Article

Investigating the Performance of Gammatone Filters and Their Applicability to Design Cochlear Implant Processing System

Rumana Islam *[ID] and Mohammed Tarique [ID]

Department of Electrical and Computer Engineering, University of Science and Technology of Fujairah, Fujairah P.O. Box 2202, United Arab Emirates; m.tarique@ustf.ac.ae
* Correspondence: r.islam@ustf.ac.ae; Tel.: +971-50-973-6727

Abstract: Commercially available cochlear implants are designed to aid profoundly deaf people in understanding speech and environmental sounds. A typical cochlear implant uses a bank of bandpass filters to decompose an audio signal into a set of dynamic signals. These filters' critical center frequencies f_0 imitate the human cochlea's vibration patterns caused by audio signals. Gammatone filters (GTFs), with two unique characteristics: (a) an appropriate "pseudo resonant" frequency transfer function, mimicking the human cochlea, and (b) realizing efficient hardware implementation, could demonstrate them as unique candidates for cochlear implant design. Although GTFs have recently attracted considerable attention from researchers, a comprehensive exposition of GTFs is still absent in the literature. This paper starts by enumerating the impulse response of GTFs. Then, the magnitude spectrum, $|H(f)|$, and bandwidth, more specifically, the equivalent rectangular bandwidth (ERB) of GTFs, are derived. The simulation results suggested that optimally chosen filter parameters, e.g., critical center frequencies, f_0; temporal decay parameter, b; and order of the filter, n, can minimize the interference of the filter bank frequencies and very likely model the filter bandwidth (ERB), independent of $\frac{f_0}{b}$. Finally, these optimized filters are applied to delineate a filter bank for a cochlear implant design based on the Clarion processor model.

Keywords: basilar membrane; clarion processor; cochlear implant; equivalent rectangular bandwidth (ERB); gammatone filter (GTF) bank; pseudo-resonant frequency

1. Introduction

Speech communication is an integral part of our daily life. During speech production, a speaker encodes information into a continuously time-varying wave propagated through a medium [1]. The wave propagates from a speaker to a listener through the vibration of air particles. Finally, a listener perceives the information contained in the sound wave.

The human peripheral auditory system is one of the most critical components of speech perception. The human peripheral auditory system consists of three major parts [2]: the outer, middle, and inner ears, as shown in Figure 1. Sound enters the outer ear through the pinna, travels down to the auditory canal, and vibrates the eardrum. The middle ear, consisting of three bones, transports the vibration of the eardrum to the inner ear. The main component of the inner ear is the snail-shaped cochlea, a coiled tube filled with fluid [3]. Within the cochlear fluid, there exists a basilar membrane. The sound vibration at the eardrum ultimately generates a compressed sound wave in the cochlear fluid and causes a vertical vibration in the basilar membrane. The basilar membrane is mechanically tuned at different frequencies. It plays a vital role in distributing sound energy in frequencies along the cochlea's length, as shown in Figure 2. The wavelengths of audible sound can cover a wide range of scales. As depicted in this figure, the lower frequencies are located near the 'apex'. In contrast, the higher frequencies are at the far end, called 'base'. The low-frequency waves can have wavelengths of up to 17 m (20 Hz), while the highest frequencies can be as small as 1.7 cm (20,000 Hz).
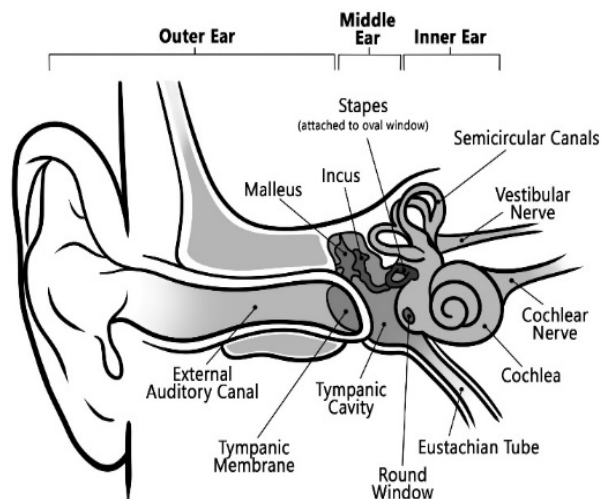
Figure 1. The components of human ear [1]. Sound enters the outer ear through the pinna and travels down to the middle and inner ears. Finally, it reaches the cochlea and vibrates the basilar membrane.
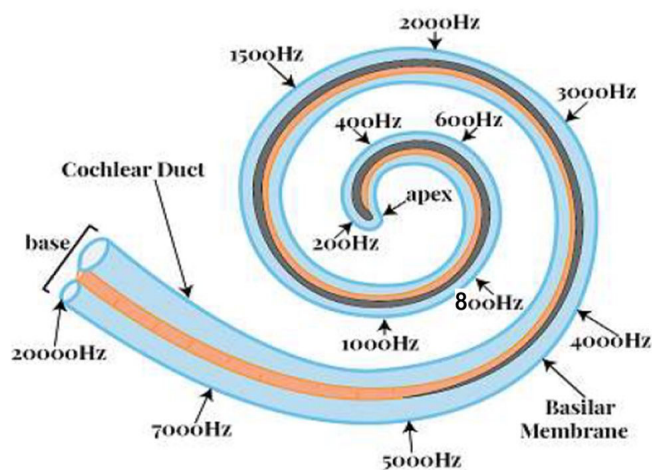

Figure 2. The tuning frequencies of the basilar membrane [4]. The basilar membrane is tuned to different frequencies from the apex to the base. The lower frequencies are near the apex, whereas the higher frequencies are near the base. The tuning frequency spacing also increases towards the base.

In conjunction with the basilar membrane, the hair cells translate mechanical information into neural information. If the hair cells are damaged, the auditory system cannot transform sound into neural impulses. The sound never reaches the brain because of the damaged hair cells. Many causes can damage hair cells, including diseases, congenital disorders, and specific drug treatments. Damaged hair cells can degenerate adjacent auditory neurons, too. Damaged neurons and hair cells can make a person profoundly deaf. However, recent research [5] has shown that the most common cause of deafness is the loss of hair cells rather than the loss of auditory neurons.

Hearing loss is the third most common health problem affecting the elderly population after heart disease and arthritis, according to some statistics [6]. Hair cells can be damaged over time, being open to continuous mechanical stress from environmental issues, including sounds. Various factors, e.g., aging, genetic defects, and ototoxic drugs, also cause additional risks of cell damage [7]. This damage can be mild to severe, causing even the death of hair cells. Unfortunately, human hair cells do not regenerate. The repair of hair cells is crucial for continued auditory function throughout life. Clinicians recommend several drugs to restore the proper functioning of hair cells when the damage is minor. Sudden hearing loss resulting from viral infection is medically treated with corticosteroids. Corticosteroids may also be used to reduce cochlear hair cell swelling and inflammation

when exposed to loud noise. However, clinical restoration of damaged hair cells from aging and genetic causes remains a research issue. Recently, gene therapy has been proven effective in restoring the functionality of damaged hair cells due to genetic factors in several animal models [8]. Even the perceptual quality of voice signals is greatly affected by hearing mechanisms [9]. A degraded voice can, in turn, be a biomarker of a human's health status, including both structural and neurological malfunctions in speech and hearing mechanisms [10–13].

A cochlear implant can play an important role here as it can excite the neurons through electrical stimulation to restore the hearing ability of a deaf person. The main idea is bypassing the standard hearing mechanism and electrically stimulating the auditory neurons.

Researchers have proved that frequency analysis performed concerning the cochlea can be modeled as a bandpass filter bank. Various filters have been proposed to implement bandpass filter banks. One of the earliest proposed filters is the rounded exponential function ('roex') [14]. The authors have shown that an exponential function can represent the auditory filter shape successfully. A novel reverse correlation technique has been introduced to better model the auditory filter [15]. Another function called 'revcor' has been introduced to define the impulse response of the peripheral auditory filters. Peculiarly, this function provides the impulse response of a sharp bandpass filter. Consequently, the GTF has been introduced in [16] to provide an analytic mathematical function approximating the 'revcor' function. Other researchers have further developed the GTF to make it suitable for practical design purposes [17,18]. One of the main merits of the GTF is its convenient mathematical form. Hence, its properties can be easily derived analytically compared to similar filters, including 'roex' filters. One of the pioneering works that investigated various properties of the GTF has been presented in [19]. The authors have defined the GTF as an infinite impulse response (IIR) filter in the time domain and described its provenance and some of its elementary properties. They also examined the behavior of the GTF in the frequency domain. They provided a way of calculating the parameters needed for a GTF to have a specified ERB. The authors provided an efficient digital implementation of the GTF on a general-purpose computer. A digital multiple-pass IIR filter technique has also been proposed to implement the GTF [20] for practical designs.

Recently, the GTF has drawn researchers' attention to sound event detection, speech signal processing, voice pathology detection, and speech recognition. In [21], the authors have proposed a GTF-based automatic speech recognition (ASR) technique. They demonstrated that GTFs are promising in terms of improving the robustness of ASR systems against noise compared to the Mel-Frequency Cepstral Coefficient (MFCC) and Perceptual Linear Prediction (PLP). GTF-based parametric filter banks have been proposed in [22] to detect speech. Three filter banks based on Mel, Gammatone, and Gaussian filters have been investigated in that work. The comparative investigation showed that the GTFs provided the highest speech detection accuracy compared to the Gaussian and Mel filters. A GTF-based sound event detection and localization (SEDL) system has been presented in [23]. The authors demonstrated that GTFs could boost the performance of state-of-the-art SEDL algorithms. In [24], the authors have applied the GTF to produce an image representation of sound signals for audio surveillance. They called this image representation a Cochleagram. The authors have shown that the proposed Cochleagram provided more noise robustness than cepstral features, namely Mel-Frequency Cepstral Coefficients and the spectrogram image feature (SIF). A learnable GTF bank is proposed to classify environmental sounds in [25]. The authors demonstrated that the learnable filter parameters of the GTFs could preserve the spectro-temporal domain features of environmental sound and can achieve high classification accuracy. In [26], the authors have shown that the GTF could enhance the performance of hearing aids. They concluded that the GTFs could provide a high hearing aid speech quality index (HASQI). A GTF-based speaker recognition system has been proposed in [27]. The authors have argued that conventional speaker recognition systems perform poorly under noisy conditions. They introduced a novel spectral feature called the Gammatone frequency cepstral coefficient (GFCC). They showed that this feature

captured speaker characteristics and performed substantially better than conventional spectral features under noisy conditions. The results showed significant performance improvements over related systems under a wide range of signal-to-noise ratios. In [28], the performances of the cochlear implants (CIs) have been investigated by using three different filters, namely GTF, DAPGF (Differentiated All-Pole GTF), OZGF (One-Zero GTF) and BUTF (Butterworth). Filter parameters, including the filter order (N), the filter quality factor (Q), and the number of channels (C) and their combinations, were tested using objective and subjective metrics in that work. The simulation results concluded that the Q and N parameters are crucial for designing cochlear implants.

