Properties and Implementation of the GammaTone Filter: A Tutorial

A. M. Darling

Abstract

Starting from the definition of the impulse response of the GammaTone filter, its complex spectrum, power spectrum, phase spectrum, the equivalent rectangular bandwidth, and the 3dB bandwidth are derived. A digital IIR filter is designed using the impulse invariance method to implement the GammaTone filter digitally through a multiple pass technique, and the approximations involved in the IIR filter implementation are clarified. Although most of these results have been published elsewhere, they are derived here in considerable detail as a tutorial.

1 Introduction

It is generally accepted that the frequency analysis performed by the peripheral auditory system can be modelled to a reasonable degree of accuracy by a bank of linear bandpass filters. Various filters, e.g. the family of roex filters (Patterson and Moore 1986), have been investigated for this purpose. A technique known as 'reverse correlation' (de Boer and Kuyper 1968), in which the responses of a primary auditory nerve fiber to white noise stimuli are measured and correlated with the input, is a method for measuring the auditory filter shape fairly directly from physiological preparations. This procedure gives rise to the so-called 'revcor' function, which within limits can be considered to be an estimate of the impulse response of the peripheral auditory filter that precedes the spike generation mechanism of that fiber. The GammaTone Filter (GTF) is an analytic mathematical function that approximates measured revcor functions (Johannesma 1972). The GTF has a convenient mathematical form that allows various filter properties to be derived analytically. As the GTF is derived from measured impulse responses, it has complete amplitude and phase information, not just amplitude information as in the case of filters derived from psychoacoustical masking experiments, such as the roex filter.

Holdsworth et al. (Holdsworth et al. 1988) stated various properties of the GTF and presented a digital multiple pass IIR filter scheme for its implementation. This technique enables GTF filter banks to be implemented much more efficiently than an FIR filter equivalent, and thus has important implications for practical applications (Patterson et al. 1988).

We are interested in using a GammaTone filter bank as the frequency analysis component in a computer model of speech perception. The purpose of this note is to derive the results of Holdsworth et al. (Holdsworth et al. 1988) as a tutorial, and to clarify some of the approximations involved. Emphasis is on the mathematical derivation of the GTF properties and the IIR filter design, rather than analysis of the filter parameters. To make this note reasonably self contained, standard Fourier Transform results used are listed in Appendix A.

2 The complex spectrum of the GammaTone filter

The impulse response of the GammaTone filter is defined as

$$h(t) = \begin{cases} c\,t^{n-1}\exp(-2\pi bt)\cos(2\pi f_0 t + \phi), & t > 0 \\ 0, & t < 0 \end{cases} \tag{2.1}$$

The GTF derives its name from the observation that h(t) is in the form of a carrier wave (tone) $\cos(2\pi f_0 t + \phi)$ amplitude modulated with an envelope that is proportional to $t^{n-1}\exp(-2\pi bt)$, which is the same functional form as the Gamma distribution in statistics. These two components gave rise to the name 'GammaTone'. The GTF parameters are:

c		-	proportionality constant
n		-	the filter order (controls the relative shape of the envelope, which becomes less skewed as n is increased [for fixed b])
b		-	temporal decay coefficient ($b > 0$) (increasing b shortens the duration of h(t))
f_0	(in Hz)	-	frequency of the carrier wave (determines the center frequency of the filter)
ϕ	(in radians)	-	carrier phase (determines the relative position of the envelope to the fine structure in the carrier wave)

The frequency response of the GTF is now derived to see what the corresponding effects of the GTF parameters are in the frequency domain.

For ease of mathematical manipulation, it is convenient to rewrite (2.1) as

$$h(t) = c\,r(t)\,s(t), \qquad -\infty \le t \le \infty \tag{2.2a}$$

where

$$s(t) = \cos(2\pi f_0 t + \phi) \tag{2.2b}$$

$$r(t) = t^{n-1}\exp(-2\pi bt)\,u(t) \tag{2.2c}$$

and $u(t)$ is the step function (see (A.8) in Appendix A). The introduction of $u(t)$ allows some standard Fourier Transform (FT) results to be applied (see

Appendix A). (Note: While ϕ is usually taken to be zero, it is included in the analysis here for completeness. It will be seen that for practical cases of interest, ϕ makes a negligible contribution to the power spectrum of the GTF.)

The frequency response of the GTF is obtained by taking the FT of $h(t)$, i.e.

$$H(f) = \int_{-\infty}^{\infty} h(t)\exp(-i2\pi ft)\,dt = c\int_{-\infty}^{\infty} r(t)\,s(t)\exp(-i2\pi ft)\,dt \tag{2.3}$$

Using the convolution theorem (see Appendix A), it follows from (2.3) that

$$H(f) = c[R(f) \otimes S(f)] \tag{2.4}$$

where $\otimes$ denotes the convolution operator and upper and lower case letters denote FT pairs. From (2.2c) and (A.13), substituting $m = n-1$ and $a = 2\pi b$

$$R(f) = (n-1)!\,[2\pi b + i2\pi f]^{-n} = (n-1)!\,(2\pi b)^{-n}\,[1 + if/b]^{-n} \tag{2.5}$$

(Note: Since $b > 0$, condition $a > 0$ in (A.6) is satisfied.) From (2.2b) and (A.10)

$$S(f) = \frac{1}{2}\,\{\exp(i\phi)\,\delta(f - f_0) + \exp(-i\phi)\,\delta(f + f_0)\} \tag{2.6}$$

