The Journal of Mathematical Neuroscience
a SpringerOpen Journal

Signal processing in the cochlea: the structure equations

Hans Martin Reimann

Abstract *Background:* Physical and physiological invariance laws, in particular time invariance and local symmetry, are at the outset of an abstract model. Harmonic analysis and Lie theory are the mathematical prerequisites for its deduction.
Results: The main result is a linear system of partial differential equations (referred to as the structure equations) that describe the result of signal processing in the cochlea. It is formulated for phase and for the logarithm of the amplitude. The changes of these quantities are the essential physiological observables in the description of signal processing in the auditory pathway.
Conclusions: The structure equations display in a quantitative way the subtle balance for processing information on the basis of phase versus amplitude. From a mathematical point of view, the linear system of equations is classified as an inhomogeneous $\bar{\partial}$-equation. In suitable variables the solutions can be represented as the superposition of a particular solution (determined by the system) and a holomorphic function (determined by the incoming signal). In this way, a global picture of signal processing in the cochlea emerges.

1 Background

At the outset of this work is the quest to understand signal processing in the cochlea.

HM Reimann (✉)
Institute of Mathematics, University of Berne, Sidlerstrasse 5, 3012 Berne, Switzerland
e-mail: martin.reimann@math.unibe.ch

1.1 Linearity and scaling

It has been known since 1992 that cochlear signal processing can be described by a wavelet transform (Daubechies 1992 [1], Yang, Wang and Shamma, 1992 [2]). There are two basic principles that lie at the core of this description: Linearity and scaling.

In the cochlea, an incoming acoustical signal $f(t)$ in the form of a pressure fluctuation (t is the time variable) induces a movement $u(x, t)$ of the basilar membrane at position x along the cochlea. At a fixed level of sound intensity, the relation between incoming signal and movement of the basilar membrane is surprisingly linear. However as a whole this process is highly compressive with respect to levels of sound - and thus cannot be linear.

In the present setting this is taken care of by a 'quasilinear model'. This is a model that depends on parameters, for example, in the present situation the level of sound intensity. For fixed parameters the model is linear. It is interpreted as a linear approximation to the process at these fixed parameter values. Wavelets give rise to linear transformations. The description of signal processing in the cochlea by wavelet transformations, where the wavelets depend on parameters, is compatible with this approach.

Scaling has its origin in the approximate local scaling symmetry (Zweig 1976 [3], Siebert 1968 [4]) that was revealed in the first experiments (Békésy 1947 [5], Rhode 1971 [6]).

The scaling law can best be formulated with the basilar membrane transfer function $\hat{g}(x, \omega)$. This is the transfer function that is defined from the response of the linear system to pure sounds. To an input signal

$$\cos(\omega t) = \operatorname{Re} e^{i\omega t}, \quad \omega > 0, \tag{1}$$

that is, to a pure sound of circular frequency ω there corresponds an output $u(x, t)$ at the position x along the cochlea that on the basis of linearity has to be of the form

$$u(x, t) = \operatorname{Re}\left\{\hat{g}(x, \omega)e^{i\omega t}\right\}. \tag{2}$$

The basilar membrane transfer function is thus a complex valued function of x and $\omega > 0$. Its modulus $|\hat{g}(x, \omega)|$ is a measure of amplification and its argument is the phase shift between input and output signals. The experiments of von Békésy [5] showed that the graphs of $|\hat{g}(x, \omega)|$ and $|\hat{g}(x, c\omega)|$ as functions of the variable x are translated against each other by a constant multiple of $\log c$. By choosing an appropriate scale on the x-axis, the multiple can be taken to be 1. The scaling law is then expressed as

$$\left|\hat{g}(x - \log c, c\omega)\right| = \left|\hat{g}(x, \omega)\right|. \tag{3}$$

The scaling law will be extended - with some modifications - to include the argument of $\hat{g}$.

Intimately connected to scaling is the concept of a tonotopic order. It is a central feature in the structure of the auditory pathway. Frequencies of the acoustic signal are associated to places, at first in the cochlea and in the following stages in the various neuronal nuclei. The assignment is monotone, it preserves the order of the

frequencies. In the cochlea, to each position x along the cochlear duct a circular frequency $\sigma = \xi(x)$ is assigned. The function ξ is the position-frequency map. Its inverse is called the tonotopic axis. At the stand of von Békésy's results, the frequency associated to a position x along the cochlea is simply the best frequency (BF), that is the frequency σ at which $|\hat{g}(x, \omega)|$ attains its maximum. The refined concept takes care of the fact that the transfer function and with it the BF changes with the level of sound intensity, at which $\hat{g}$ is determined. The characteristic frequency (CF) is then the low level limit of the best frequency. The position-frequency map ξ assigns to the position x its CF.

Scaling according to von Békésy's results implies the exponential law

$$\xi(x) = Ke^{-x} \tag{4}$$

for the position-frequency map. The constant K is determined by inserting a special value for x. The scaling law tells us that the function $|\hat{g}(x, \omega)|$ is actually a function of the 'scaling variable'

$$\frac{1}{K}\omega e^x = \frac{\omega}{\xi(x)}. \tag{5}$$

At the outset of the present investigation it will be assumed that the transfer function $\hat{g}$ is a function of the scaling variable $\frac{\omega}{\xi(x)}$. This is not strictly true, but it simplifies the exposition. In subsequent sections a general theory will be developed that incorporates quite general scaling behavior. With the availability of advanced experimental data (Rhode 1971 [6], Kiang and Moxon 1974 [7], Liberman 1978 [8], 1982 [9], Eldredge *et al.* 1981 [10], Greenwood 1990 [11]), the position-frequency map is now known precisely for many species. Shera 2007 [12] gives the formula

$$CF(x) = [CF(0) + CF_1]e^{x/l} - CF_1. \tag{6}$$

The constant l and the 'transition frequency' CF_1 vary from species to species. The scaling variable that goes with it is

$$v(x, f) = \frac{f + CF_1}{CF(x) + CF_1}. \tag{7}$$

In the present setting, x is the normalized variable (x instead of x/l) and the precise position-frequency map is expressed in the form

$$\xi(x) = Ke^{-x} - S. \tag{8}$$

ξ denotes circular frequency and $K = \xi(0) + S$. The constant S is referred to as the shift.

In the abstract model as it will be developed, much will depend on the definition of the function σ that specifies the frequency location. In the present treatment the frequency localization of a function will be defined as an expectation value in the frequency domain.

1.2 Wavelets

The response to a general signal $f(t)$ with Fourier representation

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\, dt \tag{9}$$

is given as

$$u(x,t) = \sqrt{\frac{2}{\pi}}\, \mathrm{Re} \int_{0}^{\infty} \hat{f}(\omega)\hat{g}(x,\omega)e^{i\omega t}\, d\omega. \tag{10}$$

Note that the Fourier transform of the real valued signal f satisfies $\hat{f}(\omega) = \overline{\hat{f}(-\omega)}$. If the definition of $\hat{g}$ is extended to negative values of ω by $\hat{g}(x,-\omega) = \overline{\hat{g}(x,\omega)}$ then $u(x,t)$ can be written as

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega)\hat{g}(x,\omega)e^{i\omega t}\, d\omega. \tag{11}$$

The transfer function will be described by a function h in the scaling variable:

$$\hat{g}(x,\omega) = h\left(\frac{\omega}{\xi(x)}\right).$$

The response of the cochlea to a general signal f can then be expressed as

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega)h\left(\frac{\omega}{\xi(x)}\right)e^{i\omega t}\, d\omega.$$

Setting $a = \frac{1}{\xi(x)} = \frac{1}{K}e^{x}$ and thus $x = k + \log a$ with $k = \log K$, the scaling function is simply $h(a\omega)$. This leads to the equivalent formulation

$$u(k + \log a, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega)h(a\omega)e^{i\omega t}\, d\omega. \tag{12}$$

This is recognized as a wavelet transform. Indeed, with the standard L^2-normalization a wavelet transform Wf with wavelet ψ is defined by

$$Wf(a,t) = \int_{-\infty}^{\infty} f(s)\frac{1}{\sqrt{a}}\overline{\psi\left(\frac{s-t}{a}\right)}\, ds$$

$$= \int_{-\infty}^{\infty} \hat{f}(\omega)\sqrt{a}\,\overline{\hat{\psi}(a\omega)}e^{i\omega t}\, d\omega.$$

If $\frac{1}{\sqrt{2\pi}}h(\omega)$ is identified with $\overline{\hat{\psi}(\omega)}$ then

$$u(x,t) = u(k + \log a, t) = \frac{1}{\sqrt{a}}Wf(a,t). \tag{13}$$

The fact, that the cochlea - in a first approximation - performs a wavelet transform appears in the literature in 1992, both in [1] and in [2].

1.3 Uncertainty principle

The natural symmetry group for signal processing in the cochlea is built on the affine group Γ. It derives from the scaling symmetry in combination with time-invariance. In addition, there is the circle group S that is related to phase shifts. Its action commutes with the action of the affine group. The full symmetry group for hearing is thus $\Gamma \times S$. For this group, the uncertainty principle can be formulated. The functions for which equality holds in the uncertainty inequalities are called the extremal functions. They play a special role, similar as in quantum physics the coherent states (the extremals for the Heisenberg uncertainty principle). The starting point in the present work is the tenet that these functions provide an approximation for the cochlear transfer function.

That the extremal functions should play a special role is not a new idea. In signal processing the extremal functions first appeared in Gabor's work (1946) [13] in connection with the Heisenberg uncertainty principle and then in Cohen's paper (1993) [14] in the context of the affine group. In a paper by Irino 1995 [15] the idea is taken up in connection with signal processing in the cochlea. It is further developed by Irino and Patterson [16] in 1997. The presentation in this paper is based on previous work (Reimann, 2009 [17]). The concept pursued is to determine the extremals in the space of real valued signals and to use a setup in the frequency domain, not in the time domain. Different representations of the affine group give different families E_c of extremal functions. The parameter c is used to adjust to the sound level and hence to provide linear approximations at different levels to the non-linear behavior of cochlear signal processing.

