Scan the QR code for a copy of this poster and the link to open-source code (or email me)

Direct Characteristics-Based Design of Filterbanks for Perceptual Studies and Speech and Hearing Technologies

Samiya A Alkhairy

samiya@mit.edu, samiya@alum.mit.edu

1. Goal

- **Auditory filters** (e.g. Gammatone filters, APGFs, OZGFs, P/GEFs, V, …) are typically used in filterbank configurations.
- Auditory filterbanks are useful for sound source localization models, hearing loss simulations, cochlear implants, hearing aids, speech feature extraction (e.g. pitch estimation), speech recognition, and spike-based pattern recognition. Associated multiband filters may be used for equalizers and for denoising signals.

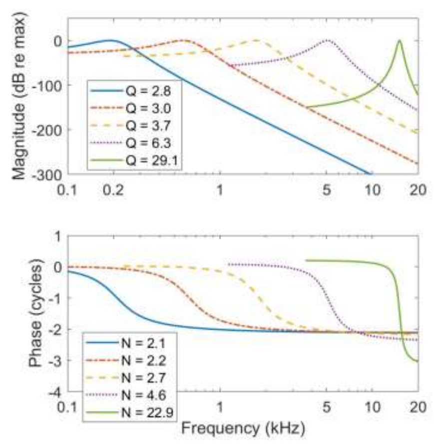


Figure - Auditory Filterbanks. Figure from Alkhairy & Shera (2019).

- Auditory filters are naturally described by **filter characteristics** (Ψ) such as peak frequency, 3dB quality factor, equivalent rectangular bandwidth (ERB), convexity and group delay at the peak.

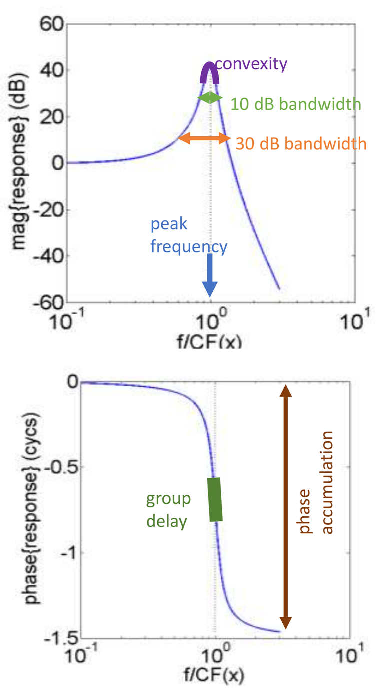


Figure – Filter Characteristics

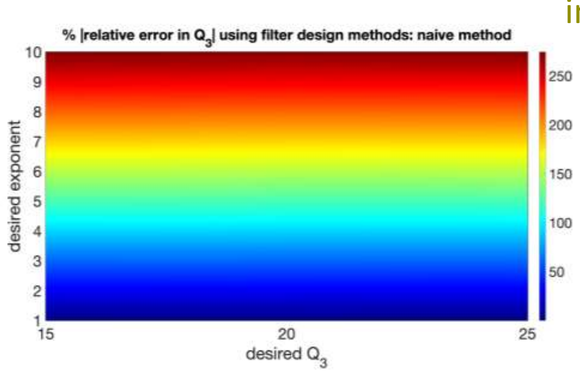
- **Goal**: is to develop methods to design auditory filters given specifications on sets of filter characteristics.

- **Achieving this goal enables**:
  1) Designing filters to accurately achieve specifications on characteristics – e.g. to mimic normal or pathological behavior of humans or other species
  2) Having direct control over the values of filter characteristics to systematically study the dependence of outcomes of perceptual studies and technological contributions on sets of filter characteristics

2. Current Methods

- Current frequency-domain methods for designing auditory filters

(A) Naïve method (based on parameterization by second order system half-power quality factor, $Q_{3,sos}$, and filter order) → (1) parameterized by **filter order or exponent which isn't particularly useful** in terms of dictating desired behavior + (2) can only **control one** filter characteristic + (3) *highly* inaccurate.