Using the sifting property of the delta function (see Appendix A), it follows from (2.4-6) that

$$\begin{aligned} H(f) &= c[R(f) \otimes S(f)] \\ &= \frac{c}{2}(n-1)!\,(2\pi b)^{-n}\,\{\exp(i\phi)\,[1 + i(f - f_0)/b]^{-n} \\ &\qquad + \exp(-i\phi)\,[1 + i(f + f_0)/b]^{-n}\} \end{aligned} \tag{2.7}$$

Thus $H(f)$ is of the form

$$H(f) = k\,[P(f) + P^*(-f)] \tag{2.8a}$$

where

$$P(f) = \exp(i\phi)\,[1 + i(f - f_0)/b]^{-n} \tag{2.8b}$$

and

$$k = \frac{c}{2}(n-1)!(2\pi b)^{-n} \tag{2.8c}$$

i.e. $H(f)$ consists of the sum of the two components $kP(f)$ and $kP^*(-f)$, which are formed by centering the functions $B_1(f)$ and $B_2(f)$ at $f_0 Hz$ and $-f_0 Hz$ respectively, where

$$\left.\begin{aligned} B_1 &= \tfrac{c}{2}R(f)\exp(i\phi) \\ B_2 &= \tfrac{c}{2}R(f)\exp(-i\phi) \end{aligned}\right\} \tag{2.9}$$

Since the power spectrum of $R(f)$ has its maximum at $f = 0$, the power spectra of the individual components $kP(f)$ and $kP^*(-f)$ of $H(f)$ have their maximum at f_0Hz and $-f_0Hz$ respectively. As will be seen later however, this does not necessarily imply for the general case that the maximum of the power spectrum of $H(f)$ occurs at $\pm f_0Hz$.

Note: If x(t) is real, its FT must be Hermitian, i.e. $X(-f) = X^*(f)$, where * denotes complex conjugate. It can easily be verified that $R(f)$, $S(f)$ and $H(f)$ are all Hermitian as expected.

3 The power spectrum and phase spectrum of the GammaTone filter

In the previous section, the complex spectrum $H(f) = |H(f)|exp[i\psi(f)]$ of the GTF was derived. $|H(f)|$ and $\psi(f)$ are the amplitude spectrum and phase spectrum respectively. In this section the power spectrum $|H(f)|^2$ and the phase spectrum of the GTF are derived from the complex spectrum.

The power spectrum:

The power spectrum $|H(f)|^2 = H(f)H^*(f)$ of the GTF is given by

$$\begin{aligned} |H(f)|^2 &= k^2[P(f) + P^*(-f)][P^*(f) + P(-f)] \\ &= k^2[|P(f)|^2 + |P(-f)|^2 + P(f)P(-f) + P^*(f)P^*(-f)] \\ &= k^2[|P(f)|^2 + |P(-f)|^2 + 2\Re\{P(f)P(-f)\}] \qquad (3.1) \end{aligned}$$

where k and $P(f)$ are as defined in (2.8) and $\Re\{X\}$ denotes the real part of X. $P(f)$ can be expressed as

$$P(f) = exp(i\phi)[Q(f)]^{-n} = |Q(f)|^{-n}exp[-in\theta_1(f)]exp(i\phi) \qquad (3.2a)$$

where

$$Q(f) = [1 + i(f - f_0)/b] \qquad and \qquad \theta_1(f) = ARG[(f - f_0)/b] \qquad (3.2b)$$

$ARG[y/x]$ denotes the principal phase value of the complex variable z with real and imaginary parts μx and μy respectively, where μ is an arbitrary scaling factor. Hence from (3.2)

$$P(-f) = exp(i\phi)[Q(-f)]^{-n} = |Q(-f)|^{-n}exp[-in\theta_2(f)]exp(i\phi) \qquad (3.3a)$$

where

$$Q(-f) = [1 - i(f + f_0)/b] \qquad and \qquad \theta_2(f) = ARG[-(f + f_0)/b] \qquad (3.3b)$$

From (3.2-3)

$$\begin{aligned} |P(f)|^2 &= |Q(f)|^{-2n} = [1 + (f - f_0)^2/b^2]^{-n} \qquad (3.4a) \\ |P(-f)|^2 &= |Q(-f)|^{-2n} = [1 + (f + f_0)^2/b^2]^{-n} \qquad (3.4b) \end{aligned}$$

$$\begin{aligned} P(f)P(-f) &= [Q(f)Q(-f)]^{-n}exp(i2\phi) \\ &= [(1 + i(f - f_0)/b)(1 - i(f + f_0)/b)]^{-n}exp(i2\phi) \\ &= [1 + (f^2 - f_0^2)/b^2 - i2f_0/b]^{-n}exp(i2\phi) \qquad (3.4c) \\ \Re\{P(f)P(-f)\} &= \\ & [(1 + (f^2 - f_0^2)/b^2)^2 + (2f_0/b)^2]^{-n/2}cos[2\phi - n\theta(f)] \qquad (3.4d) \end{aligned}$$

where

$$\theta(f) = ARG\left\{\frac{-2f_0}{b[1 + (f^2 - f_0^2)/b^2]}\right\} = ARG\left\{\frac{-2f_0b}{b^2 + (f^2 - f_0^2)}\right\} \qquad (3.4e)$$