Although the GTF has attracted considerable attention from researchers, as mentioned above, a comprehensive exposition on the GTF is still absent in the literature other than the work presented in [29]. That work has provided a tutorial introduction to the GTF without much detail. The main goals of this investigation are as follows:

a. Explore the effects of filter parameters: order n, carrier frequency f_c, carrier phase φ, temporal decay coefficient b, and Gammatone distribution function $r(t)$, on the impulse response $h(t)$ of the GTF.
b. Derive the transfer function of the GTF from the definition of the $h(t)$ by using the Fourier transform and its properties.
c. Investigate the effects of the above filter parameters on the transfer function $H(f)$ of the GTF.
d. Derive the expression for ERB of the GTF.
e. Design a filter bank using the GTF for a given pseudo-resonant f_c.
f. Demonstrate the application of the GTF in cochlear implant design.

The structure of this investigation is organized as follows: Section 2 explains the impulse response and spectrum of GTFs. Section 3 elaborates on the ERBs. Section 4 describes the possible application of GTFs in cochlear implant design. Section 5 addresses some underlying design issues and challenges. Finally, the paper concludes with a synopsis of key findings and explores possible routes for future research.

2. Impulse Response and Spectrum of GTFs

In [30], a Gammatone function has been used to model the basilar membrane displacement in the human ear. It has been further investigated, and it was shown that a GTF can be used to approximate responses recorded from the cochlear nucleus in cats [16]. In a similar work [31], a Gammatone function was used to model the impulse responses based on the auditory nerve fiber recordings in cats. Finally, the term "Gammatone filter" was introduced in [32], and its impulse response was defined as follows:

$$h(t) = ct^{n-1}e^{-2\pi bt}\cos(2\pi f_c t + \varphi)u(t) \tag{1}$$

where c is the proportionality constant, n is the filter order, b the temporal decay constant, f_c is the carrier frequency, φ is the carrier phase, and $u(t)$ is the unit step function. The expression of $h(t)$ can be broken down into two components, namely the carrier component and the Gamma distribution function.

Let us assume that the carrier component is denoted by

$$c(t) = \cos(2\pi f_c t + \varphi) \tag{2}$$

and the Gammatone distribution function is defined by

$$r(t) = t^{n-1}e^{-2\pi bt}u(t) \tag{3}$$

Hence, the impulse response of the GTF can be expressed as

$$h(t) = cs(t)r(t) \tag{4}$$

Figure 3 shows the plot for $h(t)$ of the GTF with its constituent components. In the plot, factor c has been set to $\frac{b^n}{(n-1)!}$ to make the integration under the curve of Gamma distribution equal to one. The other parameters are arbitrarily set to $b = 125$, and $f_c = 1000$ Hz. The filter order n of a GTF is an important design parameter. It controls the relative shape of the filter impulse response, as demonstrated in Figure 4. The relative shape becomes less skewed as the filter order n increases. The carrier phase, φ is also an important property that determines the relative position of the envelope.
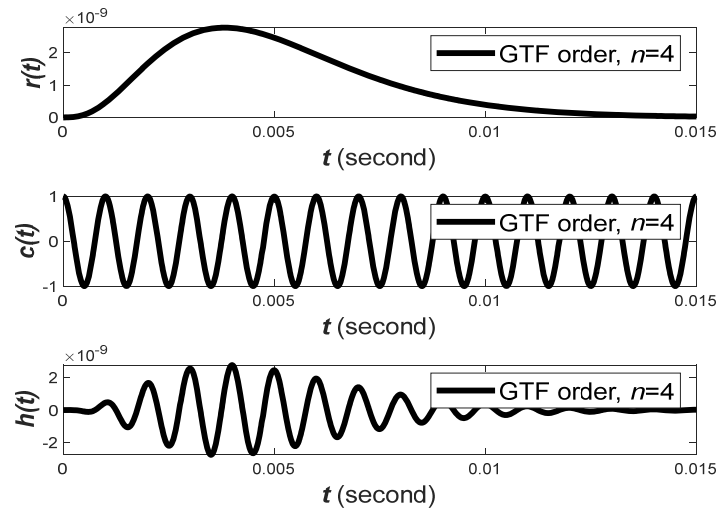

Figure 3. The components of a GTF impulse response: Gammatone distribution $r(t)$ (**top**), the carrier tone $c(t)$ (**middle**), and the impulse response $h(t)$ (**bottom**). The amplitude of the carrier tone is modulated according to the Gammatone distribution function.
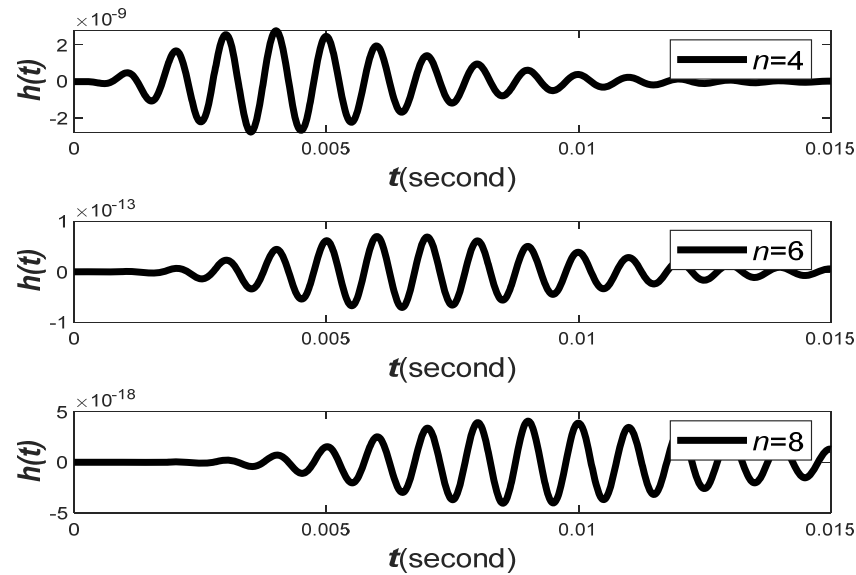

Figure 4. The Gammatone impulse response, $h(t)$ of the GTF with varying order, n. The relative shape becomes less skewed as the filter order n increases.

Observation 1. *When the filter order is higher, the impulse response of the GTF becomes less skewed and vice versa.*

The Fourier transform of the GTF's impulse response $h(t)$ will be derived to investigate its frequency domain behaviors. From the convolution property of the Fourier transform,

we know that if $z(t) = x(t)y(t)$, then $Z(f) = X(f) * Y(f)$ (Appendix A). By applying this property to (4), we can express $H(f)$ as

$$H(f) = c[S(f) * R(f)]\tag{5}$$

where $S(f)$ is the Fourier transform of $s(t)$, $R(f)$ is the Fourier transform of $r(t)$, and $H(f)$ is the Fourier transform of $h(t)$. We can find the Fourier transform of $r(t)$ by using the Fourier transform identity, $x(t) = e^{-a(t)}u(t) \leftrightarrow X(f) = \frac{1}{a+j2\pi f}$ (Appendix B). Substituting $a = 2\pi b$, we can find the Fourier transform of $x(t)$ as

$$x(t) = e^{-2\pi bt}u(t) \leftrightarrow X(f) = \frac{1}{2\pi b + j2\pi f}\tag{6}$$

From the property of the Fourier transform (Appendix A), we also know that if

$$x(t) \leftrightarrow X(f)\text{, then } t^m x(t) \leftrightarrow (-j2\pi)^{-m}\frac{d^m}{df^m}X(f).$$

By substituting $m = (n-1)$, we can express the Fourier transform of $r(t)$ as

$$r(t) = t^{n-1}e^{-j2\pi bt}u(t) \leftrightarrow R(f) = (-j2\pi)^{-(n-1)}\frac{d^{n-1}}{df^{n-1}}\left(\frac{1}{2\pi b + j2\pi f}\right)\tag{7}$$

If we substitute $n = 2$ in (7), we can find $te^{-j2\pi bt}u(t) \leftrightarrow \frac{1}{(2\pi b+j2\pi f)^2}$. Similarly, if we substitute $n = 3$, we can find, $t^2 e^{-j2\pi bt}u(t) \leftrightarrow \frac{(2)!}{(2\pi b+j2\pi f)^3}$. Proceeding in the same way, we can find the Fourier transform of $r(t)$ as

$$r(t) = t^{n-1}e^{-j2\pi bt}u(t) \leftrightarrow R(f) = \frac{(n-1)!}{(2\pi b + j2\pi f)^4}\tag{8}$$

The Fourier transform of $r(t)$ can be alternatively expressed as

$$R(f) = (n-1)!(2\pi b)^{-n}\left(1 + j\frac{f}{b}\right)^{-n}\tag{9}$$

Now, let us find the Fourier transform of the carrier signal, $c(t)$. The carrier signal is given by $c(t) = \cos(2\pi f_c t + \varphi)$, which can be alternatively expressed as $c(t) = \frac{1}{2}e^{j2\pi f_c t}e^{j\varphi} + \frac{1}{2}e^{-j2\pi f_c t}e^{-j\varphi}$. By using the Fourier transform identities $e^{j2\pi f_c} \leftrightarrow \delta(f - f_c)$, and $e^{-j2\pi f_0} \leftrightarrow \delta(f + f_0)$, the Fourier transform of $c(t)$ can be expressed as

$$C(f) = \frac{1}{2}e^{j\varphi}\delta(f - f_c) + \frac{1}{2}e^{-j\varphi}\delta(f - f_c)\tag{10}$$