2 Results and discussion

2.1 Uncertainty principle

This section starts with the specification of the symmetry group $\Gamma \times S$ that underlies the hearing process. The basic uncertainty inequalities for this group are then explicitly derived. The analysis builds on previous results (Reimann [17]). A modification is necessary because the treatment of the phase in [17] was not satisfactory. An improvement can be achieved with the inclusion of the term $\alpha \hat{H}$ in the uncertainty inequality. This term comes in naturally and it will influence the argument - but not the modulus - of the extremal functions associated to the uncertainty inequalities. It is claimed that the extremal functions derived in this section are a first approximation to the basilar membrane transfer function $\hat{g}$. The extremal functions for the basic uncertainty principle are interpreted as the transfer function at high levels of sound. This situation corresponds to the parameter value $c = 1$. With increasing parameter values the extremal functions for the general uncertainty inequality are then taken as approximations to the cochlear response at decreasing levels of sound.

2.1.1 The symmetry group

The affine group Γ is the group of affine transformations of the real line $\mathbf{R}$. It is generated by the transformation group $\tau_b(t) = t + b$ ($b \in \mathbf{R}$) and the dilation group

$\delta_a(t) = at$ ($a \in \mathbf{R}$, $A \neq 0$). Under the Fourier transform, the action of the dilation group on $L^2(\mathbf{R}, \mathbf{C})$ is intertwined to the action of the inverse dilation group $\hat{\delta}$. This group also acts directly in frequency space:

$$\hat{\delta}_a(\omega) = \frac{\omega}{a}. \tag{14}$$

The induced unitary action on $L^2(\mathbf{R}, \mathbf{C})$ is

$$\hat{\delta}_a h(\omega) = \sqrt{a} h\big(\hat{\delta}_a^{-1}(\omega)\big) = \sqrt{a} h(a\omega). \tag{15}$$

(With this convention, the group action and the induced action are denoted with the same symbol.) Clearly, the invariance property of the basilar membrane transfer function directly reflects this group action.

The action

$$\tau_b(f) = f(t - b)$$

of the translation group intertwines under the Fourier transform to the unitary action

$$\hat{\tau}_b h(\omega) = e^{-ib\omega} h(\omega). \tag{16}$$

Of relevance to our considerations is the space $L^2(\mathbf{R}, \mathbf{R})$ of real valued signals of finite energy. Under the Fourier transform it is mapped onto

$$L^2_{sym}(\mathbf{R}, \mathbf{C}) = \big\{ h \in L^2(\mathbf{R}, \mathbf{C}) : h(\omega) = \overline{h(-\omega)} \big\}. \tag{17}$$

Both $\hat{\delta}$ and $\hat{\tau}_b$ act on L^2_{sym}. The only action that commutes with both of them is the action

$$\hat{\varepsilon}_\varphi h(\omega) = e^{-i\varphi \, \mathrm{sgn}(\omega)} h(\omega) \tag{18}$$

of the circle group S. This is the third distinguished group action.

The infinitesimal operators associated with the unitary actions of $\hat{\delta}$, $\hat{\tau}_b$ and $\hat{\varepsilon}$ are the skew hermitian operators

$$\hat{A} \hat{f}(\omega) = \frac{d}{ds}\bigg|_{s=0} \hat{\delta}(e^s) \hat{f}(\omega) = \frac{d}{ds}\bigg|_{s=0} e^{s/2} \hat{f}(e^s \omega) \tag{19}$$

$$= \frac{1}{2} \hat{f}(\omega) + \omega \frac{d\hat{f}}{d\omega}(\omega), \tag{20}$$

$$\hat{B} \hat{f}(\omega) = \frac{d}{db}\bigg|_{b=0} e^{-ib\omega} \hat{f}(\omega) = -i\omega \hat{f}(\omega), \tag{21}$$

$$\hat{H} \hat{f}(\omega) = \frac{d}{d\varphi}\bigg|_{\varphi=0} e^{-i\varphi \, \mathrm{sgn}(\omega)} \hat{f}(\omega) = -i \, \mathrm{sgn}(\omega) \hat{f}(\omega). \tag{22}$$

The commutator relations are

$$[\hat{A}, \hat{B}] = \hat{B}, \tag{23}$$

$$[\hat{A}, \hat{H}] = [\hat{B}, \hat{H}] = 0. \tag{24}$$

The operators $\hat{A}$, $\hat{B}$ and $\hat{H}$ span the Lie algebra of the 'hearing group' $\Gamma \times S$, with Γ the affine group and S the circle group. The basic variables in cochlear signal processing are time t and position x along the cochlea. Clearly $\hat{B}$ is related to time, whereas $\hat{A}$ - as will be shown presently - is related to the position. In our approach the tonotopic axis is given by the exponential law

$$\xi(x) = Ke^{-x}.$$

Under the tonotopic axis dilations $\hat{\delta}_a$ are conjugated to translations (by $\log a$) in x:

$$\tau_{\log a} = \xi^{-1} \circ \hat{\delta}_a \circ \xi.$$

Here, $\xi^{-1}(\omega) = -\log \frac{\omega}{K}$ is the inverse function with respect to the composition law. The intertwining action

$$\xi h(\omega) = \frac{1}{\sqrt{\omega}} h(\xi^{-1}\omega), \quad \omega > 0,$$

is an isometry in the sense that for all $h \in L^2(\mathbf{R}, \mathbf{C})$

$$\int_0^\infty |\xi h(\omega)|^2 \, d\omega = \int_0^\infty |h(\xi^{-1}\omega)|^2 \frac{d\omega}{\omega} = \int_{-\infty}^\infty |h(x)|^2 \, dx.$$

It intertwines $\hat{A}$ with $-\frac{d}{dx}$:

$$\hat{A}\xi = -\xi \frac{d}{dx} \tag{25}$$

as the following calculation shows:

$$\hat{A}\xi h(\omega) = \frac{1}{2}\xi h(\omega) + \omega \frac{d}{d\omega}\left(\frac{1}{\sqrt{\omega}} h(\xi^{-1}\omega)\right)$$

$$= \frac{1}{2}\xi h(\omega) - \frac{1}{2}\xi h(\omega) + \frac{\omega}{\sqrt{\omega}} \frac{dh}{dx}(\xi^{-1}\omega) \frac{d\xi^{-1}}{d\omega}$$

$$= -\frac{1}{\sqrt{\omega}} \frac{dh}{dx}(\xi^{-1}\omega) = -\xi\left(\frac{dh}{dx}\right)(\omega).$$

The uncertainty principle that goes with the group $\Gamma \times S$ can thus been seen as an uncertainty for the determination of time and position.

2.1.2 The basic uncertainty inequality

The commutator relation

$$[\hat{A}, \hat{B}] = \hat{B} \tag{26}$$

is at the basis of the uncertainty principle for the affine group. From the inequality

$$0 \le \|\hat{A}h + \kappa \hat{H}\hat{B}h\|^2$$
$$= \|\hat{A}h\|^2 + \kappa(\hat{A}h, \hat{H}\hat{B}h) + \kappa(\hat{H}\hat{B}h, \hat{A}h) + \kappa^2\|\hat{H}\hat{B}h\|^2$$
$$= \|\hat{A}h\|^2 + \kappa(h, [\hat{H}\hat{B}, \hat{A}h]h) + \kappa^2\|\hat{H}\hat{B}h\|^2,$$

that has to hold for all $\kappa \in \mathbf{R}$, it follows that

$$\left|(h, \hat{H}\hat{B}h)\right| \le 2\|\hat{A}h\|\|\hat{H}\hat{B}h\|.$$

In this calculation, the operators $\hat{A}$ and $\hat{B}$ can be replaced by $\hat{A} - \alpha\hat{H} - \beta\hat{B}$ and $\hat{B} - v\hat{H}$ respectively. This leads to the new inequality

$$\left|(h, \hat{H}\hat{B}h)\right| \le 2\left\|(\hat{A} - \alpha\hat{H} - \beta\hat{B})h\right\|\left\|(\hat{H}\hat{B} + v)h\right\|. \tag{27}$$

This inequality is of the same nature as the previous inequality. It can be considered as a more precise inequality, because it holds for all parameter values of α, β and v. The expression $\|(\hat{H}\hat{B} + v)h\|$ is minimal for

$$v = v(h) = -\frac{(h, \hat{H}\hat{B}h)}{\|h\|^2} = \frac{1}{\|h\|^2} \int_{-\infty}^{\infty} |\omega|\left|h(\omega)\right|^2 d\omega. \tag{28}$$

This v is the decisive parameter. It has the interpretation of an expectation value for the frequency. Later it will be associated with the place along the cochlea.

The uncertainty inequality can thus be stated as

$$v\|h\|^2 \le 2\left\|(\hat{A} - \alpha\hat{H} - \beta\hat{B})h\right\|\left\|(\hat{H}\hat{B} + v)h\right\|. \tag{29}$$

The minimality condition for the parameters α and β in the expression

$$\left\|(\hat{A} - \alpha\hat{H})h - \beta\hat{B}\right\|$$

is given by the linear system

$$\alpha\|h\|^2 - \beta(h, \hat{H}\hat{B}h) + (h, \hat{H}\hat{A}h) = 0, \tag{30}$$

$$-\alpha(h, \hat{H}\hat{B}h) + \beta\|\hat{B}h\|^2 - \mathrm{Re}(\hat{A}h, \hat{B}h) = 0. \tag{31}$$

The coefficients are

$$(h, \hat{H}\hat{B}h) = -v\|h\|^2,$$
$$(h, \hat{H}\hat{A}h) = -(\hat{A}h, \hat{H}h)$$
$$= -\int_{-\infty}^{\infty} \left(\frac{h}{2} + \omega\frac{dh}{d\omega}\right)i\,\mathrm{sgn}(\omega)\overline{h}\,d\omega$$
$$= \mathrm{Re}\int_{-\infty}^{\infty} i|\omega|h'\overline{h}\,d\omega = \frac{1}{2}\int_{-\infty}^{\infty} i|\omega|(h\overline{h}' - \overline{h}h')\,d\omega$$
$$= \int_{-\infty}^{\infty} |h|^2|\omega|\frac{d}{d\omega}\arg h\,d\omega.$$

In this calculation, the fact that

$$\frac{d}{d\omega} \arg h = \frac{1}{2i} \frac{d}{d\omega}(\log h - \log \bar{h}) = \frac{1}{2i}\left(\frac{h'}{h} - \frac{\bar{h}'}{\bar{h}}\right)$$

has been used.