From (3.1-4)

$$\begin{aligned} |H(f)|^2 &= k^2\{|Q(f)|^{-2n} + |Q(-f)|^{-2n} + \\ & 2\Re\{[Q(f)Q(-f)]^{-n}exp(i2\phi)\}\} \qquad (3.5a) \\ &= k^2\{[1 + (f - f_0)^2/b^2]^{-n} + [1 + (f + f_0)^2/b^2]^{-n} + \\ & 2\Re\{[1 + (f^2 - f_0^2)/b^2 - i2f_0/b]^{-n}exp(i2\phi)\}\} \qquad (3.5b) \\ &= k^2\{[1 + (f - f_0)^2/b^2]^{-n} + [1 + (f + f_0)^2/b^2]^{-n} + \\ & 2[(1 + (f^2 - f_0^2)/b^2)^2 + (2f_0/b)^2]^{-n/2}cos[2\phi - n\theta(f)]\} \qquad (3.5c) \end{aligned}$$

where

$$\begin{aligned} k &= \frac{c}{2}(n - 1)!(2\pi b)^{-n} \\ \theta(f) &= ARG\left\{\frac{-2f_0b}{b^2 + (f^2 - f_0^2)}\right\} \end{aligned}$$

as defined above.

As can be seen from (3.5), the smaller b, the more rapid the decay of $|H(f)|^2$ away from $\pm f_0$, which corresponds to a slower decay of the GTF impulse response. The larger the ratio f_0/b, the less the components $kP(f)$ and $kP^*(-f)$ of $H(f)$ overlap, and the less the contribution of $kP(f)$ and $kP^*(-f)$ to $|H(f)|^2$ away from f_0 and $-f_0$ respectively. Although the power spectra of $kP(f)$ and $kP^*(-f)$ have their maximum at f_0 and $-f_0$ respectively, for the general case with no restrictions on the GTF parameters, the power spectrum of $H(f)$ does not necessarily have maxima at $\pm f_0Hz$. When the components

$kP(f)$ and $kP^*(-f)$ overlap significantly, $|H(f)|^2$ has the character of a low-pass filter with a single peak at the origin. As f_0/b is increased (for fixed n), this single peak splits and the maxima move outwards and eventually converge on $\pm f_0$. Since the purpose of the GTF in auditory modelling is to model a bandpass filter, the components $kP(f)$ and $P^*(-f)$ must be well separated, and this in turn requires that the ratio f_0/b is large. When f_0/b is large enough, $H(f)$ can be approximated as

$$H(f) \approx \begin{cases} kP(f) & f \geq 0 \\ kP^*(-f) & f < 0 \end{cases} \tag{3.6}$$

and

$$|H(f)|^2 \approx k^2\,[1+(|f|-f_0)^2/b^2]^{-n} \tag{3.7}$$

It is seen that the full form of the power spectrum, but not the approximate form, contains ϕ. For small f_0/b, ϕ influences where the maxima of the power spectrum occur, but has negligible effect for large f_0/b.

Initial estimates show that the approximation is probably quite acceptable for $f_0/b > 2$ when $n = 4$ (a typical value for n in auditory modelling). In general, for fixed f_0/b, the larger n, the better the approximation. According to Holdsworth et. al. (1988), when modelling the human auditory system, f_0/b is typically in the range $4 < f_0/b < 8$. Hence for auditory modelling, this approximation appears to be appropriate.

The phase spectrum:

The phase spectrum of the GTF is given by

$$\psi(f) = ARG\left(\frac{\Im\{H(f)\}}{\Re\{H(f)\}}\right) \tag{3.8}$$

where $\Re\{X\}$ and $\Im\{X\}$ denote the real and imaginary part of X respectively. From (3.2-3)

$$\left.\begin{array}{lcll} \Re\{P(f)\} & = & |Q(f)|^{-n} & cos[\phi - n\theta_1(f)] \\ \Im\{P(f)\} & = & |Q(f)|^{-n} & sin[\phi - n\theta_1(f)] \\ \Re\{P^*(-f)\} & = & |Q(-f)|^{-n} & cos[-\{\phi - n\theta_2(f)\}] \\ \Im\{P^*(-f)\} & = & |Q(-f)|^{-n} & sin[-\{\phi - n\theta_2(f)\}] \end{array}\right\} \tag{3.9}$$

Expanding (3.9) using (3.4), then from (2.8) and (3.8) the phase spectrum can be expressed as

$$\psi(f) = ARG\left(\frac{\Im\{P(f)\} + \Im\{P^*(-f)\}}{\Re\{P(f)\} + \Re\{P^*(-f)\}}\right) = ARG(\beta) \tag{3.10a}$$

where

$$\beta = \frac{[1+(f-f_0)^2/b^2]^{-n/2} sin[\phi - n\theta_1(f)] - [1+(f+f_0)^2/b^2]^{-n/2} sin[\phi - n\theta_2(f)]}{[1+(f-f_0)^2/b^2]^{-n/2} cos[\phi - n\theta_1(f)] + [1+(f+f_0)^2/b^2]^{-n/2} cos[\phi - n\theta_2(f)]} \tag{3.10b}$$

[using $sin(-x) = -sin(x)$ and $cos(-x) = cos(x)$]. It is noted that $\psi(0) = 0$ and $\psi(-f) = -\psi(f)$, as expected, due to the Hermitian property of $H(f)$.