Substituting the Fourier transform of $s(t)$ and $r(t)$ in (5), we can determine the expression of $H(f)$ as

$$H(f) = c(n-1)!(2\pi b)^{-n}\left(1 + j\frac{f}{b}\right)^{-n} * \left[\frac{1}{2}e^{j\varphi}\delta(f - f_c) + \frac{1}{2}e^{-j\varphi}\delta(f + f_c)\right]\tag{11}$$

By using the convolutional property of the delta Dirac function (Appendix C),

$$x(f) * \delta(f - f_c) = x(f - f_c),$$

We can find the final expression of the $H(f)$ as

$$H(f) = c(n-1)!(2\pi b)^{-n} \left[\frac{1}{2} e^{j\varphi} \left(1 + j\frac{f - f_c}{b} \right)^{-n} \right] + c(n-1)!(2\pi b)^{-n} \left[e^{-j\varphi} \left(1 + j\frac{f + f_c}{b} \right)^{-n} \right] \tag{12}$$

The plots for the $R(f)$ and $H(f)$ are shown in Figure 5. This figure shows that the $H(f)$ produces two copies of the $R(f)$ separated by the two times the carrier frequency, f_c.
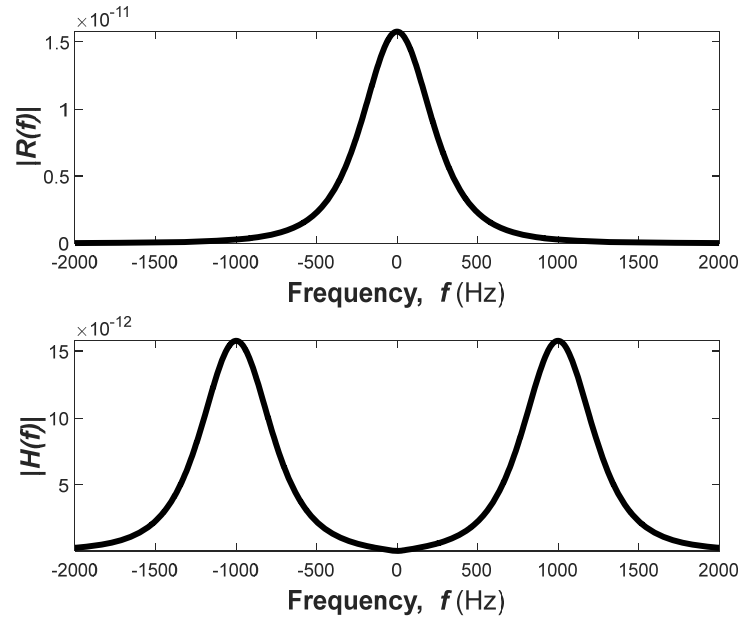

Figure 5. The magnitude spectrum of $r(t)$, i.e., $R(f)$ (**top**) and the magnitude spectrum of $h(t)$, i.e., $H(f)$ (**bottom**). $R(f)$ has a maximum value at, $f = 0$. On the other hand, $H(f)$ has two maxima, and they are located at $\pm f_c$.

Observation 2. *The Fourier transform of the Gamma distribution component of the GTF impulse response has one maximum value at $f_0 = 0$ Hz. However, the Fourier transform of the impulse response has two maxima, and the location of these two maxima depends on the carrier frequency. To avoid interference, these two frequency components shall be sufficiently separated by selecting a high carrier frequency.*

The Fourier transform of $h(t)$ can be expressed in terms of the Fourier transform of $r(t)$. From (5), we can write

$$H(f) = c \left[R(f) * \left\{ \frac{1}{2} e^{j\varphi} \delta(f - f_c) + \frac{1}{2} e^{-j\varphi} \delta(f + f_c) \right\} \right] \tag{13}$$

$$H(f) = \frac{1}{2} \left[\left\{ e^{j\varphi} R(f - f_c) + e^{-j\varphi} R(f + f_c) \right\} \right] \tag{14}$$

where $R(f)$ is given by (9) and can be alternatively expressed as

$$R(f) = \frac{(n-1)!}{[2\pi b + j2\pi f]^n} \tag{15}$$

The expression of $H(f)$ can be further simplified by assuming

$$k = \frac{c}{2}(n-1)!(2\pi b)^{-n} \tag{16}$$

$$P(f) = e^{j\varphi}[1 + j(f - f_c)/b]^{-n} \tag{17}$$

$$P^*(f) = e^{-j\varphi}[1 - j(f - f_c)/b]^{-n} \tag{18}$$

Replacing f with $-f$, we can modify (18) as

$$P^*(-f) = e^{-j\varphi}[1 + j(f + f_c)/b]^{-n} \tag{19}$$

Then, we can express $H(f)$ as

$$H(f) = k[P(f) + P^*(-f)] \tag{20}$$

The impulse response of the GTFs and their transfer functions are plotted in Figure 6 for varying parameters of f_c/b. The figure shows that, f_c/b is another critical parameter that affects the decaying behavior of the filter impulse response and filter transfer function.
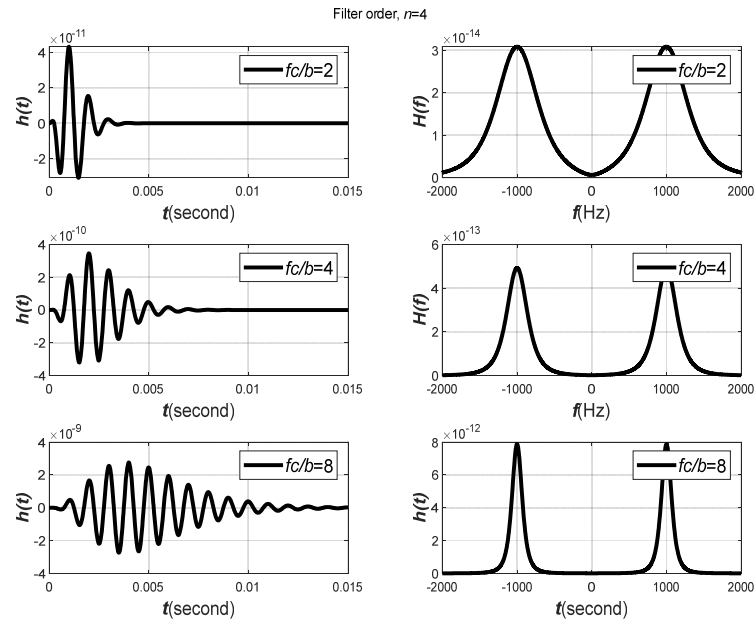

Figure 6. The filter impulse responses, $h(t)$, and their spectrums $H(f)$, with the varying magnitude of f_c/b. The value of $\frac{f_c}{b} \geq 4$ ensures that the two frequency components of GTF do not interfere with each other.

Observation 3. *When b is small (i.e., f_c/b is large), $h(t)$ will decay slowly. On the other hand, $|H(f)|$ will decay more rapidly and vice versa.*

Observation 4. *The larger the ratio f_c/b, the less the components of $KP(f)$ and $KP^*(f)$ overlap, and less interference between $KP(f)$ and $KP^*(f)$ will occur.*

In general, $H(f)$ can be expressed in terms of magnitude and phase spectrum as $H(f) = |H(f)|e^{-j\varnothing}$, where $|H(f)|$ is the magnitude of $H(f)$, and $\varnothing$ = phase spectrum of $H(f)$. The power spectrum of $H(f)$ is expressed by

$$\begin{aligned} |H(f)|^2 &= H(f)H^*(f) \\ &= k^2[P(f) + P^*(-f)][P^*(f) + P(-f)]] \\ &= k^2[P(f)P^*(f) + P^*(f)P^*(-f) + P(f)P(-f) + P^*(-f)P(-f)] \end{aligned} \tag{21}$$

By using the Fourier transform property (Appendix B), we can simplify the following expressions as $P(f)P^*(f) = |P(f)|^2$ and $P^*(-f)P(-f) = |P(-f)|^2$. Hence, $P(f)P(-f) + P^*(f)P^*(-f) = P(f)P(-f) + [P(f)P(-f)]^* = 2Re[P(f)P(-f)]$ and $|H(f)|^2$ can be expressed as

$$|H(f)|^2 = k^2\left[|P(f)|^2 + |P^*(-f)|^2 + 2Re\{P(f)P(-f)\}\right] \tag{22}$$

Assume $Q(f) = 1 + j(f - f_c)/b$. The $Q(f)$ can be expressed in terms of magnitude and phase as $Q(f) = |Q(f)|e^{j\theta_1(f)}$, where the magnitude is defined by $|Q(f)| = \sqrt{1 + \frac{(f-f_c)^2}{b^2}}$, and the phase spectrum is defined by $\theta_1(f) = \tan^{-1}\left[\frac{f-f_c}{b}\right]$, now, $P(f)$ can be expressed in terms of $Q(f)$, as

$$P(f) = e^{j\varphi}[Q(f)]^{-n} = e^{j\varphi}\left[|Q(f)|e^{j\theta_1(f)}\right]^{-n} \tag{23}$$

Substituting f with $-f$, we can find the expression of $Q(-f)$ as $Q(-f) = 1 + j(-f - f_c)$ $= 1 - j(f + f_c)/b$. Now, $Q(-f)$ can be expressed in terms of magnitude and phase as follows $Q(-f) = |Q(-f)|e^{j\theta_2(f)}$, where the magnitude of $Q(f)$ is defined as $|Q(-f)| = \sqrt{1 + \frac{(f-f_c)^2}{b^2}}$, and the phase $\theta_2(f) = \tan^{-1}\left[-\frac{f-f_c}{b}\right]$. Now, $P(-f)$ can be expressed in terms of $Q(-f)$ as

$$\begin{aligned} P(-f) &= e^{j\varphi}[Q(-f)]^{-n} = e^{j\varphi}\left[|Q(-f)|e^{j\theta_2(f)}\right]^{-n} \\ &= e^{j\varphi}|Q(-f)|^{-n}e^{-jn\theta_2(f)} \end{aligned} \tag{24}$$