***Large* errors** in desired characteristic Q_3 for higher-order auditory filters

$$\beta = \frac{f}{CF}$$

(1) Specify the CF

$$Im\{s\} = i\beta$$

(2) Specify the exponent

Example transfer function:

$$H(s) = \left(\frac{1}{s^2 + \frac{1}{Q_{3,sos}}s + 1}\right)^n$$

(3) Specify $Q_{3,sos}$ of the corresponding second order system / the pole pair – this is *not* the Q_3 of the auditory filter

Figure – Error in Characteristic for Auditory Filters Designed using Naïve Method. To design filters using this method (regardless of the choice of filter form / transfer function), one specifies the CF, the exponent, and the $Q_{3,sos}$ of the related second order filter – which is used as a (very inaccurate) proxy for Q_3 of the auditory filter.

(B) Classical generic filter design methods and other iterative and optimization-based methods (often require specified frequency response or iterative tuning of filter constants and can't design filters based on characteristics) → entirely **unsuitable** for designing auditory filters

(C) Primary auditory filter design method (based on parameterization by filter order and either equivalent rectangular bandwidth **(ERB)** of gammatone filter (GTF) due to Parseval's theorem or Q_3 [Darling 1991]) → (1) parameterized by **filter order or exponent which isn't useful** in terms of directly dictating desired behavior + (2) generally **control one** filter characteristic (ERB or Q_3) + (3) **accuracy due to limitation of assumption** has not been rigorously studied + (4) **limited** to GTFs and filters for which we have an explicit approximate mapping between the filter coefficients and those of the GTF

3. Our Approach

- **(D) We develop methods for designing auditory filters directly and accurately given simultaneous specifications on filter characteristics.**
- Our **approach** involves deriving **parameterizations for the transfer function in terms of sets of filter characteristics** describing the auditory filter *itself* H(s; Ψ). This enables directly designing auditory filters by specifying values for its characteristics, Ψ.

< < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < < <

- To derive these characteristics—based parameterizations of the auditory filters, we do the following:

Step 1: Start with the original formulation of the transfer functions - parameterized by the filter constants Θ = [A_p, b_p, B_u]

Step 2: Derive expressions for the filter characteristics Ψ = [β_{peak}, Q_3, Q_{10}, Q_{ERB}, N, S, …] in terms of the filter constants Θ = [A_p, b_p, B_u]

Step 3: 'Invert' the above expressions to arrive at expressions for filter constants Θ in terms of sets of filter characteristics Ψ

Step 4: Replace the filter constants in the original transfer functions by the chosen set of filter characteristics to arrive at the transfer functions parameterized by filter characteristics Ψ

Figure – Schematic Diagram Illustrating Our Approach for deriving characteristics-based auditory filter parameterizations and developing characteristics-based methods for auditory filter design.

- We derive the characteristics-based parameterizations for two types of auditory filters GEFs/P (similar to APGF) and V (related to OZGF).

Step 1 Original transfer functions parameterized by filter constants[1]

$$\beta = \frac{f}{CF}$$

$$Im\{s\} = i\beta$$

GEFs/P auditory filters
$$P(s;\vec{\Theta}) = C\left(\frac{1}{s^2 + 2A_p s + A_p^2 + b_p^2}\right)^{B_u}$$

V auditory filters
$$V(s;\vec{\Theta}) = C\frac{s + A_p}{(s^2 + 2A_p s + A_p^2 + b_p^2)^{B_u}}$$

$$\vec{\Theta} = [A_p, b_p, B_u]$$
Three filter constants

$$(A_p, b_p, B_u) \in \mathbf{R}^+$$

$$p = -A_p + ib_p$$
$$\bar{p} = -A_p - ib_p$$
(repeated) pair of complex conjugate poles

4. Direct Characteristics-Based Design of Auditory Filters

- We identify the filter characteristics of interests and derive the expressions for the filter characteristics Ψ in terms of the filter constants Θ

Step 2 Derive expressions for characteristics in terms of filter constants [Alkhairy 2025 for expressions] [2]

Filter characteristics Ψ include:

$$\Psi : \begin{cases} \beta_{peak} & \text{peak normalized frequency - normalized to CF} \\ Q_n & \text{n dB quality factor (e.g. 3 dB, 15 dB)} \\ Q_{erb} & \text{quality factor based on the equivalent rectangular bandwidth} \\ N & \text{normalized group delay in the peak region} \\ S_\beta & \text{convexity at the peak} \\ \phi_{accum} & \text{phase accumulation} \\ |H|_{peak} & \text{magnitude at the peak} \\ \text{CF} & \text{characteristic frequency} \end{cases}$$