As for the power spectrum, the expression for the phase spectrum can be simplified when the ratio f_0/b is large. In this case for $f > 0$, the terms in β containing $f + f_0$ become negligible when f is in the region of f_0 or larger, and $\psi(f)$ reduces to

$$\psi(f) \approx ARG\left(\frac{sin[\phi - n\theta_1(f)]}{cos[\phi - n\theta_1(f)]}\right) = \phi - n\theta_1(f) = \phi - n\; ARG[(f-f_0)/b] \tag{3.11a}$$

Similarly, for $f < 0$, the terms in β containing $f - f_0$ become negligible when f is in the region of $-f_0$ or smaller, and $\psi(f)$ reduces to

$$\psi(f) \approx -[\phi - n\theta_2(f)] = -(\phi - n\; ARG[(-(f+f_0)/b)]) \tag{3.11b}$$

Thus when f_0/b is large enough,

$$\psi(f) \approx \begin{cases} \phi & -n\; ARG[(|f|-f_0)/b]) \quad f > 0 \\ -\phi & +n\; ARG[(|f|-f_0)/b]) \quad f < 0 \end{cases} \tag{3.12}$$

Although for the approximate phase spectrum $\psi(-f) = -\psi(f)$, in general $\psi(0) \neq 0$. Even for large f_0/b, the approximate phase spectrum is not necessarily a good approximation to the true phase spectrum near the origin. However, regions near the origin where the phase approximation is not good corresponds to regions where the power spectrum is negligible.

It is seen that both the full and approximate phase spectrum contain ϕ, whereas it only features in the full, but not the approximate, power spectrum.

4 The equivalent rectangular bandwidth and 3dB bandwidth of the GammaTone filter

In addition to the power spectrum, two filter parameters that are of interest are the equivalent rectangular bandwith (ERB) and the 3dB bandwidth BW_{3dB}.

The equivalent rectangular bandwidth of the GTF:

The equivalent rectangular bandwidth (ERB) of a filter $X(f)$ is typically defined as the width of a rectangular filter whose height equals the maximum of the power spectrum of $X(f)$ and passes the same amount of power as $X(f)$. The concept of the ERB is illustrated in Fig. 4.1 for a hypothetical bandpass filter with well separated bands and power spectrum maxima at $\pm f_0$. A slightly more general definition of the ERB allows the height of the rectangular filter to be merely predefined at some fixed level. The ERB for the general GTF where the height of the rectangular filter is $|H(f_0)|^2$ can be calculated as follows, making no assumptions at present as to whether this actually is the maximum of the power spectrum of $H(f)$.

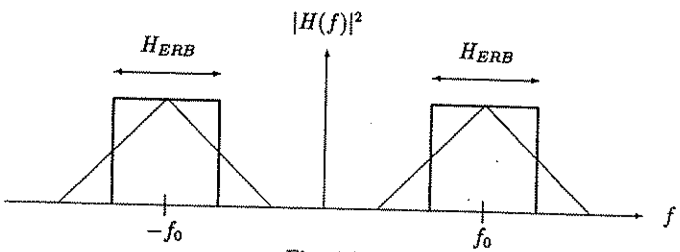

Fig. 4.1

The area of each rectangular box in Fig. 4.1 is $H_{ERB}|H(f_0)|^2$, which is equal to half of the power passed by the GTF. Since the power passed by the GTF is the integral of the GTF power spectrum, the relationship between H_{ERB} and $|H(f)|^2$ is

$$\begin{aligned} 2H_{ERB}|H(f_0)|^2 &= \int_{-\infty}^{\infty} |H(f)|^2 df \qquad (4.1) \\ &= 2\int_{0}^{\infty} |H(f)|^2 df \\ &\quad \text{(since } |H(f)|^2 \text{ is symmetric)} \end{aligned}$$

Applying Parseval's theorem and the 'Summation' property of the FT (see Appendix A) to (4.1),

$$H_{ERB} = \frac{\int_{-\infty}^{\infty} |H(f)|^2 df}{2|H(f_0)|^2} = \frac{\int_{-\infty}^{\infty} |h(t)|^2 dt}{2|H(f_0)|^2} = \frac{\breve{H}(0)}{2|H(f_0)|^2} \qquad (4.2a)$$

where

$$\breve{h}(t) = |h(t)|^2 \qquad (4.2b)$$

From (2.2)

$$\breve{h}(t) = c^2\, \breve{r}(t)\breve{s}(t) \qquad (4.3a)$$

$$\breve{r}(t) = t^{2n-2} exp(-4\pi bt)u(t) \qquad (4.3b)$$

$$\breve{s}(t) = cos^2(2\pi f_0 t + \phi) \qquad (4.3c)$$

(Note that $u(t)^2 = u(t)$.) Thus

$$\breve{H}(f) = c^2[\breve{R}(f) \otimes \breve{S}(f)] \qquad (4.4)$$

Applying (A.13) to (4.3b)

$$\begin{aligned} \breve{R}(f) &= (2n-2)!\,(4\pi b + i2\pi f)^{-(2n-1)} \\ &= (2n-2)!\,(4\pi b)^{-(2n-1)}\,(1 + if/2b)^{-(2n-1)} \qquad (4.5) \end{aligned}$$

and applying (A.9) and the convolution theorem to (4.3c)

$$\begin{aligned} \breve{S}(f) &= \left[\frac{exp(i\phi)}{2}\,\delta(f-f_0) + \frac{exp(-i\phi)}{2}\,\delta(f+f_0)\right] \otimes \\ &\qquad \left[\frac{exp(i\phi)}{2}\,\delta(f-f_0) + \frac{exp(-i\phi)}{2}\,\delta(f+f_0)\right] \\ &= \frac{exp(i2\phi)}{4}\delta(f-2f_0) + \frac{exp(-i2\phi)}{4}\delta(f+2f_0) + \frac{1}{2}\delta(f) \qquad (4.6) \end{aligned}$$

as can be verified using (A.3). From (4.4-6)

$$\begin{aligned} \breve{H}(f) &= c^2[\breve{R}(f) \otimes \breve{S}(f)] \\ &= c^2(2n-2)!\,(4\pi b)^{-(2n-1)} \left\{\frac{exp(i2\phi)}{4}[1 + i(f-2f_0)/2b]^{-(2n-1)} \right. \\ &\qquad \left. + \frac{exp(-i2\phi)}{4}[1 + i(f+2f_0)/2b]^{-(2n-1)} + \frac{1}{2}[1 + i(f/2b]^{-(2n-1)}\right\} \\ &\qquad (4.7) \end{aligned}$$