We can express the terms presented in (22) as follows

$$\begin{aligned} |P(f)|^2 &= |Q(f)|^{-2n} = \sqrt{1 + (f - f_c)^2/b^2}^{-2n} \\ &= \left[1 + (f - f_c)^2/b^2\right]^{-n} \end{aligned} \tag{25}$$

Similarly,

$$|P(-f)|^2 = |Q(-f)|^{-2n} = \sqrt{1 + (f + f_c)^2/b^2}^{-2n} = \left[1 + \frac{(f + f_c)^2}{b^2}\right]^{-n} \tag{26}$$

$$P(f)P(-f) = e^{j\varphi}[Q(f)]^{-n}e^{j\varphi}[Q(-f)]^{-n} = [Q(f)Q(-f)]^{-n}e^{j2\varphi} \tag{27}$$

Let us find the expression of $Q(f)Q(-f)$ by

$$\begin{aligned} Q(f)Q(-f) &= [1 + j(f - f_c)/b][1 - j(f - f_c)/b] \\ &= 1 - \frac{j(f - f_c)}{b} - j^2\frac{f - f_c}{b}\frac{f + f_c}{b} + j\frac{f - f_c}{b} \\ &= 1 + j\frac{f - f_c - f + f_c}{b} + \frac{f^2 - f_c^2}{b^2} \\ &= 1 + \frac{f^2 - f_c^2}{b^2} - j2\frac{f}{b} \end{aligned} \tag{28}$$

The expression of $Q(f)Q(-f)$ can be expressed in terms of magnitude and phase as $Q(f)Q(-f) = |A(f)|e^{j\theta(f)}$, where the magnitude $|A(f)|$ can be expressed as $|A(f)| = \sqrt{\left[1 + \frac{f^2 - f_c^2}{b^2}\right]^2 + \left[\frac{2f_c}{b}\right]^2}$ and $\theta(f) = \tan^{-1}\left[\frac{-2f_c/b}{1 + \frac{f^2 - f_c^2}{b^2}}\right]$. Hence, $P(f)P(-f)$ can be expressed as

$$P(f)P(-f) = \left[\left(1 + \frac{f^2 - f_c^2}{b^2}\right)^2 + \left(\frac{2f_c}{b^2}\right)^2\right]^{-n/2} e^{j[2\varphi - n\theta(f)]} \tag{29}$$

Taking the real part of $P(f)P(-f)$, we can write

$$Re[P(f)P(-f)] = \left[\left(1 + \frac{f^2 - f_c^2}{b^2}\right)^2 + \left(\frac{2f_c}{b^2}\right)^2\right]^{-n/2} \cos\left[2\varphi - n\theta(f)\right] \quad (30)$$

Substituting all the derived terms in (22) we write the final expression of the power spectrum as

$$|H(f)|^2 = k^2 \left\{\left[\frac{1}{1 + \frac{(f-f_c)^2}{b^2}}\right]^n + \left[\frac{1}{1 + \frac{(f+f_c)^2}{b^2}}\right]^n + 2\left[\frac{1}{\sqrt{\left(1 + \frac{f^2 - f_c^2}{b^2}\right)^2 + \left(\frac{2f_c}{b}\right)^2}}\right]^n \cos\left[2\varphi - n\theta(f)\right]\right\} \quad (31)$$

$$= k^2 \left\{\left[\frac{b^2}{b^2 + (f - f_c)^2}\right]^n + \left[\frac{b^2}{b^2 + (f + f_c)^2}\right]^n + 2\left[\frac{b}{\sqrt{(b^2 + f^2 - f_c)^2 + (2f_c b)^2}}\right]^n \cos\left[2\varphi - n\theta(f)\right]\right\} \quad (32)$$

The power spectrum, $|H(f)|^2$ is plotted in Figure 7 for varying f_c/b. Based on the expression of $|H(f)|^2$ in (32) and the plot in Figure 7, we can make the following observations:

Observation 5. *When b is small, $h(t)$ will decay slowly; however, $|H(f)|^2$ will decay more rapidly.*

Observation 6. *Although $KP(f)$ and $KP^*(f)$ have their maximum at $\pm f_c$, the power spectrum $|H(f)|^2$ does not necessarily have a maximum at $\pm f_c$ Hz.*

Observation 7. *When $KP(f)$ and $KP^*(f)$ overlap significantly for small f_c/b, $|H(f)|^2$ has the character of a low pass filter with the peak at the origin.*

Observation 8. *As f_c/b is increased (for fixed order), the single peak splits and the maxima move outwards and eventually converges to $\pm f_c$.*
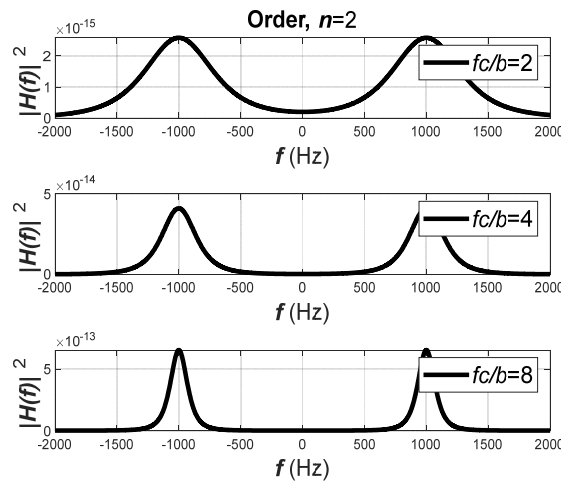

Figure 7. The plot of the power spectrum of the GTF, $|H(f)|^2$ with varying $\frac{f_c}{b}$ for $n = 2$. The plot shows that the power spectrum decays rapidly with a higher value of $\frac{f_c}{b}$. This faster decay reduces the interference between the frequency components of the GTF.

Since the purpose of the GTF in auditory modeling is to model a bandpass filter, the components $KP(f)$ and $KP^*(f)$ must be well separated, and it is required to make f_c/b large enough. In this case, we can simplify the expression of $H(f)$ as

$$H(f) = KP(f) \text{ when } f \geq 0 \quad (33)$$

$$= KP^*(-f) \text{ when } f < 0 \tag{34}$$

In addition, $P(f)P^*(-f) \approx 0$. The power spectrum expressed in (32) will be simplified as

$$|H(f)|^2 \approx k^2 \left[|P(f)|^2 \right] \text{ when } f \geq 0 \tag{35}$$

$$|H(f)|^2 \approx k^2 \left\{ \left[\frac{1}{1 + \frac{(f-f_c)^2}{b^2}} \right]^n \right\} \text{ when } f \geq 0 \tag{36}$$

Similarly, we can find

$$|H(f)|^2 \approx k^2 \left[|P^*(-f)|^2 \right] \text{ when } f < 0 \tag{37}$$

$$|H(f)|^2 \approx k^2 \left\{ \left[\frac{1}{1 + \frac{(f+f_c)^2}{b^2}} \right]^n \right\} \text{ when } f < 0 \tag{38}$$

For large f_c/b, the carrier phase φ does not have any effect on the maximum value of the power spectrum. For small f_c/b, the carrier phase φ influences where the maximum power spectrum occurs. Holdsworth shows that the optimum range of f_c/b should be $4 < f_c/b < 8$ for auditory modeling [15].

3. Equivalent Rectangular Bandwidth (ERB)

The ERB is a measure commonly used in psychoacoustics that approximates the bandwidth of the filters in human hearing. The ERB of a filter $H(f)$ is typically defined as the width of a rectangular filter whose height equals the maximum of the power spectrum of $H(f)$ and possesses the same amount of power. Based on this definition, the rectangular bandwidth H_{ERB} can be expressed as

$$H_{ERB} = \frac{\int_{-\infty}^{+\infty} |H(f)|^2 df}{2|H(f_0)|^2} \tag{39}$$

where $|H(f_0)|^2$ is the maximum value of the power spectrum, which occurs at $\pm f_0$. By using Perseval's theorem, the energy of a signal, $h(t)$, can be expressed as

$$E = \int_{-\infty}^{+\infty} |H(f)|^2 df = \int_{-\infty}^{+\infty} |h(t)|^2 dt \tag{40}$$

Hence, the expression of the rectangular bandwidth can be expressed as

$$H_{ERB} = \frac{\int_{-\infty}^{+\infty} |H(f)|^2 df}{2|H(f_0)|^2} = \frac{\int_{-\infty}^{+\infty} |h(t)|^2 dt}{2|H(f_0)|^2} \tag{41}$$

Let us assume $\check{h}(t) = |h(t)|^2$; hence, H_{ERB} can be expressed as

$$H_{ERB} = \frac{\int_{-\infty}^{+\infty} \check{h}(t) dt}{2|H(f_0)|^2} = \frac{\int_{-\infty}^{+\infty} |h(t)|^2 dt}{2|H(f_0)|^2} \tag{42}$$

From the definition of the Fourier transform of $\check{h}(t)$, we can write

$$\check{H}(f) = \int_{-\infty}^{+\infty} \check{h}(t) e^{-j2\pi ft} dt \tag{43}$$

The dc component of $\check{H}(f)$ can be found by substituting $f = 0$ in (43) and can be expressed as

$$\check{H}(0) = \int_{-\infty}^{+\infty} \check{h}(t)dt \tag{44}$$

By substituting $\int_{-\infty}^{+\infty} \check{h}(t)dt$ by $\check{H}(0)$ in (39), we can find an alternative expression for the equivalent rectangular bandwidth, H_{ERB} as

$$H_{ERB} = \frac{\int_{-\infty}^{+\infty} |H(f)|^2 df}{2|H(f_0)|^2} = \frac{\int_{-\infty}^{+\infty} |h(t)|^2 dt}{2|H(f_0)|^2} = \frac{\check{H}(0)}{2|H(f_0)|^2} \tag{45}$$