- We then 'invert' our expressions to derive expressions for filter constants Θ that originally parameterize the transfer functions in terms of filter characteristics Ψ that are specified when designing the auditory filters

Step 3 Invert for expressions for filter constants (which parameterize the original representations of the transfer functions) in terms of sets of filter characteristics (which we want to specify when designing auditory filters)

	Θ / Ψ subset	$\mathbf{b_p}$	$\mathbf{A_p}$	$\mathbf{B_u}$
II.1	$\beta_{peak}, N, \phi_{accum}$	β_{peak}	$\frac{\phi_{accum}}{\pi N}$	$2\phi_{accum}$
II.2	β_{peak}, N, Q_{erb}	β_{peak}	$= \frac{1}{2\pi N}B_u$ or, $= \frac{\beta_{peak}}{\sqrt{\pi}Q_{erb}}\frac{\Gamma(B_u)}{\Gamma(B_u - \frac{1}{2})}$	$\approx e^{\frac{b}{a}}\left(\frac{Q_{erb}}{\beta_{peak}N}\right)^{-\frac{1}{a}}$
II.3	$\beta_{peak}, Q_{erb}, \phi_{accum}$	β_{peak}	$\frac{\beta_{peak}}{\sqrt{\pi}Q_{erb}}\frac{\Gamma(2\phi_{accum})}{\Gamma(2\phi_{accum} - \frac{1}{2})}$	$2\phi_{accum}$
II.4	$\beta_{peak}, Q_n, \phi_{accum}$	β_{peak}	$\frac{\beta_{peak}}{2Q_n}\left(10^{\frac{n}{20\phi_{accum}}} - 1\right)^{-\frac{1}{2}}$	$2\phi_{accum}$
II.5	β_{peak}, S_β, N	β_{peak}	$\frac{40\pi}{\log(10)}\frac{N}{S_\beta}$	$\frac{80\pi^2}{\log(10)}\frac{N^2}{S_\beta}$
II.6	$\beta_{peak}, S_\beta, \phi_{accum}$	β_{peak}	$\sqrt{\frac{40\pi}{\log(10)}\frac{\phi_{accum}}{S_\beta}}$	$2\phi_{accum}$
II.7	β_{peak}, Q_n, N	β_{peak}	$\frac{1}{2\pi N}B_u$	solve the implicit equation $\frac{Q_n}{\beta_{peak}N} = \frac{\pi}{B_u}\left(10^{\frac{n}{10B_u}} - 1\right)^{-\frac{1}{2}}$

Table – Examples of Expressions for Filter Constants in Terms of Sets of Filter Characteristics. Table from Alkhairy (2015).

> >

- This allows us to simply express the transfer function in terms of desired sets of filter characteristics

Step 4 Transfer functions parameterized by filter characteristics (which are specified when designing auditory filters)

*Examples of transfer functions **parameterized by various sets of filter characteristics**:*

Parameterized by peak frequency, group delay at the peak, and phase accumulation

$$P(s;\vec{\Psi}_1) = C\left(s^2 + 2\frac{\phi_{accum}}{\pi N}s + \beta_{peak}^2 + \left(\frac{\phi_{accum}}{\pi N}\right)^2\right)^{-2\phi_{accum}}$$

Parameterized by peak frequency, group delay at the peak, and convexity at the peak

$$P(s;\vec{\Psi}_2) = C\left(s^2 + \frac{80\pi N}{\log(10)S_\beta}s + \beta_{peak}^2 + \left(\frac{40\pi N}{\log(10)S_\beta}\right)^2\right)^{-\frac{80\pi^2 N^2}{\log(10)S_\beta}}$$

- To design auditory filters, we simply and **directly** use the set of desired specifications on filter characteristics and 'plug them' into the expressions parameterized by filter characteristics (from step 4). Equivalently, we may convert the specifications on filter characteristics to specifications on filter constants (using the table in step 3), then construct the filters using the original expressions parameterized by filter constants (of step 1).