For $f = 0$

$$\begin{aligned} \breve{H}(0) &= c^2\,(2n-2)!\,(4\pi b)^{-(2n-1)} \left\{\frac{exp(i2\phi)}{4}[1 - if_0/b]^{-(2n-1)} \right. \\ &\qquad \left. + \frac{exp(-i2\phi)}{4}[1 + if_0/b]^{-(2n-1)} + \frac{1}{2}\right\} \qquad (4.8) \end{aligned}$$

Using $X^m + [X^*]^m = 2\Re\{X^m\} = 2|X|^m cos(m\theta)$, with $\theta = ARG(\Im\{X\}/\Re\{X\})$, (4.8) becomes

$$\begin{aligned} \breve{H}(0) &= c^2\,(2n-2)!\,(4\pi b)^{-(2n-1)} \\ &\quad \left\{\frac{2}{4}(1 + f_0^2/b^2)^{-(2n-1)/2} cos(2\phi - [2n-1]ARG[-f_0/b]) + \frac{1}{2}\right\} \qquad (4.9) \end{aligned}$$

($ARG()$ is as defined in the previous section.) From (3.5)

$$|H(f_0)|^2 = k^2 \{1 + [1 + (2f_0/b)^2]^{-n} + 2[1 + (2f_0/b)^2]^{-n/2} \cos(2\phi - n\, ARG[-2f_0/b])\} \tag{4.10}$$

(4.9) and (4.10) in (4.2)

$$H_{ERB} = \eta\, \mu \tag{4.11a}$$

where

$$\mu = \frac{\frac{2}{4}[1 + (f_0/b)^2]^{-(2n-1)/2} cos(2\phi - [2n-1]ARG[-f_0/b]) + \frac{1}{2}}{1 + [1 + (2f_0/b)^2]^{-n} + 2[1 + (2f_0/b)^2]^{-n/2} cos[2\phi - n\, ARG(-2f_0/b)]} \tag{4.11b}$$

and

$$\eta = \frac{c^2(2n-2)!(4\pi b)^{-(2n-1)}}{2c^2[(n-1)!]^2(2\pi b)^{-2n}2^{-2}} = \frac{(2n-2)!2^{-(2n-1)2}(\pi b)^{-(2n-1)}}{[(n-1)!]^2 2^{-(2n+1)}(\pi b)^{-2n}}$$
$$= \frac{(2n-2)!\, \pi b\, 2^{3-2n}}{[(n-1)!]^2} \tag{4.11c}$$

(4.11) gives the ERB of a general GTF where the height of the rectangular filter is $|H(f_0)|^2$. As mentioned in Section 3, in practice we are interested in the case where f_0/b is sufficiently large so that the components of $H(f)$ are well separated and the maximum value $|H(f)|^2$ occurs at $\pm f_0$. In this case the factors in (4.11b) containing the ratio f_0/b become negligible and μ tends to 1/2, so that (4.11) can be simplified to

$$H_{ERB} \approx \eta/2 = \frac{(2n-2)!\, \pi b\, 2^{2-2n}}{[(n-1)!]^2} \tag{4.12}$$

Thus for large f_0/b, H_{ERB} is seen to be proportional to b and independent of f_0, i.e.

$$H_{ERB} \approx \alpha(n) b \tag{4.13a}$$

where

$$\alpha(n) = \frac{(2n-2)!2^{2-2n}\pi}{[(n-1)!]^2} \tag{4.13b}$$

(Note that this approximation to H_{ERB} could have been derived directly from the approximate form of $|H(f)|^2$ for large f_0/b (see (3.7)), rather than deriving it from the more general case as done above. As mentioned in Section 3, for any realistic GTF for auditory modelling this approximation appears to be adequate. However, the general case is given for completeness.)

The 3dB bandwidth of the GTF:

Let

$$|H(\varepsilon)|^2 = |H(f_0)|^2/2 \tag{4.14}$$

(where $|H(f)|^2$ has a maximum at f_0, then $2|(\varepsilon - f_0)|$ is the 3dB bandwidth H_{3dB} of the GTF. Using the assumption now that f_0/b is large, then from (3.7)

$$|H(f)|^2 \approx k^2[1 + (f - f_0)^2/b^2]^{-n} \qquad for f \geq 0 \tag{4.15}$$

hence

$$|H(f_0)|^2 \approx k^2 \tag{4.16}$$

and

$$|H(\varepsilon)|^2 \approx k^2[1 + (\varepsilon - f_0)^2/b^2]^{-n} \tag{4.17}$$