The remaining coefficient is

$$\text{Re}(\hat{A}h, \hat{B}h) = \int_{-\infty}^{\infty}\left(\frac{h}{2} + \omega h'\right) i\omega \bar{h}\, d\omega = \int_{-\infty}^{\infty}|\omega|^2 \frac{i}{2}\left(\frac{h'}{h} - \frac{\bar{h}'}{\bar{h}}\right) d\omega$$

$$= \int_{-\infty}^{\infty} |h|^2 |\omega|^2 \frac{d}{d\omega} \arg h\, d\omega.$$

With $h = \hat{f}$ a different meaning can be given to it:

$$\text{Re}(\hat{A}\hat{f}, \hat{B}\hat{f}) = \text{Re}\int_{-\infty}^{\infty}\left(-\frac{f(t)}{2} - t\frac{df}{dt}(t)\right)\left(-\frac{d\bar{f}}{dt}(t)\right) dt$$

$$= \int_{-\infty}^{\infty} t\left|\frac{df}{dt}(t)\right|^2 dt.$$

The integrals $\int_{-\infty}^{\infty}|h|^2|\omega|\frac{d}{d\omega}\arg h\, d\omega$ and $\int_{-\infty}^{\infty} t|\frac{df}{dt}(t)|^2\, dt$ can be interpreted as expectation values of $|h|^2\frac{d}{d\omega}\arg h$ in the frequency space and for $|\frac{df}{dt}(t)|^2$ in the time space. Roughly, in combination with $\hat{H}$ the operator $\hat{A}$ controls $\frac{d}{d\omega}\arg h$ and $\hat{B}$ the time derivative.

We will assume that the parameters α, β and ν are always chosen such that the right hand side in the uncertainty inequality is minimal, that is, the inequality is formulated in its sharpest form.

The mean deviation from the expectation value ν for the modulus of the frequency is

$$\tau^2 = \frac{\|(HB + \nu I)\hat{f}\|^2}{\|\hat{f}\|^2} = \frac{1}{\|f\|^2}\int_{-\infty}^{\infty}\left(|\omega| - \nu\right)^2 |\hat{f}(\omega)|^2\, d\omega. \qquad (32)$$

The factor

$$\left\|(\hat{A} - \alpha\hat{H} - \beta\hat{B})h\right\|$$

does not have such a simple interpretation except in the special case $\alpha = 0$. This is treated in [17].

A function h is called extremal, if equality holds for it in the uncertainty relation. The extremal functions are expected to play a special role in the signal processing of the cochlea. In the context of the classical Heisenberg uncertainty relation, the extremal functions are translates of the Gaussian function e^{-x^2} under the action of the Heisenberg group. They are called 'coherent states'. Their significance in signal processing is well established since the appearance of Gabor's work in 1946 [13]. At the outset of the present discussion is however the fact that the cochlea performs a wavelet transform - and not a Fourier transform. The invariance group is $\Gamma \times S$ and

not the Heisenberg group. It should therefore be expected that the extremal functions as discussed below play the crucial role in the hearing process.

The extremal functions h (in frequency space) satisfy the equation

$$(\hat{A} - \alpha \hat{H} - \beta \hat{B})h + \kappa(\hat{H}\hat{B} + \nu)h = 0. \tag{33}$$

This is in fact a differential equation:

$$\left(\frac{1}{2} + \omega \frac{d}{d\omega}\right)h = \left(-i\alpha \operatorname{sgn}(\omega) - i\beta\omega - \kappa|\omega| + \kappa\nu\right)h. \tag{34}$$

The solutions are

$$h(\omega) = ke^{i\varepsilon \operatorname{sgn}(\omega) - i\alpha \operatorname{sgn}(\omega)\log|\omega| - i\beta\omega}e^{-\kappa|\omega|}|\omega|^{\kappa\nu - \frac{1}{2}}, \tag{35}$$

with real constants k, ε, α, β, κ and ν. Square integrability implies $\kappa > 0$ and ν is the positive frequency expectation value.

From the explicit form it is clear that the space of solutions is invariant under the action of $\Gamma \times S$. The tenet is now:

2.2 The basilar membrane transfer function is given by extremal functions

To be more precise, there exist an extremal function h, normalized by the condition $\nu(h) = 1$, such that $h(\frac{\omega}{\xi})$ adequately describes the basilar membrane transfer function $\hat{g}$:

$$\hat{g}(x, \omega) = h\left(\frac{\omega}{\xi}\right) = ke^{i\varepsilon \operatorname{sgn}(\omega) - i\alpha \operatorname{sgn}(\omega)\log|\frac{\omega}{\xi}| - i\beta\frac{\omega}{\xi}}e^{-\kappa|\frac{\omega}{\xi}|}\left|\frac{\omega}{\xi}\right|^{\kappa - \frac{1}{2}}. \tag{36}$$

In this formula, $\xi = \xi(x)$ is the position-frequency map. Note further that

$$h\left(\frac{\omega}{\xi}\right) = h \circ \hat{\delta}_\xi(\omega) = \sqrt{\xi}\hat{\delta}_\xi^{-1}h(\omega)$$

such that

$$\nu(h \circ \hat{\delta}_\xi) = \nu(\hat{\delta}_\xi^{-1}h) = \hat{\delta}_\xi^{-1}(\nu(h)) = \xi. \tag{37}$$

The frequency expectation of $\hat{g}(x, \omega)$ at x is thus $\xi(x)$. The question then arises whether the experiments confirm the tenet. To arrive at a preliminary conclusion, graphs of the modulus and of the real part of the function h are displayed in Figure 1. The parameters are $\alpha = -\pi$, $\beta = 2\pi$ and $\kappa = 4$.

The classical results by von Békésy (1947) [5] seem to be in favor of such a statement. However the situation is of course not so simple. The basic problem is the non-linearity of the process that associates the movement $u(x, t)$ of the basilar membrane to the input signal $f(t)$. This process is highly compressive and therefore its description by a transfer function can at best be looked at as an approximation.

Von Békésy's result stem from experiments on dead animals. The outcome can be compared to the experimental results obtained with life animals, yet at high intensities
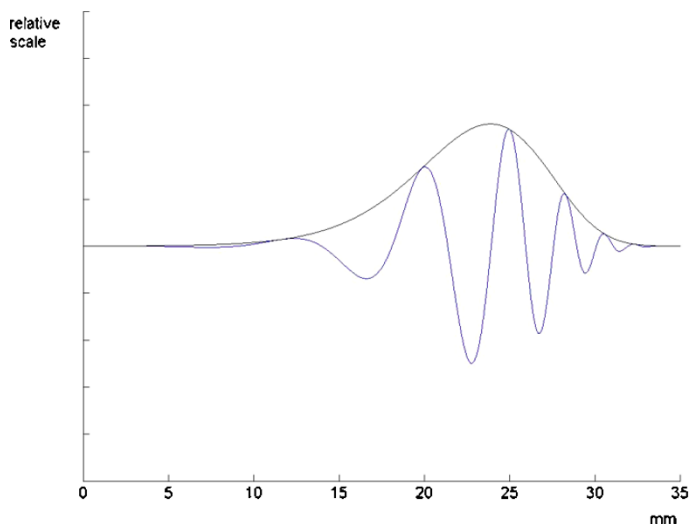
Fig. 1 The extremal function on a relative scale. The real part and the modulus of the extremal function h to the basic uncertainty principle ($c = 1$). In the drawings, the parameters are distance d in mm from the stapes ($x = \frac{d}{l} = \frac{d}{6.6}$) and frequency $f = \frac{\omega}{2\pi}$ in Hz. The extremal function is shown for a fixed frequency f as a function of the distance d (in mm) to the stapes. The parameters are: $\varepsilon = 2\pi$, $\alpha = -\pi$, $\beta = 2\pi$, $\kappa = 4$.

of sound pressure. The above description of the basilar membrane transfer function is therefore taken to be a linear approximation at high levels of sound pressure. In the following section the approach will be modified with the aim of obtaining linear approximations at all levels of sound pressure.

2.2.1 General uncertainty inequality for $\Gamma \times S$

There are various ways that the abstract group $\Gamma \times S$ can act on the space L^2_{sym}. Apart from the natural representation that associates to the corresponding basis elements in the Lie algebra of Γ the operators $\hat{A}$ and $\hat{B}$, the general representation considered below is built on the operators $\frac{1}{c}\hat{A}$ and $\hat{B}_c = -i|\omega|^c \, \text{sgn}(\omega)$. The representation of Γ induced by this algebra representation retains the crucial scaling behavior known from the experimental results. It seems to be suitable in the present context, despite the fact that the operator $\hat{B}_c$ does not stand for the time derivative any more.