[1] Limitation (from step 1): $B_u \geq 1.5$ (can be rational) to result in impulse responses that mimic that of the cochlea

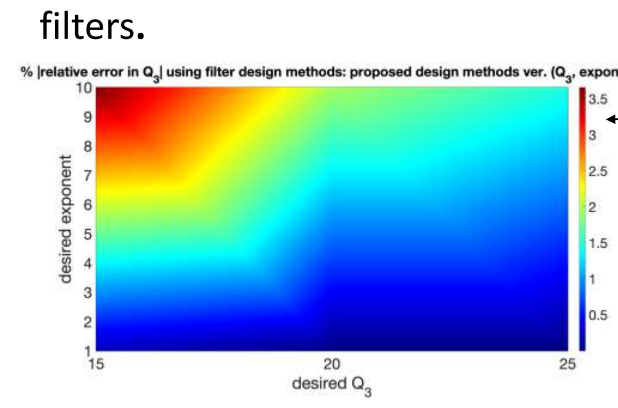
[2] Uses a sharp-filter approximation → limitation (from step 2): $A_p < 0.2$ for high accuracy

5. Choosing the Set of Characteristics for Design

- Our approach enables specification of various sets of characteristics depending on the needs of a particular study. For instance,
  a) It enables **simultaneous specification** of mixed magnitude- and phase-based characteristics (e.g., ERB and group delay), allowing for **sharply tuned responses without excessive group delay** and for control over both frequency **selectivity and synchronization** in filterbanks.
  b) It also enables **fine control over the shape** of the frequency response magnitude – e.g. through **simultaneous specification** of 3dB and 15dB quality factors.

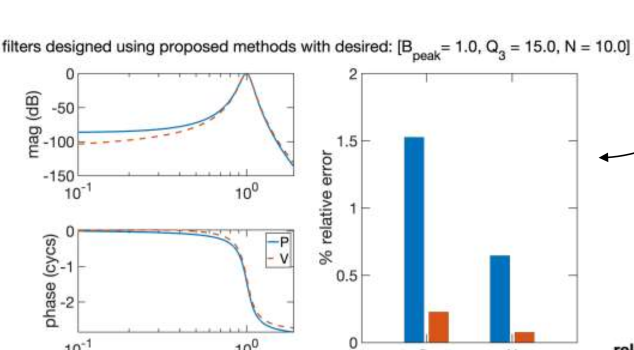
6. Accuracy

- Our methods for designing auditory filters given specifications on various sets filter characteristics are **highly accurate** for sharp filters.



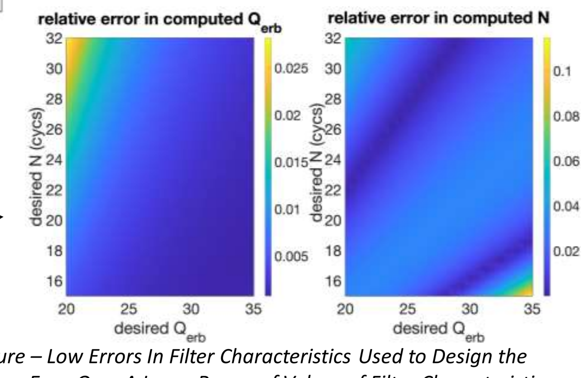
***Small* errors** (e.g. 3%) in desired characteristic Q_3 for higher-order auditory filters (compare with that of the naïve method).

Figure – Error in Characteristic for Auditory Filters Designed using Our Methods. We designed these GEFs/P filters by specifying the Q_3 directly, along with the filter exponent (used in order to compare against the accuracy of existing methods which use the filter order or exponent).



***Small* errors** (e.g. 1.5%) desired filter characteristics regardless of whether designing GEFs/P filters or V filters.

Figure – Left: Frequency Response of Filter Designed Given Specifications on Filter Characteristics for Two Types of Auditory Filters (GEFs/P, V). Right: Low Relative Error Between Desired and Achieved Filter Characteristics. We designed these filters by specifying the peak normalized frequency, 3 dB quality factor, and group delay at the peak



The **small errors** in achieving desired filter characteristics holds true over a large range of characteristics' values

Figure – Low Errors In Filter Characteristics Used to Design the Filters Even Over A Large Range of Values of Filter Characteristics. We designed these GEFs/P filters by specifying the peak normalized frequency, and a range of values for each of 3 dB quality factor and group delay at the peak (using the second option/expression to compute A_p)