Substituting (4.16 and 4.17) into (4.14)

$$k^2 2^{-1} = k^2[1 + (\varepsilon - f_0)^2/b^2]^{-n} \tag{4.18}$$

from which it follows that

$$(\varepsilon - f_0)^2 = [(2^{1/n} - 1)b^2] \tag{4.19}$$

and

$$\varepsilon - f_0 = b(2^{1/n} - 1)^{1/2} \tag{4.20}$$

Thus the approximate H_{3dB} is proportional to n and independent of f_0, and given by

$$H_{3dB} \approx \beta(n) b \tag{4.21a}$$

where

$$\beta(n) = 2(2^{1/n} - 1)^{1/2} \tag{4.21b}$$

5 Implementation of the GammaTone filter through a multiple pass IIR filtering technique

In this section it is shown how a multiple pass first order IIR filter can be designed to implement the GTF. It was shown in Section 2 (see (2.8)) that $H(f)$ can be expressed as

$$H(f) = A(f) + A^*(-f) \tag{5.1a}$$

where k has now been absorbed into $P(f)$, i.e.

$$A(f) = kP(f) \tag{5.1b}$$

Since $x^*(t) \xleftrightarrow{\mathcal{FT}} X^*(-f)$,

$$h(t) = a(t) + a^*(t) = 2\Re\{a(t)\} \tag{5.2}$$

From (5.2) it follows that if $x(t)$ is real, as is the case for speech signals,

$$x(t) \otimes h(t) = 2\Re\{x(t) \otimes a(t)\} \tag{5.3}$$

i.e. for real input, filtering with filter H(f) is equivalent to filtering with $A(f)$, followed by taking the real part of the filtered signal in the time domain and multiplying by 2.

$A(f)$ can be expressed as

$$A(f) = exp(i\phi)\,[k^{1/n}\tilde{A}(f)]^n = exp(i\phi)\,[\gamma\tilde{A}(f)]^n \tag{5.4a}$$

where

$$\tilde{A}(f) = [1 + i(f - f_0)/b]^{-1} \tag{5.4b}$$

From (5.4a) it is seen that filtering by $A(f)$ is equivalent to filtering by $\gamma\tilde{A}(f)$ n times, and then post-multiplying the output by $exp(i\phi)$. $k^{1/n}$ has been absorbed into the gain parameter γ, which can be adjusted to give the desired gain. (This is easier than adjusting c to obtain the desired gain.)

A first order IIR filter for implementing $\gamma\tilde{A}(f)$ can be designed using the impulse invariance method (Bozic 1981). A full discussion of this method is beyond the scope of this communication, and the main steps of this design technique are merely stated without proof.

Basic steps for designing a first order IIR filter to implement the filter $\gamma\tilde{A}(f) \overset{\mathcal{FT}}{\longleftrightarrow} \gamma\tilde{a}(t)$ using the impulse invariance method:

a) Replace $\gamma\tilde{a}(t)$ with its sampled form $\gamma\tilde{a}(mT)$, where T is the sampling interval, and express it in the form

$$\gamma\tilde{a}(t) = \gamma U w^m \tag{5.5}$$

where U is a constant.

b) Take the z-transform $\gamma\tilde{A}_z(z)$ of $\gamma\tilde{a}(mT)$

$$\gamma\tilde{A}_z(z) = \gamma\sum_{m=0}^{\infty} a(mT)z^{-m} = \gamma U\sum_{m=0}^{\infty} w^m z^{-m} = \frac{\gamma U}{1 - wz^{-1}} \tag{5.6}$$

(provided $|wz^{-1}| < 1$).

c) Write the first order IIR filter as

$$y_j = \gamma U x_j + w y_{j-1} \tag{5.7}$$

where x_j and y_j are the jth input and output samples respectively.

Steps (a - c) can be carried out for the GTF as follows: From (5.4b) and (A.5-6)

$$\begin{aligned}\gamma\tilde{A}(f) &= \gamma\, 2\pi b[2\pi b + i2\pi(f - f_0)]^{-1} \overset{\mathcal{FT}}{\longleftrightarrow} \\ &\quad \gamma\, 2\pi b\, exp(i2\pi f_0 t)exp(-2\pi bt)u(t) = \gamma\tilde{a}(t)\end{aligned} \tag{5.8}$$

Following Step (a),

$$\gamma\tilde{a}(mT) = \gamma\, 2\pi b\, exp(i2\pi f_0 mT)exp(-2\pi bmT), \quad m \geq 0 \tag{5.9}$$

(u(t) is dropped as m is restricted to be non-negative.) It is seen that $\gamma\tilde{a}(mT)$ is in the form of (5.5). (It is noted that $\gamma h(t)$ for $n > 1$ can not be expressed in this form due to the term $(mT)^{n-1}$.) Step (b) then leads to

$$\begin{aligned}\gamma\tilde{A}_z(z) &= \gamma\, 2\pi b\sum_{m=0}^{\infty}\{exp[2\pi T(if_0 - b)]\}^m z^{-m} \\ &= \frac{\gamma\, 2\pi b}{1 - exp[2\pi T(if_0 - b)]z^{-1}}\end{aligned} \tag{5.10}$$

(Note that the condition $|wz^{-1}| < 1$ is satisfied.) According to Step (c), the IIR filter is now given by

$$y_j = \gamma\, 2\pi b\, x_j + exp[2\pi T(if_0 - b)]y_{j-1} \tag{5.11}$$

Calculation of γ for unit gain at $f = f_0$:

Since $z = exp(i2\pi fT)$, $f = f_0$ leads to $z = exp(i2\pi f_0 T)$. Hence

$$\begin{aligned}\gamma\tilde{A}_z(exp[i2\pi f_0]) = 1 &= \frac{\gamma\, 2\pi b}{1 - exp[2\pi T(if_0 - b)]\, exp[-i2\pi f_0 T]} \\ &= \frac{\gamma\, 2\pi b}{1 - exp(-i2\pi bT)}\end{aligned} \tag{5.12}$$

and

$$\gamma = \frac{1 - exp(-2\pi bT)}{2\pi b} \tag{5.13}$$

(5.10) now becomes

$$y_j = [1 - exp(-2\pi bT)]x_j + exp[2\pi T(if_0 - b)]y_{j-1} \tag{5.14}$$

For each output sample y_j, there is one multiplication by a real constant, one multiplication by a complex constant and one complex addition.