Squaring the (4), we can find the expression of $\check{h}(t)$ as

$$\check{h}(t) = [cr(t)s(t)]^2$$
$$= c^2 r^2(t) s^2(t) \tag{46}$$

This expression can be further simplified as

$$\check{h}(t) = c^2 \check{r}(t) \check{s}(t) \tag{47}$$

where

$$\check{r}(t) = r^2(t)$$
$$= \left[t^{n-1} e^{-2\pi b t} \right]^2$$
$$= t^{2n-2} e^{-j4\pi b t} u(t) \tag{48}$$

and

$$\check{s}(t) = s^2(t) = \cos^2(2\pi f_c t + \varphi) \tag{49}$$

By taking the Fourier transform of both sides of (47) and applying the convolution property of the Fourier transform, we can write

$$\check{H}(f) = c^2 \left[\check{R}(f) * \check{S}(f) \right] \tag{50}$$

where $\check{R}(f)$ = Fourier transform of $\check{r}(t)$, and $\check{S}(f)$ = Fourier transform of $\check{s}(t)$. Now, we need to find the Fourier transform of $\check{r}(t)$ and $\check{s}(t)$ and substitute in (50). Applying the Fourier transform property $t^m e^{-at} u(t) \leftrightarrow m!(a + j2\pi f)^{-(m+1)}$. Substituting $m = 2n - 2$, and $a = 4\pi b$, we can find the $\check{r}(t) = t^{2n-2} e^{-4\pi b t} u(t) \leftrightarrow \check{R}(f) = (2n-2)!(4\pi b + j2\pi f)^{-(2n-1)}$. This expression can be further simplified as

$$\check{R}(f) = (2n-2)! \left[4\pi b \left(1 + j\frac{2\pi f}{4\pi b} \right) \right]^{-(2n-1)}$$

$$\check{R}(f) = (2n-2)!(4\pi b)^{-(2n-1)} \left[1 + j\frac{f}{2b} \right]^{-(2n-1)} \tag{51}$$

Now, the expression of $\check{s}(t)$ can be simplified as

$$\check{s}(t) = \cos^2(2\pi f_c t + \varphi)$$
$$= \tfrac{1}{2}[1 + \cos(4\pi f_c t + 2\varphi)]$$
$$= \tfrac{1}{2} + \tfrac{1}{2} \left[\frac{e^{j(4\pi f_c t + 2\varphi)} + e^{-j(4\pi f_c t + 2\varphi)}}{2} \right]$$
$$= \tfrac{1}{2} + \frac{e^{j2\varphi}}{4} e^{j2(2\pi f_c t)} + \frac{e^{-j2\varphi}}{4} e^{-j2(2\pi f_c t)}$$

Taking the Fourier transform, we can express the Fourier transform of $\check{s}(t)$ as

$$\check{S}(f) = \frac{1}{2}\delta(f) + \frac{1}{4}e^{j2\varphi}\delta(f - 2f_c) + \frac{1}{4}e^{-j\varphi}\delta(f + 2f_c) \tag{52}$$

Substituting the value of $\check{R}(f)$ and $\check{S}(f)$ in (50), we find the expression of $\check{H}(f)$ as

$$\check{H}(f) = c^2(2n-2)!(4\pi b)^{-(2n-1)}\left[1 + j\frac{f}{2b}\right]^{-(2n-1)} * \left[\frac{1}{2}\delta(f) + \frac{1}{4}e^{j2\varphi}\delta(f - 2f_c) + \frac{1}{4}e^{-j2\varphi}\delta(f + 2f_c)\right]$$

By using the convolution property of the delta dirac function, the above expression can be further simplified as

$$\check{H}(f) = c^2(2n-2)!(4\pi b)^{-(2n-1)}\left\{\left[\frac{1}{2}\left(1 + j\frac{f}{2b}\right)\right]^{-(2n-1)} + \frac{1}{4}e^{j2\varphi}\left[1 + j\frac{(f-2f_c)}{2b}\right]^{-(2n-1)}\right.$$
$$\left. + \frac{1}{4}e^{-j2\varphi}\left[1 + j\frac{(f+2f_c)}{2b}\right]^{-(2n-1)}\right\} \tag{53}$$

By substituting $f = 0$, we can find the expression of $\check{H}(0)$ as

$$\check{H}(0) = c^2(2n-2)!(4\pi b)^{-(2n-1)}\left\{\begin{array}{l}\frac{1}{2} + \frac{1}{4}e^{j\varphi}\left[1 - j\frac{f_c}{b}\right]^{-(2n-1)} \\ + \frac{1}{4}e^{-j\varphi}\left[1 + j\frac{f_c}{b}\right]^{-(2n-1)}\end{array}\right\} \tag{54}$$

Let us assume $X = \frac{e^{j2\varphi}}{4}\left[1 - j\frac{f_c}{b}\right]$ and $X^* = \frac{e^{-j2\varphi}}{4}\left[1 + j\frac{f_c}{b}\right]$. We can express X as

$$X = \frac{e^{j2\varphi}}{4}\left[1 - j\frac{f_c}{b}\right], \text{ substituting } \theta_1(f) = \tan^{-1}\left[-\frac{f_c}{b}\right]$$
$$= \frac{e^{j2\varphi}}{4}\left[\sqrt{1 + \frac{f_c^2}{b^2}}\right]e^{\theta_1(f)} \tag{55}$$

$$= \frac{1}{4}\left[\sqrt{1 + \frac{f_c^2}{b^2}}\right]e^{\theta_1(f)+j2\varphi}$$
$$= \frac{1}{4}\left[\sqrt{1 + \frac{f_c^2}{b^2}}\right]e^{j[2\varphi+\theta_1(f)]} \tag{56}$$

Similarly, it can be proved that X^* can be expressed as

$$X^* = \frac{1}{4}\left[\sqrt{1 + \frac{f_c^2}{b^2}}\right]e^{j[-2\varphi+\theta_2(f)]}$$

where $\theta_2(f) = \tan^{-1}\left[\frac{f_c}{b}\right]$. However, $\theta_2(f) = \pi - \theta_1(f)$. Hence, X^* can be expressed as

$$X^* = \frac{1}{4}\left[\sqrt{1 + \frac{f_c^2}{b^2}}\right]e^{-j[2\varphi+\theta_1(f)]}$$

By using the complex variable identity $X^m + (X^*)^m = 2Re[X^m] = 2|X|^m\cos(m\theta)$. We can write the expression of $\check{H}(0)$ as

$$\check{H}(0) = c^2(2n-2)!(4\pi b)^{-(2n-1)}\left\{\begin{array}{l}\frac{2}{4}\left(1 + \frac{f_c^2}{b^2}\right)^{-\frac{2n-1}{2}} \\ \cos[2\varphi - (2n-1)\theta_1(f)] + \frac{1}{2}\end{array}\right\} \tag{57}$$

Substituting $f = f_0$ in (32) we can find the expression of $|H(f_0)|^2$ as

$$|H(f_0)|^2 = k^2 \left\{ \left[\frac{1}{1 + \frac{(f_0 - f_c)^2}{b^2}} \right]^n + \left[\frac{1}{1 + \frac{(f_0 + f_c)^2}{b^2}} \right]^n + 2 \left[\frac{1}{\sqrt{\left(1 + \frac{f_0^2 - f_c^2}{b^2}\right)^2 + \left(\frac{2f_c}{b}\right)^2}} \right]^n \cos\left[2\varphi - n\theta(f)\right] \right\} \quad (58)$$

Substituting $\check{H}(0)$ and $\left| H\left(f_{0)}\right| \right.$ in (45), we can find the final expression of the H_{ERB} as

$$H_{ERB} = \frac{\check{H}(0)}{2|H(f_0)|^2} = \frac{c^2(2n-2)!(4\pi b)^{-(2n-1)} \left\{ \frac{2}{4}\left(1 + \frac{f_c^2}{b^2}\right)^{-(2n-1)/2} \cos\left[2\varphi - (2n-1)\theta_1(f)\right] + \frac{1}{2} \right\}}{k^2 \left\{ \left[\frac{1}{1 + \frac{(f_0 - f_c)^2}{b^2}} \right]^n + \left[\frac{1}{1 + \frac{(f_0 + f_c)^2}{b^2}} \right]^n + 2 \left[\frac{1}{\sqrt{\left(1 + \frac{f_0^2 - f_c^2}{b^2}\right)^2 + \left(\frac{2f_c}{b}\right)^2}} \right]^n \cos\left[2\varphi - n\theta(f)\right] \right\}} \quad (59)$$

According to the definition of the H_{ERB}, we substitute $f_c = f_0$ in (59) and can find the final expression of H_{ERB} as

$$H_{ERB} = \frac{c^2(2n-2)!(4\pi b)^{-(2n-1)} \left\{ \frac{2}{4}\left(1 + \frac{f_c^2}{b^2}\right)^{-(2n-1)/2} \cos\left[2\varphi - (2n-1)\theta_1(f)\right] + \frac{1}{2} \right\}}{2k^2 \left\{ 1 + \left[\frac{1}{1 + \frac{(2f_c)^2}{b^2}} \right]^n + 2 \left[\frac{1}{\sqrt{1 + \left(\frac{2f_c}{b}\right)^2}} \right]^n \cos\left[2\varphi - n\theta(f)\right] \right\}} \quad (60)$$

Defining two more design parameters η and μ as

$$\eta = \frac{c^2(2n-2)!(4\pi b)^{-(2n-1)}}{2k^2} \quad (61)$$

$$\mu = \frac{\left\{ \frac{2}{4}\left(1 + \frac{f_c^2}{b^2}\right)^{-(2n-1)/2} \cos\left[2\varphi - (2n-1)\theta_1(f)\right] + \frac{1}{2} \right\}}{\left\{ 1 + \left[\frac{1}{1 + \frac{(2f_c)^2}{b^2}} \right]^n + 2 \left[\frac{1}{\sqrt{1 + \left(\frac{2f_c}{b}\right)^2}} \right]^n \cos\left[2\varphi - n\theta(f)\right] \right\}} \quad (62)$$