The general representation of $\Gamma \times S$ on the space L^2_{sym} that will be considered is determined by the representation of its Lie algebra and as such by the operators $\frac{1}{c}\hat{A}$, $\hat{B}_c = -i|\omega|^c \, \text{sgn}(\omega)$ and $\hat{H}$. There is the single non trivial commutator relation:

$$\left[\frac{1}{c}\hat{A}, \hat{B}_c\right] = \hat{B}_c. \tag{38}$$

The uncertainty inequality that goes with it is

$$\left|(h, \hat{H}\hat{B}_c h)\right| \leq \frac{2}{c}\left\|(\hat{A} - \alpha\hat{H} - \beta\hat{B})h\right\|\left\|(\hat{H}\hat{B}_c + \nu)h\right\|. \tag{39}$$

At this point it is not clear whether the term $\frac{1}{c}(\hat{A} - \alpha\hat{H} - \beta\hat{B})$ should rather be $\frac{1}{c}(\hat{A} - \alpha\hat{H} - \beta\hat{B}_c)$. In fact the inequality is true for both variants and in both cases, families of inequalities depending on the parameters are obtained. The question is how the extremal functions that are associated to these inequalities vary with the parameters. Yet by choosing the present version one finds that the set of extremal functions is invariant under the action of $\Gamma \times S$. The expression $\alpha\hat{H} + \beta\hat{B}$ should been seen as the linear approximation of the skew hermitian operator $\hat{A}$. As a result of using both $\hat{B}_c$ and $\hat{B}$, the first as an operator and the latter as an approximation term, this has the effect, that the argument of the extremal function appears in a slightly different context than the modulus. The extremal functions are obtained from the relation

$$\lambda\frac{1}{c}(\hat{A} - \alpha\hat{H} - \beta\hat{B})h + \mu(\hat{H}\hat{B}_c + \nu)h = 0.$$

They satisfy the differential equation

$$\left(\frac{1}{2} + \omega\frac{d}{d\omega}\right)h = \left(-i\alpha\,\text{sgn}(\omega) - i\beta\omega - \kappa|\omega|^c + \kappa\nu\right)h. \tag{40}$$

The proportionality factor is $\kappa = -\frac{c\mu}{\lambda}$. Its choice is arbitrary. All the constants α, β, and κ can in fact be chosen in dependence of the parameter c. This gives a possibility for fine adjustment of the extremal function h_c that describes the linear approximation at level c of the basilar membrane filter.

The solutions of the equation are

$$h_c(\omega) = ke^{i\varepsilon\,\text{sgn}(\omega) - i\alpha\,\text{sgn}(\omega)\log|\omega| - i\beta\omega}e^{-\frac{\kappa}{c}|\omega|^c}|\omega|^{\kappa\nu - \frac{1}{2}}, \tag{41}$$

with real constants k, ε, α, β, κ and ν. These solutions are in L^2 if both $\kappa > 0$ and $\nu > 0$.

The parameter in the uncertainty inequality is

$$\nu = \nu_c(h) = -\frac{(h, \hat{H}\hat{B}_c h)}{\|h\|^2} = \frac{1}{\|h\|^2}\int_{-\infty}^{\infty}|\omega|^c|h(\omega)|^2\,d\omega. \tag{42}$$

This time, the frequency localization of the function h is

$$\nu^{\frac{1}{c}}(h) = \left(\frac{1}{\|h\|^2}\int_{-\infty}^{\infty}|\omega||h(\omega)|^2 d\omega\right)^{\frac{1}{c}} \tag{43}$$

and

$$\nu^{\frac{1}{c}}(h \circ \hat{\delta}_a) = a\nu^{\frac{1}{c}}(h). \tag{44}$$

In accordance with our tenet, the basilar membrane filter is described as

$$\hat{g}(x, \omega) = h_c \circ \hat{\delta}_\xi(\omega), \tag{45}$$

with an extremal function h_c, normalized by the condition $\nu(h_c) = 1$.

$$\hat{g}(x,\omega) = h_c\left(\frac{\omega}{\xi}\right) = k e^{i\varepsilon\,\mathrm{sgn}(\omega) - i\alpha\,\mathrm{sgn}(\omega)\log\left|\frac{\omega}{\xi}\right| - i\beta\frac{\omega}{\xi}} e^{-\frac{\kappa}{c}\left|\frac{\omega}{\xi}\right|^c} \left|\frac{\omega}{\xi}\right|^{\kappa - \frac{1}{2}}. \qquad (46)$$

As before, $\xi = \xi(x)$ is the tonotopic axis. The frequency localization of $\hat{g}(x,\omega)$ as a function of ω is ξ:

$$\nu^{\frac{1}{c}}\big(\hat{g}(x,\cdot)\big) = \xi(x). \qquad (47)$$

The parameter c allows to express at which level of sound intensity the linearization is specified. Parameters $c \sim 1$ indicate high levels and parameters $c \gg 1$ small levels of intensity.

Experimental results on the basilar membrane transfer function are reviewed in Robles and Rugggero, 2001 [18]. As already pointed out in the previous paper 2009 [17], the shape of the modulus of the transfer function determined at various intensities as given by Rhode and Recio (2000 [19], Figure 1C) is approximately described by the modulus of the extremal functions at the corresponding parameter values. In particular, at a fixed position, the modulus of the transfer function has its peak below the position of the frequency localization and with decreasing intensity of sound level it approaches this position. Similarly, for fixed x, the extremal functions $|\hat{g}_c(x,\omega)|$ attain their maxima at

$$\omega = \xi(x)\left(1 - \frac{1}{2\kappa}\right)^{\frac{1}{c}}. \qquad (48)$$

With increasing values of c this approaches the frequency localization $\xi(x)$.

With the present setup the argument of the basilar membrane filter is independent of c. The experimental results by Rhode and Recio (2000) show minor changes of phase in dependence of the intensity level. With increasing intensity there is a small phase lag below the characteristic frequency and an equally small phase lead for frequencies above the characteristic frequency. Studies of the impulse response also confirm that the phase is almost invariant under changes in sound level (Recio and Rhode (2000) [20], Shera (2001) [21]). In order to obtain a fine adjustment of the phase data, the parameters α and β would have to be chosen in dependence of c.

In any case, the phase function has to be a decreasing function both when considered as a function of the frequency ω and as a function of the place x. The phase of the extremal function does not satisfy this requirement because of the logarithmic term. Yet still, the phase of the extremal function serves as an approximation of the physiological phase function on the interval in which the the absolute value is relevant. At places at which the absolute value is close to zero, the argument is of no significance. In Figure 2 the phase function is pictured for the fixed circular frequency $\omega = 1,000$ as a function of the distance d to the stapes, on the interval that is of physiological relevance. In Figure 3 the phase is pictured as a function of frequency (in Hz). In this figure, the characteristic frequency is 7,000 Hz. The part above about 3,000 Hz is of physiological relevance. The approximation holds in this range. It should be compared with the experimental results by Rhode and Recio (2000 [19], Figure 2E).
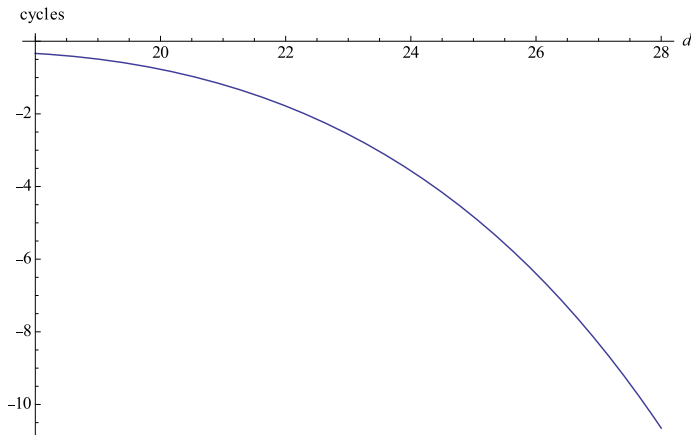
Fig. 2 Phase as a function of d. The phase (in cycles) of the extremal function h as a function of the distance d (in mm) to the stapes. The frequency is 500 Hz. Under the tonotopic axis this value corresponds to $d = 24.4$ mm. The parameters are: $\varepsilon = 16\pi$, $\alpha = -7.8\pi$, $\beta = 24\pi$.

The part below 3,000 Hz is the mathematical expression for the phase function. It is physiologically not correct, but this is of no significance.

The factor $e^{-i\beta\omega}$ in the extremal functions h_c stands for a pure time delay by β. Dividing the extremal functions by this factor, one is left with the extremal functions as they would appear if β had been set equal to zero in the general uncertainty
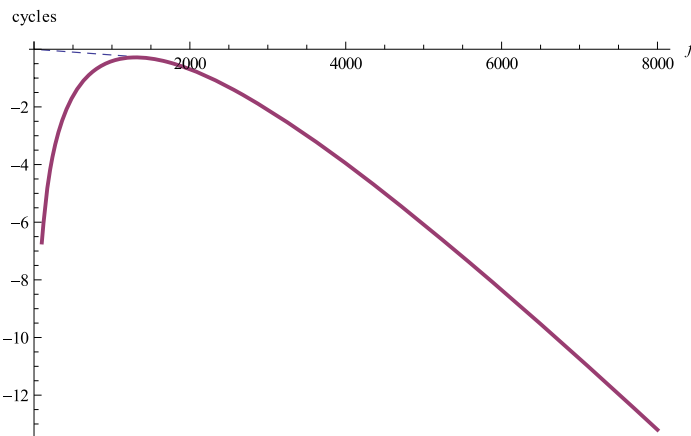
Fig. 3 Phase as a function of f. The phase (in cycles) of the extremal function h as a function of frequency $f = \frac{\omega}{2\pi}$ at the distance $d = 10.6$ mm to the stapes. The solid line is the phase as determined by the extremal function. The dashed line is the physiologically correct substitute at low frequencies. Note that the phase values in this region are practically irrelevant for signal processing, because the amplitude values in this region are negligible. Under the tonotopic axis d corresponds to the frequency $f = 4,000$ Hz. The parameters are the same as in Figure 2.

inequality. They are the extremal functions for the uncertainty inequality

$$v_c \|h\|^2 \leq \frac{2}{c} \left\| (\hat{A} - \alpha \hat{H}) h \right\| \left\| (\hat{H} \hat{B}_c + v_c) h \right\|. \tag{49}$$

2.3 The structure equations

Extremals for the uncertainty principle satisfy differential equations. Since the membrane transfer function is described by an extremal function and its transforms under the symmetry group and since the extremal functions are preserved under this action, it is possible to derive differential equations for the output of the signal. The resulting equations are called the structure equations.

The derivation starts with the simple case $c = 1$. In this situation differentiation of the wavelet transform $Wf(a, t)$ with respect to the parameters of the symmetry group directly leads to a differential equation. In the case $c > 0$ however, the resulting equation is actually a pseudo differential equation. A linearization process for the kernel brings it back to a differential equation that is then satisfied approximately.