Multiple pass IIR algorithm for implementing the GTF:

The algorithm for implementing a GTF of order n can now be stated as follows:

1. Set iteration counter N to 0
2. Increment the iteration counter N by 1
3. Pass the input array $x\{j\}$ through the digital IIR filter as expressed in (5.14) to get the output array $y\{j\}$ (Nth iteration).
4. If $n > N$, treat the current output array $y\{j\}$ as new input $x\{j\}$ and go to step 2. (After n iterations through steps 2 - 4, the original input has been filtered by $[\gamma\tilde{A}(f)]^n$.)
5. If $\phi \neq 0$, multiply the last output $y\{j\}$ by $exp(i\phi)$ (Filtering by $A(f)$ has now been achieved.)
6. Take the real part of the last output $y\{j\}$and multiply by 2.

Since on all but the first pass through the IIR filter the filter input is complex, calculation of y_j in (5.14) requires the multiplication of x_j by a real constant, the multiplication of y_{j-1} by a complex constant and one complex addition, leading to a total of 6 real multiplications and 4 real additions per sample per pass. For an nth order filter, the total number of calculations required per sample for steps 1-4 of the algorithm is 6n real multiplications and 4n real additions (treating the input on the first pass as complex).

Holdsworth et al. (Holdsworth et al. 1988) designed an IIR filter to implement the low-pass filter $\gamma\,[1+if/b]^{-1}$ rather than $\gamma\,[1+i(f-f_0)/b]^{-1}$, leading to the IIR filter

$$\begin{aligned} y_j &= [1-exp(-2\pi bT)]x_j + exp(-2\pi Tb)y_{j-1} && (5.15a) \\ &= x_j + exp(-2\pi bT)(y_{j-1}-x_j) && (5.15b) \end{aligned}$$

Note that (5.15b) is more efficient than (5.15a), requiring one multiplication less but one extra addition. However to implement a GTF with this IIR filter, the original input has to be frequency shifted by $-f_0Hz$ prior to filtering (i.e. prior to step 1 in the above algorithm) and the output after step 4 has to be frequency shifted back by f_0Hz. This is achieved by producing the new input array

$$x_j = exp(-i2\pi f_0Tj)x_j \qquad (5.16)$$

prior to step 1 and the new output array

$$y_j = exp(i2\pi f_0Tj)y_j \qquad (5.17)$$

after step 4 (see (A.5)).

Calculation of y_j in (5.15b) requires 2 real multiplications and 4 real additions. For an nth order GTF, the total number of calculations for performing steps 1-4 of the algorithm (i.e. excluding the frequency shifting) is 2n real multiplications and 4n real additions. The total number of calculations with the up and down shifting is (2n + 6) real multiplications and (4n + 2) real additions, taking into account that the original input is real. If $\phi = 0$ so that the complex multiplication by $exp(i\phi)$ is not required, only the real part of the output after shifting back by f_0Hz is required and the total number of real multiplications and additions can both be reduced by 1. In terms of the number of calculations required to accomplish filtering by $[\gamma\tilde{A}(f)]^n$ the algorithm with up and down frequency shifting appears computationally more efficient. However, the overhead for calculating the complex exponentials (or retrieving them from a look up table) to effect the frequency shifting has not been included in the calculation.

It is emphasized that both versions of the algorithm for the digital multiple pass IIR filter approximation to the GTF described above are equivalent in terms of their filtering properties. Although the IIR filter designed in both cases is based on only one component of the full GTF, the final step in the algorithm of taking the real part of the filtered signal and multiplying by 2 ensures that in effect the full GTF (rather than the simplified version of the GTF obtained by assuming f_0/b is large) is approximated.

Additional aspects related to the digital IIR implementation, such as aliasing problems and techniques for avoiding them, are beyond the scope of this discussion.

6 Summary

The complex spectrum, power spectrum, phase spectrum, equivalent rectangular bandwidth and 3dB bandwidth of the GTF have been derived. While for practical cases of interest when the ratio f_0/b is large certain approximations can be made, only the calculation of the 3dB bandwidth actually made use of this assumption. The range of f_0/b over which the approximations are valid has not been rigorously examined here, but appears to be appropriate for practical applications of the GTF to auditory modelling. Furthermore, it was shown that the digital multiple pass IIR filter approximation to the GTF is an approximation to the full GTF and not its simplified form obtained by assuming f_0/b is large. (This appears to have been a source of confusion in the past.)

Acknowledgements

This work was carried out as part of a study supported by the Medical Research Council of Great Britain on computational modelling of speech perception.