References

- Alkhairy, S. A., & Shera, C. A. (2019). An analytic physically motivated model of the mammalian cochlea. The Journal of the Acoustical Society of America, 145(1), 45-60.
- * Darling, A. M. (1991). Properties and implementation of the gammatone filter: a tutorial. *Speech Hearing and Language, Work in Progress, University College London, Department of Phonetics and Linguistics*, 43-61.
- * Alkhairy, S. A. (2025). Characteristics-Based Design of Generalized-Exponent Bandpass Filters. arXiv preprint arXiv:2404.15321.
- Zhao, W., & Alkhairy, S. A. (2025) GEFs GitHub Repository.
- Warren, R. M., Bashford Jr, J. A., & Lenz, P. W. (2004). Intelligibility of bandpass filtered speech: Steepness of slopes required to eliminate transition band contributions. The Journal of the Acoustical Society of America, 115(3), 1292-1295.
- Dimitriadis, D., Maragos, P., & Potamianos, A. (2010). On the effects of filterbank design and energy computation on robust speech recognition. IEEE transactions on audio, speech, and language processing, 19(6), 1504-1516.
- Slaney, M., & Seltzer, M. L. (2014). The influence of pitch and noise on the discriminability of filterbank features. In INTERSPEECH (Vol. 14, pp. 2263-2267).
- Dietz, M., Ewert, S. D., & Hohmann, V. (2011). Auditory model based direction estimation of concurrent speakers from binaural signals. Speech Communication, 53(5), 592-605.
- Ghosh, P. K., Goldstein, L. M., & Narayanan, S. S. (2011). Processing speech signal using auditory-like filterbank provides least uncertainty about articulatory gestures. The Journal of the Acoustical Society of America, 129(6), 4014-4022.
- Cosentino, S., Falk, T. H., McAlpine, D., & Marquardt, T. (2013). Cochlear implant filterbank design and optimization: A simulation study. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 22(2), 347-353.
- Moore, B. C. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. Journal of the Association for Research in Otolaryngology, 9, 399-406.
- Irino, T., & Patterson, R. D. (2020). The gammachirp auditory filter and its application to speech perception. *Acoustical Science and Technology*, *41*(1), 99-107.
- Alkhairy, S. A. (2024). Cochlear wave propagation and dynamics in the human base and apex: Model-based estimates from noninvasive measurements. In AIP Conference Proceedings (Vol. 3062, No. 1). AIP Publishing.

7. Contributions and Properties of Characteristics-Based Design Methods

- The characteristics-based methods for direct design of auditory filters addresses **two critical needs**:
  a) Accurately designing auditory filters based on simultaneous specifications on characteristics
  b) Systematically studying how varying characteristics affects outcomes from perceptual studies and technological advances.
- The methods are entirely characteristics-based, accurate, direct, simple, inherently stable, computationally efficient, avoid iterative processes, and are appropriate for various species. To our knowledge, no existing methods exhibit this degree of accuracy, simplicity, and simultaneous control.

8. Future Directions and Areas for Collaboration

- The filter design methods may be used to **directly and accurately** design filterbanks.
- The methods may also be used to ***systematically* investigate** the dependence of perceptual and technological study outcomes on filter characteristics – e.g. by studying the dependence on **isolated** characteristics (by varying ERB while fixing group delay, or varying Q_{10} while fixing Q_{15}). This is motivated by **studies reporting sensitivity** based on (usually *ad hoc)* variation of certain parameters – e.g. ERB and filter order. The outcomes influenced by changing filter characteristics include:
  1. Intelligibility scores of bandpass-filtered speech [Warren et al 2004]
  2. Accuracy of speech recognition [Dimitriadis et al 2010; Slaney & Seltzer 2014]
  3. DOA / sound source localization models [Dietz et al 2011]
  4. Mutual information between articulatory gestures of vocal tracts and acoustic and perceptual features [Ghosh et al 2011]
  5. Accuracy of speech intelligibility models for cochlear implants [Cosentino et al 2013]
- This work may also be used to understand the cochlea's role in perception via underlying unified models [Alkhairy 2024].
- Our filter design methods may be extended to incorporate specifications on combined spectrotemporal characteristics as is relevant for studying certain perceptual functions [Moore 2008]. Future work should include considerations to handle nonlinear versions of the filters as is especially relevant for hearing loss simulations [Irino & Patterson 2020].