The quantities in the equation at first are derivatives of the output function $Wf(a, t)$ and its Hilbert transform. A further calculation then shows that the result can be formulated as an inhomogeneous system of linear partial differential equations for the phase and for the logarithm of the amplitude of the output signal. This is particularly satisfying because these are exactly the physiologically relevant quantities.

The point of departure is the tenet that the basilar membrane transfer function is given as

$$\hat{g}(x, \omega) = h_c \circ \hat{\delta}_\xi(\omega) \quad \left(= h_c \left(\frac{\omega}{\xi} \right) \right),$$

with an extremal function h_c (normalized by the condition $v(h_c) = 1$).

$$\hat{g}(x, \omega) = h_c \left(\frac{\omega}{\xi} \right) = k e^{i\varepsilon \, \mathrm{sgn}(\omega) - i\alpha \, \mathrm{sgn}(\omega) \log|\frac{\omega}{\xi}| - i\beta \frac{\omega}{\xi}} e^{-\frac{\kappa}{c} |\frac{\omega}{\xi}|^c} \left| \frac{\omega}{\xi} \right|^{\kappa - \frac{1}{2}}.$$

As mentioned in the section 'background' the response to a general signal is interpreted as a wavelet transform:

$$u(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega) \hat{g}(x, \omega) e^{i\omega t} \, d\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega) h_c \left(\frac{\omega}{\xi(x)} \right) e^{i\omega t} \, d\omega,$$

with

$$\xi(x) = K e^{-x}.$$

The parameter in the wavelet transform can be normalized such that $a = \frac{1}{\xi(x)}$. The wavelet transform is then

$$Wf(a, t) = \int_{-\infty}^{\infty} \hat{f}(\omega) \sqrt{a} h_c(a\omega) e^{i\omega t} \, d\omega. \tag{50}$$

The considerations in this section start from the formula

$$u(x, t) = \frac{1}{\sqrt{a}} W f(a, t). \tag{51}$$

First the case $c = 1$ is treated. It leads to exact results whereas in the case $c > 1$ an approximation procedure will be applied.

Differentiation with respect to the variable a gives

$$a \frac{\partial}{\partial a} W f(a, t) = \int_{-\infty}^{\infty} \hat{f}(\omega) a \frac{d}{da} \hat{\delta}_a h(\omega) e^{i\omega t} \, d\omega$$

$$= \int_{-\infty}^{\infty} \hat{f}(\omega) \hat{\delta}_a \hat{A} h(\omega) e^{i\omega t} \, d\omega.$$

Note that $\hat{A}$ commutes with $\hat{\delta}_a$. The normalized extremal function h satisfies the differential equation

$$\hat{A} h = (\alpha \hat{H} + \beta \hat{B} + \kappa \hat{H} \hat{B} + \kappa) h.$$

It follows that

$$a \frac{\partial}{\partial a} W f(a, t) = \int_{-\infty}^{\infty} \hat{f}(\omega) \hat{\delta}_a (\alpha \hat{H} + \beta \hat{B} + \kappa \hat{H} \hat{B} + \kappa) h(\omega) e^{i\omega t} \, d\omega$$

$$= \int_{-\infty}^{\infty} \hat{f}(\omega) (\alpha \hat{H} + a\beta \hat{B} + a\kappa \hat{H} \hat{B} + \kappa) \hat{\delta}_a h(\omega) e^{i\omega t} \, d\omega.$$

Under the Fourier transform, $-\frac{d}{dt}$ is mapped into $\hat{B}$ and the Hilbert transform H is mapped into $\hat{H}$. This gives the basic equation

$$a \frac{\partial}{\partial a} W f(a, t) = \kappa W f(a, t) + \alpha H W f(a, t) - a\beta \frac{\partial}{\partial t} W f(a, t)$$

$$- a\kappa \frac{\partial}{\partial t} H W f(a, t). \tag{52}$$

It is quite remarkable that the differentiation $a \frac{\partial}{\partial a}$ (this essentially is differentiation with respect to x) brings in the Hilbert transform H. This transform is a unitary operator on $L^2(\mathbf{R}, \mathbf{C})$. Its square is the negative of the identity operator: $H^2 = -I$. It extends to a bigger class of functions (to all temperate distributions). On the basic trigonometric functions it operates very simply:

$$H \cos(\omega t) = \sin(\omega t).$$

Experimentally, the basilar membrane function is - at least in many studies - determined in terms of the input signals $\cos(\omega t)$. It immediately follows that

$$W H f(a, t) = H W f(a, t) \tag{53}$$

holds for the functions $f = \cos(\omega t)$. The linearity assumption then implies that this holds for arbitrary input signals f.

Since $\hat{H}$ commutes with both $\hat{A}$ and $\hat{B}$, the basic equation tells us that

$$
a\frac{\partial}{\partial a}HWf(a,t) = \kappa HWf(a,t) - \alpha Wf(a,t) - a\beta\frac{\partial}{\partial t}HWf(a,t)
$$
$$
+ a\kappa\frac{\partial}{\partial t}Wf(a,t). \tag{54}
$$

The Hilbert transform thus appears naturally in this setting. This is the justification to study the 'analytic wavelet transform'. Taking the factor $\sqrt{a}$ into account this is

$$
Zf(a,t) = \frac{1}{\sqrt{a}}Wf(a,t) + iH\frac{1}{\sqrt{a}}Wf(a,t) = u(x,t) + iHu(x,t). \tag{55}
$$

The complex valued function Zf then satisfies

$$
a\frac{\partial}{\partial a}Zf(a,t) = \left(\kappa - \frac{1}{2}\right)Zf(a,t) + \alpha HZf(a,t) - a\beta\frac{\partial}{\partial t}Zf(a,t)
$$
$$
- a\kappa\frac{\partial}{\partial t}HZf(a,t). \tag{56}
$$

Notice the shift by $\frac{1}{2}$ that has its origin in the factor $\frac{1}{\sqrt{a}}$.

The basic equation can now be reformulated as a system of equations in the polar coordinates

$$
r(a,t) = |Zf(a,t)| = \left(\frac{1}{a}|Wf(a,t)|^2 + \frac{1}{a}|HWf(a,t)|^2\right)^{\frac{1}{2}},
$$

$$
\varphi(a,t) = \arg Zf(a,t) = \operatorname{arc\,tg}\frac{HWf(a,t)}{Wf(a,t)}.
$$

The calculation in real terms starts with the observation that on one hand

$$
\operatorname{Re}\left(a\frac{\partial}{\partial a}Zf\overline{Zf}\right) = \frac{1}{2}a\frac{\partial}{\partial a}r^2,
$$

whereas on the on the other hand

$$
a\operatorname{Re}\left(a\frac{\partial}{\partial a}Zf\overline{Zf}\right)
$$
$$
= Wf\left(\alpha HWf + \left(\kappa - \frac{1}{2}\right)Wf - a\beta\frac{\partial}{\partial t}Wf - a\kappa\frac{\partial}{\partial t}HWf\right)
$$
$$
+ HWf\left(-\alpha Wf + \left(\kappa - \frac{1}{2}\right)HWf - a\beta\frac{\partial}{\partial t}HWf + a\kappa\frac{\partial}{\partial t}Wf\right)
$$
$$
= a\left(\kappa - \frac{1}{2}\right)r^2 - a\left(Wf\frac{\partial}{\partial t}Wf + HWf\frac{\partial}{\partial t}HWf\right)
$$
$$
- a\kappa\left(Wf\frac{\partial}{\partial t}HWf - HWf\frac{\partial}{\partial t}Wf\right)
$$

$$= a\left(\kappa - \frac{1}{2}\right)r^2 - \frac{a^2\beta}{2}\frac{\partial}{\partial t}r^2 - a^2\kappa r^2\frac{\partial\varphi}{\partial t}.$$

This gives the first equation

$$a\frac{\partial}{\partial a}\log r = \left(\kappa - \frac{1}{2}\right) - a\beta\frac{\partial}{\partial t}\log r - a\kappa\frac{\partial\varphi}{\partial t}. \tag{57}$$

Similarly, on one hand

$$\mathrm{Im}\left(a\frac{\partial}{\partial a}Zf\overline{Zf}\right) = \mathrm{Re}\,Za\frac{\partial}{\partial a}\,\mathrm{Im}\,Z - \mathrm{Im}\,Za\frac{\partial}{\partial a}\,\mathrm{Re}\,Z$$

$$= r^2a\frac{\partial}{\partial a}\,\mathrm{arc\,tg}\,\frac{\mathrm{Im}\,Z}{\mathrm{Re}\,z} = r^2a\frac{\partial\varphi}{\partial a}.$$

On the other hand

$$a\,\mathrm{Im}\left(a\frac{\partial}{\partial a}Zf\overline{Zf}\right)$$

$$= Wf\left(-\alpha Wf + \left(\kappa - \frac{1}{2}\right)HWf - a\beta\frac{\partial}{\partial t}HWf - a\kappa\frac{\partial}{\partial t}Wf\right)$$

$$\quad - HWf\left(-\alpha HWf + \left(\kappa - \frac{1}{2}\right)Wf - a\beta\frac{\partial}{\partial t}Wf - a\kappa\frac{\partial}{\partial t}HWf\right)$$

$$= -\alpha ar^2 - a\beta\left(Wf\frac{\partial}{\partial t}HWf - HWf\frac{\partial}{\partial t}Wf\right)$$

$$\quad + a\kappa\left(Wf\frac{\partial}{\partial t}Wf + HWf\frac{\partial}{\partial t}HWf\right)$$

$$= -\alpha r^2 - a^2\beta r^2\frac{\partial\varphi}{\partial t} + \frac{a^2\kappa}{2}\frac{\partial}{\partial t}r^2.$$

This gives the second equation

$$a\frac{\partial\varphi}{\partial a} = -\alpha + a\kappa\frac{\partial}{\partial t}\log r - a\beta\frac{\partial\varphi}{\partial t}. \tag{58}$$