References

Bozic S.M. (1981)**Digital and Kalman Filtering**,Edward-Arnold

de Boer E. and Kuyper P. (1968)'Triggered correlation', **IEEE Transactions on Biological and Medical Engineering** , BME 15-3

Holdsworth J. et al. (1988)'Implementing a Gamma Tone Filter Bank', in **SVOS Final Report** - Part A: The Auditory Filterbank, MRC Applied Psychology Unit, Cambridge, England

Johannesma P.I.M. (1972)'The pre-response stimulus ensemble of neurons in the cochlear nucleus', **Proceedings of the Symposium on Hearing Theory**, IPO, Eindhoven, The Netherlands

Patterson R.D. and Moore B.J.C. (1986)'Auditory filters and excitation patterns as representations of frequency resolution', in B.C.J. Moore (Ed.), **Frequency Selectivity in Hearing**,Academic Press

Patterson R.D. et al. (1988) in **SVOS Final Report** - Part A: The Auditory Filterbank, MRC Applied Psychology Unit, Cambridge, England

Appendix A: Summary of notation and standard results used

Some Fourier Transform Pairs

With the Fourier Transform (FT) defined as

$$X(f) = \int_{-\infty}^{\infty} x(t)\, exp(-i2\pi ft)\, dt \qquad \text{(A.1a)}$$

$$x(t) = \int_{-\infty}^{\infty} X(f)\, exp(i2\pi ft)\, df \qquad \text{(A.1b)}$$

where upper and lower case letters denote FT pairs, i.e.

$$x(t) \xleftrightarrow{\mathcal{FT}} X(f)$$

the following FT pairs can be established:

$$\begin{aligned}
x^*(\pm t) &\xleftrightarrow{\mathcal{FT}} X^*(\mp f) && \text{(A.2)}\\
\delta(t \pm a) &\xleftrightarrow{\mathcal{FT}} exp(\pm i2\pi af) && \text{(A.3)}\\
exp(\pm i2\pi at) &\xleftrightarrow{\mathcal{FT}} \delta(f \mp a) && \text{(A.4)}\\
x(t)exp(\pm i2\pi at) &\xleftrightarrow{\mathcal{FT}} X(f \mp a) && \text{(A.5)}\\
exp(-at)\, u(t) &\xleftrightarrow{\mathcal{FT}} (a + i2\pi f)^{-1}, \quad a > 0 && \text{(A.6)}\\
t^m\, x(t) &\xleftrightarrow{\mathcal{FT}} (-i2\pi)^{-m} \frac{d^m}{df^m} X(f) && \text{(A.7)}
\end{aligned}$$

where $\delta(t)$ is the delta function, u(t) is the step function

$$u(t) = \begin{cases} 1, & t > 0 \\ 0, & t < 0 \end{cases} \qquad \text{(A.8)}$$

and x^* denotes the complex conjugate of x.

Application of (A.4) to obtain the FT of $cos(2\pi f_0 t + \phi)$:

Since $cos(y) = \frac{exp(iy)}{2} + \frac{exp(-iy)}{2}$

$$\begin{aligned}
cos(2\pi f_0 t + \phi) &= \frac{1}{2} exp(i2\pi f_0 t + i\phi) + \frac{1}{2} exp(-i2\pi f_0 t - i\phi)\\
&= \frac{1}{2} exp(i\phi)\, exp(i2\pi f_0 t) + \frac{1}{2} exp(-i\phi)\, exp(-i2\pi f_0 t)
\end{aligned} \qquad \text{(A.9)}$$

Applying (A.4) to (A.9) and using the linearity of the FT operator

$$cos(2\pi f_0 t + \phi) \xleftrightarrow{\mathcal{FT}} \frac{1}{2} exp(i\phi)\, \delta(f - f_0) + \frac{1}{2} exp(-i\phi)\, \delta(f + f_0) \qquad \text{(A.10)}$$

Application of (A.6) and (A.7) to obtain the FT of $t^m\, exp(-at)\, u(t)$:

From (A.6) and (A.7)

$$t^m\, exp(-at)\, u(t) \overset{\mathcal{FT}}{\longleftrightarrow} (-i2\pi)^{-m} \frac{d^m}{df^m}[a + i2\pi f]^{-1} \qquad \text{(A.11)}$$

Let $X(f) = [Y(f)]^{-1} = [a + i2\pi f]^{-1}$, then

$$\begin{aligned} \frac{d}{df}X(f) &= \frac{d}{dY(f)}X(f)\,\frac{d}{df}Y(f) = (-1)\,Y(f)^{-2}\,(i2\pi) \\ \frac{d^2}{df^2}X(f) &= (-1)\,(-2)\,Y(f)^{-3}\,(i2\pi)^2 \end{aligned}$$

and in general

$$\frac{d^m}{df^m}X(f) = (-1)^m\, m![a + i2\pi f]^{-(m+1)}(i2\pi)^m = m!\,[a + i2\pi f]^{-(m+1)}(-i2\pi)^m \qquad \text{(A.12)}$$

(A.12) in (A.11)

$$\begin{aligned} t^m\, exp(-at)\, u(t) &\overset{\mathcal{FT}}{\longleftrightarrow} (-i2\pi)^{-m}\,(-i2\pi)^m\, m!\,[a + i2\pi f]^{-(m+1)} \\ &= m!\,[a + i2\pi f]^{-(m+1)} \end{aligned} \qquad \text{(A.13)}$$

Some Theorems

Convolution Theorem:
If $x(t) = v(t)\, w(t)$, then

$$X(f) = V(f) \otimes W(f)$$

where $\otimes$ denotes the convolution operator.

Parseval's Theorem:

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

Some Properties of the FT and Delta Functions

'Summation' property of the FT:

$$\int_{-\infty}^{\infty} x(t)\, dt = X(0)$$

$$\int_{-\infty}^{\infty} X(f)\, df = x(0)$$

Sifting property of the delta function:

$$X(f) \otimes \delta(f - a) = X(f - a)$$