Substituting $k = \frac{c}{2}(n-1)!(2\pi b)^{-n}$ (61), we can find the final expression for η as

$$\begin{aligned} \eta &= \frac{c^2(2n-2)!(4\pi b)^{-(2n-1)}}{2\left(\frac{c}{2}(n-1)!(2\pi b)^{-n}\right)^2} \\ &= \frac{c^2(2n-2)!(4\pi b)^{-(2n-1)}}{2\frac{c^2}{2^2}\left[(n-1)!(2\pi b)\right]^2} \\ &= \frac{(2n-2)!}{[(n-1)!]^2}(\pi b)2^{3-2n} \end{aligned} \quad (63)$$

The variation in H_{ERB} with $\frac{f_0}{b}$ is plotted in Figure 8. The maximum value of μ is $1/2$ when $\frac{f_c}{b}$ is sufficiently large. With this value of μ, $H_{ERB} = \eta\mu$ becomes approximately

$$H_{ERB} \approx \eta/2$$

$$H_{ERB} \approx \frac{(2n-2)!(\pi b)2^{2-2n}}{[(n-1)!]^2} \quad (64)$$

Hence, H_{ERB} becomes proportional to b and independent of f_0. As mentioned above, the resonant frequencies along the basilar membrane vary from 20 Hz at the apex to 20,000 Hz at the base. Hence, it is essential to make the bandwidth of the filter independent

of the carrier frequency, f_0. This makes the GTF a unique candidate for auditory modeling in cochlear implants.
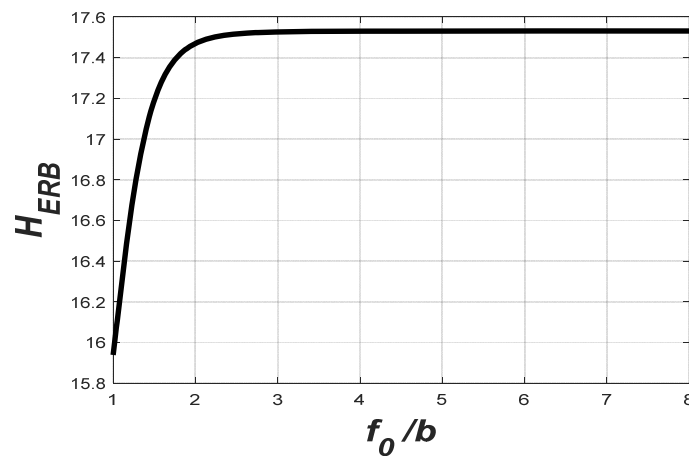

Figure 8. The variation in H_{ERB} with $\frac{f_0}{b}$. This figure shows that H_{ERB} remains independent of $\frac{f_0}{b}$, for $\frac{f_0}{b} > 3$, and $\mu \approx 1/2$.

4. Application of GTFs in Cochlear Implant Design

Almost 40 years ago, researchers initiated the restoration of normal hearing in deaf people via electrical stimulation of the auditory nerve [33]. Since then, they have been investigating different techniques for delivering electrical stimuli to the auditory nerve so that profoundly deaf people understand normal speech. Advances in signal processing largely contribute to the continuous and steady improvement of cochlear implant users. Several review papers on this topic have been published [34–37]. Recently, prosthetic devices [38–41], called cochlear implants [42,43], can be implanted in the inner ear to restore the partial hearing ability of profoundly deaf people. By using cochlear implants, some individuals can now communicate like normal people.

Initially, single-channel implants were tested in human subjects in the early 1970s [44–46]. Single-channel implants provide electrical stimulation at a single site in the cochlea using a single electrode. These implants are of interest because of their simplicity in design, as they do not require much hardware. The first experiments were discouraging as the patients reported unintelligible perception of speech. Later, related research works have been focused on multi-channel implants. Unlike single-channel implants, multi-channel implants provide electrical stimulation at multiple sites in the cochlea using an array of electrodes. An electrode array is used to stimulate different auditory nerve fibers at various places in the cochlea. Different electrodes are stimulated depending on the frequency of the signal. Electrodes near the base of the cochlea are stimulated with high-frequency signals, while electrodes near the apex are stimulated with low-frequency signals, as shown in Figure 2. In multi-channel cochlear implants, signal processing is the most important component [47], and a bank of bandpass filters is used to split the input sound signals into a set of parallel signals [47]. In this work, we are proposing to use GTFs instead.

To investigate the application of GTFs in cochlear implantation, a commercially available cochlear implant processor model called Clarion [33,42], as shown in Figure 9, is used in this work. The Clarion processor uses a microphone, worn at ear level, to capture the incoming sound. The sound is digitized and analyzed by a processor. The processor divides the signal into several channels based on frequency and translates the information in each channel into instructions that are transmitted to and control an implanted receiver that drives the implanted electrode array. The array of electrodes consists of 6–22 intra-cochlear electrodes distributed along the length of the cochlea. Stimuli delivered to an electrode preferentially excite the nerve fibers nearby.
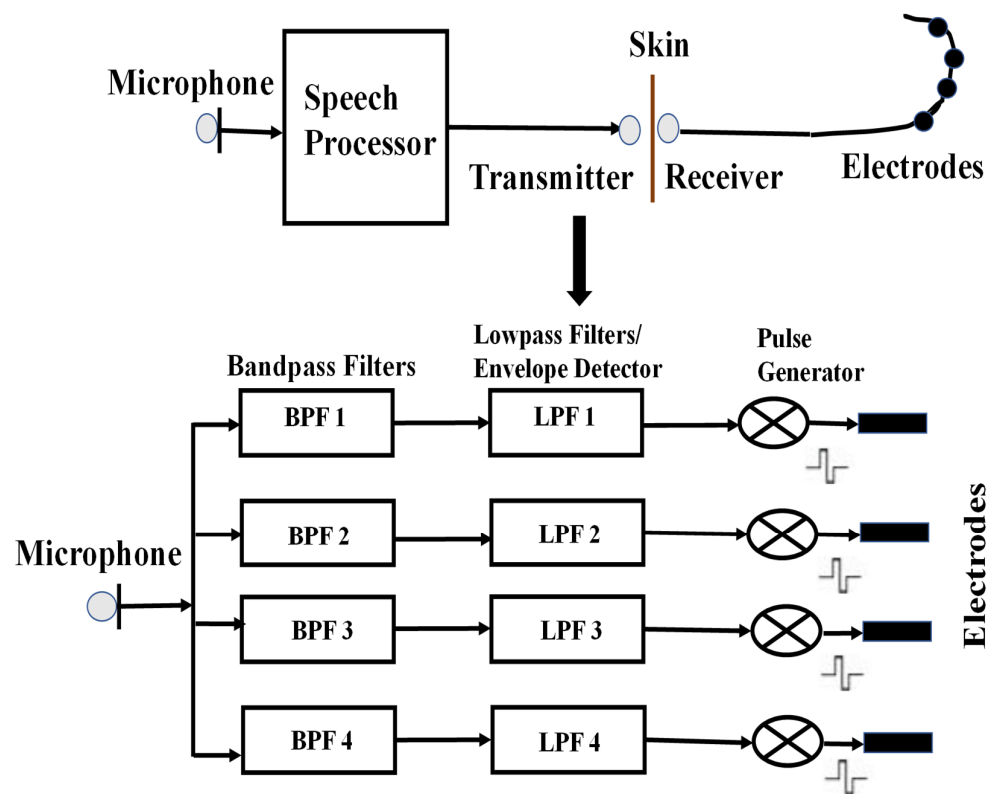
Figure 9. The signal processing steps in the Clarion processor. The main components of a cochlear implant are the microphone, speech processors and the electrodes. The speech processor consists of a bank of bandpass filters that split the incoming sounds into parallel components that are subsequently processed by a bank of lowpass filters. The pulse generator produces non-linearly mapped pulses to excite the electrodes.

The proposed model slightly varies from the Clarion processor. In the proposed model, the audio signal is first pre-emphasized [48] to boost the higher frequency components, as shown in Figure 10. The signal is then divided into channels by a set of GTFs instead of bandpass filters that are used in the Clarion processor. The main reason is that the bandpass filters do not represent the way the human auditory system responds to sounds. In addition, the hardware implementation of the bandpass filters is not as straightforward as the GTF. The next stage in the implant's processing is the extraction of the envelope of the signal from each channel. This is achieved by rectification and lowpass filtering. Full-wave rectification is used in this model. A dc component is introduced during the rectification methods, and the harmonics that typically fall above the Nyquist frequency are aliased to lower frequencies. The rectified signal is lowpass filtered using a 16th-order moving-average filter [49]. In a cochlear implant, the amplitude envelopes of each channel modulate a biphasic pulse train, which has a repetition rate of 800 to 4000 pulses per second (pps). Each modulated pulse train is delivered to a separate electrode, emulating the tonotopic arrangement of the cochlea. The GTFs are designed to cover a range of frequencies representing the basilar membrane [50–54]. In this work, these filters were designed based on the specifications mentioned in [21]. The center frequency and the bandwidth of these eight GTFs are listed in Table 1, and the magnitude spectrum of the GTF bank is shown in Figure 11. Those GTFs perform spectral analysis and convert an acoustic wave into a multichannel representation by mimicking the basilar membrane motion [55]. These GTFs have been designed in a way that $\frac{f_0}{b} = 4$, as mentioned above. The filter order n was set to 4.0. The shape of the magnitude characteristic of the GTFs with order 4 is very similar to that of the *roex* function [56] that is commonly used to represent the magnitude response of the human auditory filter [57].
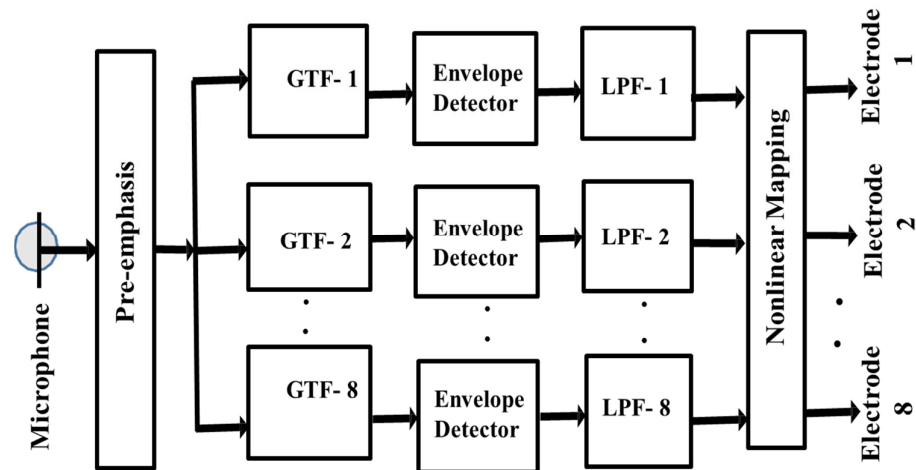
Figure 10. The proposed model uses GTF. The sound signal is pre-emphasized and is split into eight parallel signals by the GTF bank. The envelope detectors extract the signal envelops, which pass through the lowpass filter. Non-linear mapping is performed to reduce the interference among the electrodes.