The calculation in complex notation makes use of the fact that

$$HZf = Hu(x,t) + iH^2u(x,t) = -i\big(u(x,t) + iHu(x,t)\big) = -iZf.$$

The basic equation is then

$$a\frac{\partial}{\partial a}Zf = \left(\kappa - \frac{1}{2}\right)Zf + \alpha HZf - a\beta\frac{\partial}{\partial t}Zf - a\kappa\frac{\partial}{\partial t}HZf \tag{59}$$

$$= \left(\kappa - \frac{1}{2} - i\alpha\right)Zf - a(\beta - i\kappa)\frac{\partial}{\partial t}Zf. \tag{60}$$

Dividing by Zf it follows immediately that

$$\frac{\partial}{\partial a} \log Zf = \kappa - \frac{1}{2} - i\alpha - a(\beta - i\kappa)\frac{\partial}{\partial t} \log Zf. \tag{61}$$

The case $c \neq 1$ is now being treated in a similar spirit. Recall that the normalized extremal functions ($\nu = 1$) satisfy the differential equation

$$\hat{A}h_c = (\alpha\hat{H} + \beta\hat{B} + \kappa\hat{H}\hat{B}_c + \kappa)h_c$$

and that the basic equation derives from

$$a\frac{\partial}{\partial a}Wf(a,t) = \int_{-\infty}^{\infty} \hat{f}(\omega)\hat{\delta}_a \hat{A}h_c(\omega)e^{i\omega t}\,d\omega$$

$$= \int_{-\infty}^{\infty} \hat{f}(\omega)\hat{\delta}_a(\alpha\hat{H} + \beta\hat{B} + \kappa\hat{H}\hat{B}_c + \kappa)h_c(\omega)e^{i\omega t}\,d\omega.$$

The case $c \neq 1$ will not directly lead to a differential operator, because the operator $\hat{B}_c$ is not the Fourier transform of a differential operator (unless c is an odd natural number). It is however possible to use a linear approximation for $\hat{B}_c$ near the frequency expectation value of h_c, that is, at the point $\omega = 1$:

$$-i\,\text{sgn}(\omega)|\omega|^c = -i\,\text{sgn}(\omega) - ic(\omega - \text{sgn}(\omega)) + O(\omega - \text{sgn}(\omega))^2, \tag{62}$$

$$\hat{B}_c \cong \hat{H} + c(\hat{B} - \hat{H}). \tag{63}$$

The above equation is then approximated by

$$a\frac{\partial}{\partial a}Wf(a,t)$$

$$\cong \int_{-\infty}^{\infty} \hat{f}(\omega)\hat{\delta}_a(\alpha\hat{H} + \beta\hat{B} + \kappa\hat{H}(1-c)\hat{H} + c\kappa\hat{H}\hat{B} + \kappa)h_c(\omega)e^{i\omega t}\,d\omega$$

$$= \int_{-\infty}^{\infty} \hat{f}(\omega)(\alpha\hat{H} + a\beta\hat{B} + ac\kappa\hat{H}\hat{B} + c\kappa)\hat{\delta}_a h_c(\omega)e^{i\omega t}\,d\omega.$$

With the consequence that

$$a\frac{\partial}{\partial a}Wf \cong c\kappa Wf + \alpha HWf - a\beta\frac{\partial}{\partial t}Wf - ac\kappa\frac{\partial}{\partial t}HWf. \tag{64}$$

The calculation for the analytic wavelet transform $Zf = u + iHu$ then proceeds as above. Only the constants are slightly different. In the sequel the notation $\gamma = c\kappa$ will be used. Recall that in prospective refined adjustments the parameters α, β and γ may vary with c.

The structure equations are

$$a\frac{\partial}{\partial a}\log r \cong \left(\gamma - \frac{1}{2}\right) - a\beta\frac{\partial}{\partial t}\log r - a\gamma\frac{\partial\varphi}{\partial t}, \tag{65}$$

$$a\frac{\partial \varphi}{\partial a} \cong -\alpha + a\gamma \frac{\partial}{\partial t}\log r - a\beta \frac{\partial \varphi}{\partial t}. \tag{66}$$

They combine to the complex equation

$$a\frac{\partial}{\partial a}\log Zf \cong \gamma - \frac{1}{2} - i\alpha - a(\beta - i\gamma)\frac{\partial}{\partial t}\log Zf. \tag{67}$$

Equality holds if $c = 1$.

Under the tonotopic axis $\xi(x) = Ke^{-x}$, $x = k + \log a$, the derivative $a\frac{\partial}{\partial a}$ transforms into $\frac{\partial}{\partial x}$:

$$\frac{\partial}{\partial x}(u + iHu)(x, t) = a\frac{\partial}{\partial a}\log Zf(a, t). \tag{68}$$

The structure equations can be written in x, t-coordinates:

$$\frac{\partial}{\partial x}\log(u + iHu)(x, t) \cong \gamma - \frac{1}{2} - i\alpha - \frac{1}{\xi(x)}(\beta - i\gamma)\frac{\partial}{\partial t}\log(u + iHu)(x, t). \tag{69}$$

2.4 Consequences of the structure equations

Signal processing in the cochlea is non-linear. The main - but certainly not the only - source of non-linearity is the compressive nature inherent in the hearing process. In the abstract model pursued here this is taken care of with a single parameter that represents the level of sound intensity. The model then describes the linear approximations at these levels. The structure equations are at the core of this abstract model, in fact they comprise all the essential features. First of all, they are linear (as would be expected from a linear approximation). From a mathematical point of view, the equations therefore are very simple. On top, the system is quite special. With respect to suitable variables it represents an inhomogeneous $\bar{\partial}$-equation. Its solutions can be realized in complex form as products of two factors, the first of which is entirely determined by the system and the second is a holomorphic function that can be calculated from the signal. At every level c it is thus possible to associate to an input signal in a unique way a holomorphic function that describes the output signal in terms of the physiological parameters.

The phase and the logarithm of the amplitude are used in the description of the experiments and they are omnipresent in all the representations of the auditory pathway. In themselves they are of limited significance, because they are not coded as such. What really is essential in any cochlear or in any neural model are the changes of these quantities, both with respect to time and with respect to the place. The structure equations precisely relate the local and temporal derivatives of phase and (logarithm of) amplitude. The geometry of the cochlea implicitly is inherent in the extremality property of the basilar membrane filter. But in the structure equations this only shows in terms of the constants. The implicit appearance of the tonotopic axis is an expression of the basic invariance principle that stands at the outset of all considerations.

The structure equations clearly exhibit the dichotomy in cochlear signal processing. The signals can either be analyzed in terms of their phase or in terms of their

amplitudes. Assume that there is complete information on phase changes, that is, the quantities $\frac{\partial \varphi}{\partial t}$ and $\frac{\partial \varphi}{\partial a}$ are known. Then the second equation

$$a \frac{\partial \varphi}{\partial a} \cong -\alpha + a\gamma \frac{\partial}{\partial t} \log r - a\beta \frac{\partial \varphi}{\partial t}$$

can be solved for $\frac{\partial}{\partial t} \log r$. Inserted in the first equation

$$a \frac{\partial}{\partial a} \log r \cong \left(\gamma - \frac{1}{2} \right) - a\beta \frac{\partial}{\partial t} \log r - a\gamma \frac{\partial \varphi}{\partial t}$$

this then determines $\frac{\partial}{\partial a} \log r$. Conversely, the complete knowledge of amplitude information determines the phase information. From an abstract point of view, phase information and amplitude information each individually contain the full information of the signal. In the auditory pathway both phase and amplitude information is being processed. It is commonly assumed that phase information dominates in the low frequency range and amplitude information in the regions that process high frequencies. The equations tell us that phase processing and amplitude processing are equally significant.

The complex equation

$$\frac{\partial}{\partial x} \log(u + iHu)(x, t) \cong \gamma - \frac{1}{2} - i\alpha - \frac{1}{\xi(x)} (\beta - i\gamma) \frac{\partial}{\partial t} \log(u + iHu)(x, t)$$

shows that there is also a twofold way of data processing with respect to time and with respect to the place. Complete information on derivatives with respect to the position gives complete information on time derivatives - and *vice versa*.

The structure equations are so simple that they can be solved in explicit mathematical terms. In its complex form the structure equation is the linear inhomogeneous equation

$$a \frac{\partial}{\partial a} \log Y(a, t) = \gamma - \frac{1}{2} - i\alpha - a(\beta - i\gamma) \frac{\partial}{\partial t} \log Y(a, t). \tag{70}$$

The general solution of an inhomogeneous linear differential equation can be presented as the linear combination of a particular solution (any chosen solution of the equation) and the general solution of the associated homogeneous differential equation.

A particular solution $\log Y_p$ of the above complex equation is the function

$$\log Y_p = \left(\gamma - \frac{1}{2} - i\alpha \right) \log a := P_\gamma(a). \tag{71}$$

Its distinguished feature is the time independence. It follows that the general solution $\log Y$ is of the form

$$\log Y(a, t) = P_\gamma(a) + \log X(a, t) \tag{72}$$

for some function X satisfying the homogeneous equation

$$\frac{\partial}{\partial a} X(a, t) = -(\beta - i\gamma) \frac{\partial}{\partial t} X(a, t). \tag{73}$$

This leads to the product representation

$$Y(a, t) = e^{P_\gamma(a)} X(a, t). \tag{74}$$

(As a side remark, observe that the complex structure equation is obtained from the basic equation

$$a \frac{\partial}{\partial a} Zf = \left(\gamma - \frac{1}{2} - i\alpha\right) Zf - a(\beta - i\gamma) \frac{\partial}{\partial t} Zf \tag{75}$$

after division by Zf. Writing

$$Zf(a, t) = e^{P_\gamma(a)} X(a, t),$$

it is then clear that the homogeneous differential equation for X also holds at the zeros of X.) With the variable change

$$z = t - a\beta + ia\gamma,$$
$$\bar{z} = t - a\beta - ia\gamma$$

the homogeneous equation turns into a $\bar{\partial}$-equation for the transformed function

$$G(z, \bar{z}) = X(a, t). \tag{76}$$

$$
\begin{aligned}
0 &= \frac{\partial}{\partial a} X(a, t) + (\beta - i\gamma) \frac{\partial}{\partial t} X(a, t) \\
&= \frac{\partial G}{\partial z}(-\beta + i\gamma) + \frac{\partial G}{\partial \bar{z}}(-\beta - i\gamma) + (\beta - i\gamma)\left(\frac{\partial G}{\partial z} + \frac{\partial G}{\partial \bar{z}}\right) \\
&= -2i\gamma \frac{\partial G}{\partial \bar{z}}.
\end{aligned}
$$

This then shows that the solutions of the linear inhomogeneous equation have the representation

$$Y(a, t) = a^{\gamma - \frac{1}{2} - i\alpha} G(z), \tag{77}$$

with G a holomorphic function in the variable $z = t - a\beta + ia\gamma$. Since $a > 0$ (and $\gamma > 0$) it is defined in the upper half space $\{z \in \mathbf{C} : \operatorname{Im} z > 0\}$. The function $G(z)$ is uniquely defined up to a constant.

