filter

The hybrid mode compensation filter accounts for the non flat frequency response in the room. The filter has been tested and described in section 2.3.2. All noise signals were filtered with this compensation filter before their storage as wavefiles.

The final signals where stored as wavefiles. To avoid clipping and allow an easier reproduction in the low frequency room, every signal was normalized by the same factor. This factor was the single maximum value found on the notched noise stimuli of all subjects.

# Appendix G Listening experiments: Instructions sheets

As explained in the main report two different listening experiments sessions were needed during this work. The first session was performed using 14 subjects and consisted of an absolute threshold of hearing determination and an equal loudness contour determination (ELC) for pure tones. The second measurement was a threshold of hearing determination for pure tones in presence of a notched noise masker. This measurement was carried out for the 11 selected subjects from the original 14.

To perform the experiments the subjects needed some instructions. In this appendix the reader can find the instructions sheets that the subjects were given at the beginning of each of the two listening experiments run during the project. When needed they were accompanied by oral explanations.

#### APPENDIX G. LISTENING EXPERIMENTS: INSTRUCTIONS SHEETS

# **Listening Experiment: Phase 1**

Welcome to our listening experiment. This is a description of the experiment you are going to participate in. Here we describe what will be your task and all relevant information regarding the experiment. Please take some minutes to read carefully and feel free to ask if you have any doubts or questions.

Please before we start we need some information about you:

- Name:
- Age:
- Occupation:

# **Description of the Listening Experiment**

This experiment is divided in 2 parts. In each part of the experiment you will be sitting on a chair in a listening room where sounds will be presented to you. You will have a control board with two buttons to deliver your answers.

### Part 1

In the first part of the experiment two sounds will be presented to you one after the other. Each of the buttons of the control board will be associated to one of the sounds (left-first, right-second) and they will light during their presentation. Your task will be to determine the sound which sounds *louder* to you. After the presentation of the pair of sounds, both buttons will light and you will be requested to press the button associated with the sound you perceived as the loudest. Loudness can be seen as a measure of the strength of the sound, or the quality of the sound which is altered by means of a volume control, i.e in a radio receiver. The process will be repeated several times for different pairs of sounds. It may be possible that sometimes the sounds seem equally loud to you, however you are requested to choose an answer, making your best choice.

Part 1 will be divided in two similar sessions with a break in between for you to relax and have a coffee if you wish so. The overall duration of this part will be around 30 minutes (with break included).

Following the end of Part 1, a break of approx. 10 minutes will take place. We will continue with the second part of the experiment described in what is following.

### Part 2

In the second part of the experiment, a sound will be presented to you and the volume of the sound will vary to some extent from presentation to presentation. The left button of the control board will light during the whole duration of this part of the experiment. Your task will be to press the left button *if* you *hear* a sound. In some occasions you will not hear anything, in this cases you should

not press any button. After some presentations a new sound will be presented to you and the same procedure will be repeated. Only the left button of the control board will be used in this part of the experiment. The duration of this part will be approximately 10 minutes.

Before we start with the experiment, note that there will be no right or wrong answers, we are interested in how *you* perceive the sound so we rely on your best judgment. The levels of the presented sounds present no risk of hearing damage.

Don't hesitate to ask any questions or comments about the experiment. Thank you in advance.

David & Carlos

## **Listening Experiment: Phase 2**

Welcome back to our listening experiment. The experiment you are going to participate in will allow us to investigate aspects of human hearing performance that have never been investigated before. The measurements have been custom-made for every one of you so we will take into account your personal hearing features. Results could be applied in many fields and will help to predict and avoid annoyance of many of today's daily life noises. Therefore your help is of great importance and we'd like to thank you for participating. Here we describe what will be your task and all relevant information regarding the experiment. Please take some minutes to read carefully and feel free to ask if you have any doubts or questions.

Please, before we start we need some information about you:

- Name:
- Age:
- Occupation:

# **Description of the Listening Experiment**

The experiment will be divided in two similar sessions with a break in between for you to relax. The duration of each session will be around 15 minutes. During the experiment you will be sitting on a chair in a listening room where sounds will be presented to you. You will have a control board to deliver your answers.

In order to perform the experiments you will be presented two kinds of sounds. During the whole experiment you will continuously hear a *low noise* in the room. At the same time a *low frequency tone* will be periodically presented (a low frequency tone is a sound similar to a very low musical note in a wind instrument). The noise will remain constant, but the tone will change in loudness in such a way that you will sometimes hear it, and sometimes not.

The left button of the control board will light during the whole duration of each measurement. Your task will be to press that button *if* you *clearly distinguish*the sound out of the noise. This doesn't mean that you press only when the tone is powerful, as you might be able to distinguish it even if its level is not very high. Just press the button when you *notice* the tone clearly.

Twelve short sets of measurements will be carried out before the break and 12 more after. But we will first begin with one demonstration that will clarify your task.

Before we start with the experiment, note that there will be no right or wrong answers, we are interested in if *you* perceive the sound so we rely on your best judgment. The levels of the presented sounds present no risk of hearing damage.

Don't hesitate to ask any questions or comments about the experiment. Thank you in advance.

David & Carlos