Table 1. The center frequency and the bandwidth of the GTFs.

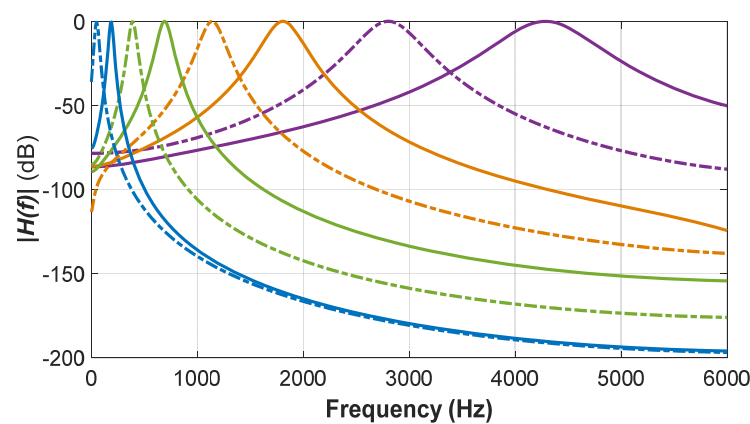| Bandwidth (Hz)	Center Frequency (Hz)
158	50
173	186
276	389
478	690
788	1139
1249	1807
1936	2802
2960	4282

Figure 11. The magnitude spectrum of the GTF. The tuning frequencies and the bandwidth of the filters are determined by the specifications mentioned in Table 1. As depicted in this figure, the filter's bandwidth increases with the tuning frequency. The eight filters are identified with different colors and line styles.

5. Design Issues and Challenges

Despite the impressive ability of cochlear implants to improve sound audibility and speech understanding in profoundly deaf people, several significant challenges remain to address to maximize the benefits of this device. One major challenge is the substantial variability of audio perception among different gender groups, demographics, and ages. Research is still ongoing to correlate neural and cognitive function in cochlear implant users.

There is a need to devise simple assessment measures to evaluate the perceptional outcomes of cochlear implant users. Poorer frequency discrimination abilities and neural deficits resulting from long-term deafness pose extra challenges to audio perception for cochlear implant users [58,59]. Rather than a physiologic point of view, some technological issues also need future investigation. A healthy cochlea transmits temporal-frequency information of audible sounds through around 3000 inner hair cells, but an implanted version could deliver a degraded version of such information resulting from signal processing (e.g., signal compression, bandpass filtering, temporal envelope extraction) and only a small number (up to eight) of electrodes in this design. As mentioned above, the number of spectral channels used for most CI users is likely less than eight due to factors including channel interactions. Signal processing also removes delicate temporal structures that may hinder normal hearing regarding melody contents [60].

While cochlear implants have proven to be beneficial for many individuals with profound hearing loss, there are some potential drawbacks to consider:

Cost: The potential physiological design challenges mentioned above could make cochlear implants expensive, and the cost may not always be fully covered by insurance. This financial aspect can be a barrier for some individuals.

Surgical Risks: Though the implantation process involves a mild surgery, like any surgical procedure, there are inherent risks like infections, bleeding, and issues related to anesthesia.

Learning Curve: Adjusting to hearing with a cochlear implant requires time and effort. Some individuals may find the initial period challenging as they learn to interpret the new auditory signals.

Maintenance and Upkeep: Cochlear implants require ongoing maintenance, including regular checks and adjustments. The external components also need to be cared for to ensure optimal functioning.

It is essential for individuals considering cochlear implants to discuss these aspects with their healthcare providers and audiologists. Despite these considerations, many people with cochlear implants experience significant improvements in their ability to hear and communicate.

6. Conclusions

The design and optimization of GTFs are proposed for cochlear implants. The complex spectrum and equivalent rectangular bandwidth of GTFs have been derived and investigated. One of the key findings is that the frequency domain behavior of the GTF strongly depends on $\frac{f_c}{b}$. Additionally, the optimal choice of a design parameter, i.e., $f_c/b \geq 4$ could minimize the interferences of the filter frequency components. A smaller value of b can cause the power spectrum, $|H(f)|^2$ to decay faster and hence can reduce the interference of the filter output. It is concluded that the ERB of the GTF has a strong influence on b, and becomes independent of f_c for $f_c/b \geq 4$.

This research provides a theoretical and simulation-based analysis of GTFs to explore their applicability for cochlear implants. The details of the hardware implementation are yet to be investigated. The presented filter bank is designed with eight filters. However, the number of filters (and hence the number of electrodes) to optimize the cochlear implants is still an open issue. Reducing the interferences among the electrodes also needs future investigation.

Author Contributions: Conceptualization, analysis, mathematical modeling, simulation, and manuscript writing—R.I.; mathematical modeling, editing, and manuscript writing—M.T. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A Properties of Fourier Transform

Operation	$x(t)$	$X(f)$
Time Convolution	$x(t) * y(t)$	$X(f)Y(f)$
Multiplication by t	$t^m x(t)$	$(-j2\pi)^{-m}\frac{d^m}{df^m}X(f).$
Modulation	$x(t)\cos 2\pi f_0 t$	$0.5X(f - f_0) + 0.5X(f + f_0)$
Frequency Convolution	$x(t)y(t)$	$X(f)Y(f)$
Time Shifting	$x(t - t_0)$	$X(f)e^{-j2\pi f t_0}$
Frequency Shifting	$x(t)e^{j2\pi f_0 t}$	$X(f - f_0)$

Appendix B Short Table of Fourier Transform

$x(t)$	$X(f)$
$e^{-a(t)}u(t)$	$\frac{1}{a+j2\pi f}.$
$e^{j2\pi f_0 t}$	$\delta(f - f_0)$
$e^{-j2\pi f_0 t}$	$\delta(f + f_0)$
$\cos 2\pi f_0 t$	$0.5\delta(f - f_0) + 0.5\delta(f + f_0)$
$\cos(2\pi f_0 t + \varphi)$	$0.5e^{j\varphi}\delta(f - f_0) + 0.5e^{-j\varphi}\delta(f - f_0)$

Appendix C Some Important Formulae

$$X(f) = \int_{-\infty}^{+\infty} x(t)e^{-j2\pi ft}dt$$

$$x(t) = \int_{-\infty}^{+\infty} X(f)e^{+j2\pi ft}df$$

$$x(t) * \delta(t - t_0) = x(t - t_0)$$

$$E = \int_{-\infty}^{+\infty} |x(t)|^2 dt = \int_{-\infty}^{+\infty} |X(f)|^2 df$$

$$X(f)X(f)^* = |X(f)|^2$$

References

1. Rabiner, L.R.; Schafer, R.W. Auditory, and Speech Perception. In *Theory and Applications of Digital Speech Processing*, 1st ed.; Prentice-Hall: Upper Saddle River, NJ, USA, 2011; pp. 138–145.
2. Chittka, L.; Brockmann, A. Perception Space—The Final Frontier. *PLoS Biol.* **2015**, *3*, 564–568. [CrossRef]
3. Quateri, T.E. Production and Classification of Speech Sounds. In *Discrete-Time Speech Signal Processing: Principles and Practices*; Prentice-Hall: Upper Saddle River, NJ, USA, 2001; pp. 72–76.
4. Islam, R.; Abdel-Raheem, E.; Tarique, M. A Novel Pathological Voice Identification Technique through Simulated Cochlear Implant Processing Systems. *Appl. Sci.* **2022**, *12*, 2398. [CrossRef]
5. Hinojosa, R.; Marion, M. Histopathology of profound sensorineural deafness. *Ann. N. Y. Acad. Sci.* **1983**, *405*, 459–484. [CrossRef]
6. Blackwell, D.L.; Lucas, J.W.; Clarke, T.C. *Summary Health Statistics for US Adults: National Health Interview Survey*; Vital and Health Statistics, Series: 10; Number 260; National Health Survey; National Library of Medicine: Bethesda, MD, USA, 2014; pp. 1–161.
7. Wagner, E.L.; Shin, J.B. Mechanisms of Hair Cell Damage and Repair. *Trends Neurosci.* **2019**, *42*, 414–424. [CrossRef]
8. Taiber, S.; Cohen, R.; Yizhar-Barnea, O.; Sprinzak, D.; Holt, J.R.; Avraham, K.B. Neonatal AAV gene therapy rescues hearing in a mouse model of SYNE4 deafness. *EMBO Mol. Med.* **2021**, *13*, e13259. [CrossRef] [PubMed]
9. Antje, H.; Helen, H.; Melanie, A.F. The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests. *Front. Psychol.* **2015**, *6*, 782. [CrossRef]
10. Islam, R.; Tarique, M.; Abdel-Raheem, E. A Survey on Signal Processing Based Pathological Voice Detection Techniques. *IEEE Access* **2020**, *8*, 66749–66776. [CrossRef]
11. Islam, R.; Tarique, M. A novel convolutional neural network based dysphonic voice detection algorithm using chromagram. *Int. J. Electr. Comput. Eng.* **2022**, *12*, 5511–5518. [CrossRef]