The situation can now be summarized as follows: An incoming signal $f(t)$ gives rise to a family of analytic wavelet transforms

$$Zf = \frac{1}{\sqrt{a}}(Wf + iHWf) \tag{78}$$

depending on the parameter $\gamma = c\kappa$. The functions Zf approximately satisfy the complex structure equation. The solutions

$$Y(a,t) = a^{\gamma - \frac{1}{2} - i\alpha} G(t - a\beta + ia\gamma) \tag{79}$$

of the equation

$$a \frac{\partial}{\partial a} \log Y = \gamma - \frac{1}{2} - i\alpha - a(\beta - i\gamma)\frac{\partial}{\partial t} \log Y \tag{80}$$

are then expected to provide approximations for Zf (with equality for $c = 1$).

The functions G are holomorphic and depend on the parameter γ. They can in fact be determined directly from the Fourier transform of the incoming signal $f(t)$. Since the system is linear, the superposition principle holds:

If $f = f_1 + f_2$ is the superposition of two incoming signals f_1 and f_2 to which the holomorphic functions $G_1(z)$ and $G_2(z)$ are associated, then the holomorphic function for f is

$$G(z) = G_1(z) + G_2(z). \tag{81}$$

All that has to be done is to calculate the holomorphic functions that correspond to to the basic functions $f(t) = A\cos(\nu t + \vartheta)$. In the following section these are identified as the functions

$$G(z) = ke^{i\varepsilon} A e^{i\vartheta} \nu^{\gamma - \frac{1}{2} - i\alpha} e^{i\nu z}. \tag{82}$$

The Fourier representation

$$f(t) = \sqrt{\frac{2}{\pi}} \operatorname{Re} \int_0^\infty \hat{f}(\omega)e^{i\omega t} \, d\omega \tag{83}$$

then tells us that the holomorphic function associated to f is

$$G(z) = ke^{i\varepsilon}\sqrt{\frac{2}{\pi}} \int_0^\infty \hat{f}(\omega)\omega^{\gamma - \frac{1}{2} - i\alpha} e^{i\omega z} \, d\omega. \tag{84}$$

The conclusion is that the holomorphic functions $G(z)$ with $z = t - a\beta + ia\gamma$ provide approximate solutions to the structure equation

$$Zf(a,t) \cong e^{P_\gamma(a)} G(z) \tag{85}$$

$$= a^{\gamma - \frac{1}{2} - i\alpha} G(t - a\beta + ia\gamma). \tag{86}$$

The relevant expressions in the structure equations can then be calculated from the derivative of $F(z) := \log G(z)$:

$$\frac{\partial}{\partial t} \log r = \frac{\partial}{\partial t} \operatorname{Re} F(z) = \operatorname{Re} \frac{\partial}{\partial t} F(z) = \operatorname{Re}\left\{ F'(z)\frac{\partial z}{\partial t} \right\} = \operatorname{Re} F'(z), \tag{87}$$

$$\frac{\partial}{\partial t} \varphi = \operatorname{Im} F'(z), \tag{88}$$

$$a\frac{\partial}{\partial a}\log r = \gamma - \frac{1}{2} + \text{Re}\{F'(z)(-a\beta + ia\gamma)\}, \tag{89}$$

$$a\frac{\partial}{\partial a}\phi = \alpha + \text{Im}\{F'(z)(-a\beta + ia\gamma)\}. \tag{90}$$

2.5 Examples

2.5.1 Pure sounds

For the input signal

$$f(t) = \text{Re } e^{ivt} = \cos vt, \quad v > 0,$$

the quantities $\log r(a, t)$ and $\varphi(a, t)$ can be calculated explicitly from the formula

$$\frac{1}{\sqrt{a}}Wf(a, t) = \text{Re}\{\hat{g}(k + \log a, v)e^{ivt}\}$$

$$= \text{Re}\{ke^{i\varepsilon\,\text{sgn}(v) - i\alpha\log|av| - i\beta av + ivt}e^{-\frac{\kappa}{c}|av|^c}|av|^{\kappa - \frac{1}{2}}\}.$$

This is

$$\log r(a, t) = \log k - \frac{\kappa}{c}|av|^c + \left(\kappa - \frac{1}{2}\right)\log av,$$

$$\varphi(a, t) = i\varepsilon - \alpha\log av - \beta av + vt,$$

and for the derivatives

$$a\frac{\partial}{\partial a}\log r = -\kappa|av|^c + \kappa - \frac{1}{2},$$

$$\frac{\partial}{\partial t}\log r = 0,$$

$$a\frac{\partial}{\partial a}\varphi = -\alpha - \beta av,$$

$$\frac{\partial}{\partial t}\varphi = v.$$

The first structure equation gives

$$-\kappa|av|^c + \kappa - \frac{1}{2} \cong \left(c\kappa - \frac{1}{2}\right) - ac\kappa v.$$

This is in fact the correct linear approximation. It is equivalent with the linear approximation of $|av|^c$ at $av = 1$:

$$|av|^c \cong 1 + c(av - 1).$$

The second structure equation is satisfied as an equality.

From the complex structure equation the holomorphic function associated to $f(t) = \cos \nu t$ can be determined:

$$\log Zf = \log k + i\varepsilon\nu - i\alpha \log |a\nu| - i\beta a\nu + i\nu t$$

$$- \frac{\kappa}{c}|a\nu|^c + \left(\kappa - \frac{1}{2}\right)\log |a\nu|,$$

$$a\frac{\partial}{\partial a}\log Zf = -\kappa |a\nu|^c + \left(\kappa - \frac{1}{2}\right) - i(\alpha + \beta a\nu).$$

With the above approximation this is

$$a\frac{\partial}{\partial a}\log Zf \cong -c\kappa a\nu + \left(c\kappa - \frac{1}{2}\right) - i(\alpha + \beta a\nu) = \gamma a\nu + \left(\gamma - \frac{1}{2}\right) - i(\alpha + \beta a\nu)$$

(with the abbreviation $c\kappa = \gamma$). Observe that for $c > 1$

$$|a\nu|^c > 1 + c(a\nu - 1),$$

unless $a\nu = 1$. The approximate value for $\log |Zf|$ therefore is an over estimation. Together with

$$a\frac{\partial}{\partial t}\log Zf = i\nu$$

the result leads to

$$\log k + i\varepsilon + \left(\gamma - \frac{1}{2} - i\alpha\right)\log |a\nu| - \gamma a\nu - ia\beta\nu + i\nu t$$

$$= \log k + i\varepsilon + P_\gamma(a) + P_\gamma(\nu) + i\nu z$$

as the approximate value of $\log Zf$.

The associated holomorphic function is thus

$$G(z) = ke^{i\nu z + P_\gamma(\nu) + i\varepsilon},$$

with

$$z = t - a\beta + ia\gamma,$$

$$e^{P_\gamma(\nu)} = \nu^{\gamma - \frac{1}{2} - i\alpha}.$$

Note that the holomorphic function G associated to the input signal

$$f(t) = \text{Re } Ae^{i\nu t + i\vartheta} = A\cos(\nu t + \vartheta), \quad \nu > 0, A > 0, \vartheta \in \mathbf{R},$$

is

$$G(z) = ke^{i\varepsilon}Ae^{i\vartheta}(\nu)^{\gamma - \frac{1}{2} - i\alpha}e^{i\nu z}. \tag{91}$$

The constants k and ε are of little importance and do not show in the structure equations. In the following calculations we set $ke^{i\varepsilon} = 1$.

2.5.2 Amplitude modulation

The amplitude modulated signal

$$f(t) = (1 + A \cos \mu t) \cos \nu t$$

$$= \mathrm{Re}\left\{ e^{i\nu t} + \frac{A}{2} e^{i(\nu + \mu t)} + \frac{A}{2} e^{i(\nu - \mu t)} \right\} \quad (0 < A < 1)$$

with $0 < \mu \ll \nu$ is described by the holomorphic function

$$G(z) = e^{P(\nu) + i\nu z} + \frac{A}{2} e^{P(\nu + \mu) + i\nu z + i\mu z} + \frac{A}{2} e^{P(\nu - \mu) + i\nu z - i\mu z}$$

$$= e^{P(\nu) + i\nu z} \left(1 + \frac{A}{2} e^{P(\nu + \mu) - P(\nu) + i\mu z} + \frac{A}{2} e^{P(\nu - \mu) - P(\nu) - i\mu z} \right).$$

The outcome depends on the ratio between the amplitudes of the coefficients. The frequency ν is dominant as long as

$$\left| \frac{A}{2} e^{P(\nu + \mu) - P(\nu) + i\mu z} \right| < \frac{1}{2},$$

$$\left| \frac{A}{2} e^{P(\nu - \mu) - P(\nu) - i\mu z} \right| < \frac{1}{2}.$$

With

$$\mathrm{Re}\{P(\nu \pm \mu) - P(\nu)\} = \left(\gamma - \frac{1}{2} \right) \log \frac{\nu \pm \mu}{\nu} \cong \pm \frac{\mu}{\nu} \left(\gamma - \frac{1}{2} \right),$$

$$\mathrm{Re}\, i\mu z = -\mu a \gamma$$

this gives the estimates

$$a > \frac{1}{\nu} \left(1 - \frac{1}{2\gamma} + \frac{\nu}{\mu \gamma} \log A \right),$$

$$a < \frac{1}{\nu} \left(1 - \frac{1}{2\gamma} - \frac{\nu}{\mu \gamma} \log A \right).$$