12. Islam, R.; Abdel-Raheem, E.; Tarique, M. A study of using cough sounds and deep neural networks for the early detection of COVID-19. *Biomedical. Eng. Adv.* **2022**, *3*, 100025. [CrossRef] [PubMed]
13. Islam, R.; Abdel-Raheem, E.; Tarique, M. Voiced Features and Artificial Neural Networks to Diagnose Parkinson's Disease Patients. In Proceedings of the International Conference on Electrical and Computing Technologies and Applications, Ras Al Khaimah, UAE, 23–25 November 2022; pp. 132–136. [CrossRef]
14. Patterson, R.D.; Moore, B.J.C. Auditory filters and excitation patterns as representations of frequency resolution. In *Frequency Selectivity in Hearing*; Moore, B.C.J., Ed.; Academic Press: Cambridge, MA, USA, 1986; pp. 123–177.
15. Boer, E.D.; Kuyper, P. Triggered Correlation. *IEEE Trans. Biomed. Eng.* **1968**, *BME-15*, 169–179. [CrossRef]
16. Johannesma, P.I.M. The pre-response stimulus ensemble of neurons in the cochlear nucleus. In Proceedings of the Symposium on Hearing Theory, Eindhoven, The Netherlands, 22–23 June 1972; pp. 58–69.
17. Boer, E.D.; Jongh, H.R.D. On cochlear encoding: Potentialities and limitations of the reverse-correlation technique. *J. Acoust. Soc. Am.* **1978**, *63*, 115–135. [CrossRef] [PubMed]
18. Boer, E.D.; Kruidenier, C. On ringing limits of the auditory periphery. *Biol. Cybern.* **1990**, *63*, 433–442. [CrossRef] [PubMed]
19. Holdsworth, J.; Patterson, R.; Nimmo-Smith, I.; Rice, P. Implementing a Gammatone Filter Bank. In *SVOS Final Report Part A: The Auditory Filterbank*; MRC Applied Psychology Unit: Cambridge, UK, 1988.
20. Patterson, R.; Nimmo-Smith, I.; Holdsworth, J.; Rice, P. The Auditory Filterbank. In *SVOS Final Report. Part A*; MRC Applied Psychology Unit: Cambridge, UK, 1988.
21. Qi, J.; Wang, D.; Jiang, Y.; Liu, R. Auditory features based on Gammatone filters for robust speech recognition. In Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Beijing, China, 19–23 May 2013; pp. 305–308. [CrossRef]
22. Cai, X.; Ko, S. Development of Parametric Filter Banks for Sound Feature Extraction. *IEEE Access* **2023**, *11*, 109856–109867. [CrossRef]
23. Jacome, K.G.R.; Grijalva, F.L.; Masiero, B.S. Sound Events Localization and Detection Using Bio-Inspired Gammatone Filters and Temporal Convolutional Neural Networks. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2023**, *31*, 2314–2324. [CrossRef]
24. Sharan, R.V.; Moir, T.J. Subband Time-Frequency Image Texture Features for Robust Audio Surveillance. *IEEE Trans. Inf. Secur.* **2015**, *10*, 2605–2615. [CrossRef]
25. Park, H.; Yoo, C.D. CNN-Based Learnable Gammatone Filterbank and Equal-Loudness Normalization for Environmental Sound Classification. *IEEE Signal Process. Lett.* **2020**, *27*, 411–415. [CrossRef]
26. Salehi, H.; Suelzle, D.; Folkeard, P.; Parsa, V. Learning-Based Reference-Free Speech Quality Measures for Hearing Aid Applications. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2018**, *26*, 2277–2288. [CrossRef]
27. Zhao, X.; Shao, Y.; Wang, D. CASA-Based Robust Speaker Identification. *IEEE Trans. Audio Speech Lang. Process.* **2012**, *20*, 1608–1616. [CrossRef]
28. Cosentino, S.; Falk, T.H.; McAlpine, D.; Marquardt, T. Cochlear Implant Filterbank Design and Optimization: A Simulation Study. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2014**, *22*, 347–353. [CrossRef]
29. Darling, A.M. Properties and Implementation of Gammatone Filters: A Tutorial. Available online: https://www.phon.ucl.ac.uk/home/shl5/Darling1991-GammatoneFilter.pdf (accessed on 4 March 2023).
30. Flanagan, J.L. Models for approximating basilar membrane displacement. *Bell Syst. Tech. J.* **1960**, *39*, 1163–1191. [CrossRef]
31. Boer, E.D. On the Principle of Specific Coding—A System Analysis of the Inner Ear Mechanism. In Proceedings of the International Federation of Automatic Control, Genova, Italy, 4–8 June 1973; Volume 6, pp. 187–194. [CrossRef]
32. Aertsen, A.M.H.J.; Johannesma, P.I.M.; Hermes, D.J. Spectro-temporal receptive fields of auditory neurons in the grass frog. *Biol. Cybern.* **1980**, *38*, 235–248. [CrossRef]
33. Dau, T.; Püschel, D.; Kohlrausch, A. A quantitative model of the effective signal processing in the auditory system. I. Model structure. *J. Acoust. Soc. Am.* **1996**, *99*, 3615–3622. [CrossRef]
34. Zeng, F.G. Trends in cochlear implants. *Trends Amplif.* **2004**, *8*, 1–34. [CrossRef]
35. Loizou, P.C. Signal-processing techniques for cochlear implants. *IEEE Eng. Med. Biol. Mag.* **1999**, *18*, 34–46. [CrossRef] [PubMed]
36. Rubinstein, J.T. How cochlear implants encode speech. *Curr. Opin. Otolaryngol. Head Neck Surg.* **2004**, *12*, 444–448. [CrossRef] [PubMed]
37. Ay, S.U.; Zeng, F.G.; Sheu, B.J. Hearing with bionic ears [cochlear implant devices]. *IEEE Circuits Devices Mag.* **1997**, *13*, 18–23. [CrossRef]
38. Loeb, G. Cochlear prosthetics. *Annu. Rev. Neurosci.* **1990**, *13*, 357–371. [CrossRef] [PubMed]
39. Millar, J.; Tong, Y.; Clark, G. Speech processing for cochlear implant prostheses. *J. Speech Hear. Res.* **1984**, *27*, 280–296. [CrossRef] [PubMed]
40. Parkins, C.; Anderson, S. *Cochlear Prostheses: An International Symposium*; New York Academy of Sciences: New York, NY, USA, 1983.
41. Loizau, P.C. Mimicking the Human Ear. *IEEE Signal Process. Mag.* **1998**, *15*, 101–130. [CrossRef]
42. Schindler, R.; Icessler, D. Preliminary results with the Clarion cochlear implant. *Laryngoscope* **1992**, *102*, 1006–1013. [CrossRef]
43. Kessler, D.; Schindler, R. Progress with a multi-strategy cochlear implant system: The Clarion. In *Advances in Cochlear Implants*; Hochmair-Desoyer, I., Hochmair, E., Eds.; Manz: Vienna, Austria, 1994; pp. 354–362.
44. House, W. A personal perspective on cochlear implants. In *Cochlear Implants*; Schindler, R., Merzenich, M., Eds.; Raven Press: New York, NY, USA, 1985; pp. 13–16.

45. House, W.; Urban, J. Long-term results of electrode implantation and electronic stimulation of the cochlea in man. *Ann. Otol. Rhinol. Laryngol.* **1973**, *82*, 504–517. [CrossRef] [PubMed]

46. House, W.; Berliner, K. Cochlear implants: Progress and perspectives. *Ann. Otol. Rhinol. Laryngol.* **1982**, *295* (Suppl. 91), 1–124.

47. Loizou, P.C.; Dorman, M.; Tu, Z. On the number of channels needed to understand speech. *J. Acoust. Soc. Am.* **1999**, *106*, 2097–2103. [CrossRef] [PubMed]

48. Bäckström, T. Introduction to Speech Processing: Pre-Emphasis. Available online: https://speechprocessingbook.aalto.fi/Preprocessing/Pre-emphasis.html (accessed on 26 January 2024).

49. Oppenheim, A.V.; Schafer, R.W. Digital Filter Design Techniques. In *Digital Signal Processing*; Prentice Hall: Upper Saddle River, NJ, USA, 1975; pp. 239–250.

50. Dau, T.; Püschel, D.; Kohlrausch, A. A quantitative model of the effective signal processing in the auditory system. II. Simulations and measurements. *J. Acoust. Soc. Am.* **1996**, *99*, 3623–3631. [CrossRef] [PubMed]

51. Patterson, R. Auditory images: How complex sounds are represented in the auditory system. *Acoust. Sci. Technol.* **2000**, *21*, 183–190. [CrossRef]

52. Cooke, M. A glimpsing model of speech perception in noise. *J. Acoust. Soc. Am.* **2006**, *119*, 1562–1573. [CrossRef] [PubMed]

53. Kubin, G.; Kleijn, W.B. Multiple-description coding (MDC) of speech with an invertible auditory model. In Proceedings of the IEEE Workshop on Speech Coding Proceedings, Model, Coders, and Error Criteria (Cat. No.99EX351), Porvoo, Finland, 20–23 June 1999; pp. 81–83. [CrossRef]

54. Kubin, G.; Kleijn, W.B. On speech coding in a perceptual domain. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Phoenix, AZ, USA, 15–19 March 1999; pp. 205–208. [CrossRef]

55. Patterson, R.D.; Holdsworth, J.A. A functional model of neural activity patterns and auditory image. *Adv. Speech Hear. Lang. Process.* **2004**, *3*, 547–563.

56. Unoki, M.; Irino, T.; Glasberg, B.; Moore, B.C.; Patterson, R.D. Comparison of the roex and gammachirp filters as representations of the auditory filter. *J. Acoust. Soc. Am.* **2006**, *120*, 1474–1492. [CrossRef]

57. Schofield, D. *Visualizations of the Speech Based on a Model of the Peripheral Auditory System*; NPL Report DITC 62/85; National Physical Laboratory: Teddington, UK, 1985.

58. Zhang, F.; Underwood, G.; McGuire, K.; Liang, C.; Moore, D.R.; Fu, Q.-J. Frequency Change Detection and Speech Perception in Cochlear Implant Users. *Hear. Res.* **2019**, *379*, 12–20. [CrossRef]

59. Medscape General Medicine. Hearing Loss: Does Gender Play a Role? Available online: https://www.medscape.com/viewarticle/719262_6?form=fpf (accessed on 21 January 2024).

60. Reich, R.D. Instrument Identification through a Simulated Cochlear Implant Processing System. Master's Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2012.