The frequency interval covered is

$$\left[\frac{\nu}{1 - \frac{1}{2\gamma} + k}, \frac{\nu}{1 - \frac{1}{2\gamma} - k} \right]$$

with

$$k = -\frac{\nu}{\mu \gamma} \log A.$$

For sufficiently small values of A and $\mu \ll \nu$ it includes the entire range along the cochlea that is involved in the processing of the amplitude modulated signal. The

function F describing this signal in the relevant range can then be estimated by using the approximation $\log(1 + x) \cong x$ for small $|x|$:

$$P(v \pm \mu) - P(v) = \left(\gamma - \frac{1}{2} - i\alpha\right)\log\left(1 \pm \frac{\mu}{v}\right) \cong \pm\left(\gamma - \frac{1}{2} - i\alpha\right)\frac{\mu}{v}$$

and

$$F(z) = P(v) + ivz + \log\left\{1 + \frac{A}{2}\left(e^{i\mu z + \frac{\mu}{v}(\gamma - \frac{1}{2} - i\alpha)} + e^{-i\mu z - \frac{\mu}{v}(\gamma - \frac{1}{2} - i\alpha)}\right)\right\} \quad (92)$$

$$\cong P(v) + ivz + A\cos\left(\mu z - \alpha - i\frac{\mu}{v}\left(\gamma - \frac{1}{2}\right)\right). \quad (93)$$

The result exhibits the basic frequency v as the carrier frequency. But it should be warned that the approximation is valid only in the frequency interval specified above.

The relevant expressions in the structure equation can be calculated from $F'(z)$:

$$F'(z) \cong iv - \mu A\sin\left(\mu z - \alpha - i\frac{\mu}{v}\left(\gamma - \frac{1}{2}\right)\right).$$

It can clearly be seen that there is the constant contribution from the carrier frequency and - as the interesting part - a slow oscillation of angular frequency μ that stems from $\sin(\mu z + const)$. Both the amplitude and the phase derivatives show this oscillation.

2.5.3 The sound of a violin

No doubt, the distinguished feature of a violin sound is the extraordinary big number of harmonics in the frequency spectrum. It is not uncommon to observe around twenty harmonics at an intensity level at which the sounds can still be detected. Except possibly for the first few, the harmonics show a gradual decrease in amplitude with some oscillation. It is conjectured that these properties are in fact characteristic for violins of good quality. Figure 4 shows the amplitudes of the individual harmonics of the violin sounds a, e' and d''. The program 'Prisma-Realtime' by Bachmann *et al.* (2007) [22] uses windowed Fourier transform for this spectrogram. The amplitudes are determined at short intervals and marked with a point. The intensity of these points is fading with the time.

The violin sound has the representation

$$f(t) = \text{Re}\left\{\sum_{m=1}^{\infty} c_m e^{imvt}\right\}.$$

Figure 4 indicates that $|c_m|$ decreases exponentially in m. In the dB-scale the decrease is roughly linear with slope -2:

$$20\log_{10}|c_m| \cong const. - 2m,$$

$$|c_m| \cong const.e^{-0.23m}$$
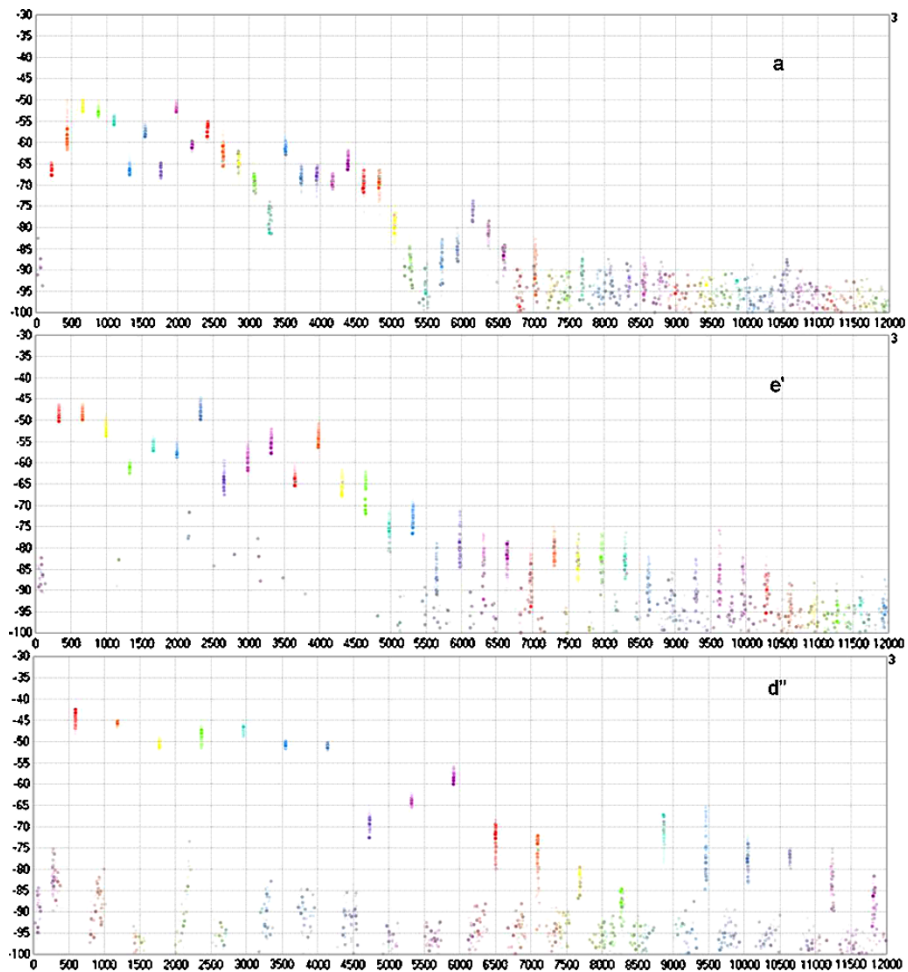
Fig. 4 Spectrograms of three sounds on a new violin. Spectrograms of the three sounds a, e' and d'' on a new violin. The amplitudes of the higher harmonics are determined at short intervals and marked with a point. The intensity of these points is fading with the time. Amplitudes of the harmonics are shown on a relative scale (in dB). More than 20 higher harmonics can be identified. In the first example the level differences of the first 20 harmonics are within a limit of 20 dB.

(with the approximation $10 \cong e^{2.3}$).

The associated holomorphic function is

$$G(z) = \sum_{m=1}^{\infty} c_m e^{imvz + P(mv)}.$$

It is a Fourier series $\sum_{m=1}^{\infty} d_m e^{imvt}$ with coefficients

$$d_m = c_m e^{-a\gamma mv - ia\beta mv + P(mv)}$$

depending on the position (represented by the variable a). The amplitudes

$$|d_m| = |c_m| e^{-a\gamma m\nu} (m\nu)^{\gamma - \frac{1}{2}} \cong e^{0.2m - a\gamma m\nu} (m\nu)^{\gamma - \frac{1}{2}}$$

are maximal for

$$m = \frac{\gamma - \frac{1}{2}}{a\gamma\nu + 0.23} \cong \frac{1}{a\nu}.$$

At the place along the cochlea that codes for the angular frequency $n\nu$, that is, for $\xi(x) = \frac{1}{a} = n\nu$, the coefficient d_n is thus dominant. In a neighborhood of this n^{th} harmonic (near $a = \frac{1}{n\nu}$) the function F can be described locally. The calculation exhibits apart from the contribution by the carrier frequency a substantial oscillatory part of angular frequency ν. In a first calculation only the influence of the two closest harmonics is taken into account. The partial signal

$$f(t) = \text{Re}\{c_n e^{in\nu t} + c_{n+1} e^{i(n+1)\nu t} + c_{n-1} e^{i(n-1)\nu t}\}$$

$$:= \text{Re}\{A e^{in\nu t + i\vartheta} + A_+ e^{i((n+1)\nu t + i\theta_+)} + A_- e^{i((n-1)\nu t + i\theta_-)}\}$$

$$= \text{Re}\left\{A e^{in\nu t + i\vartheta}\left(1 + \frac{A_+}{A} e^{i(\vartheta_+ - \vartheta)} e^{i(n+1)\nu t} + \frac{A_-}{A} e^{i(\vartheta_- - \vartheta)} e^{i(n-1)\nu t}\right)\right\}$$

corresponds to the function

$$F(z) = \log\{A e^{in\nu z + P(n\nu) + i\vartheta} + A_+ e^{(i(n+1)\nu z + P((n+1)\nu) + i\theta_+)}$$

$$+ A_- e^{(i(n-1)\nu z + P((n-1)\nu) + i\theta_-)}\}$$

$$= in\nu z + P(n\nu) + \log A + i\vartheta$$

$$+ \log\left(1 + \frac{A_+}{A} e^{i(\vartheta_+ - \vartheta)} e^{i\nu z + P(\frac{n+1}{n})} + \frac{A_-}{A} e^{i(\vartheta_- - \vartheta)} e^{-i\nu z + P(\frac{n-1}{n})}\right).$$

As before, the approximation

$$P\left(\frac{n+1}{n}\right) \cong \pm\frac{1}{n}\left(\gamma - \frac{1}{2} - i\alpha\right)$$

is used. The coefficients

$$c_\pm = \frac{A_\pm}{A} e^{i(\vartheta_\pm - \vartheta) \pm \frac{1}{n}(\gamma - \frac{1}{2} - i\alpha)}$$

are decomposed as

$$c_+ = \frac{s+d}{2},$$

$$c_- = \frac{s-d}{2},$$

with $s = c_+ + c_-$ and $d = c_+ - c_-$. This gives the approximation

$$F(z) \cong in\nu z + P(n\nu) + \log A + i\theta + s\cos(\nu z) + id\sin(\nu z)$$

near the nth harmonic. The carrier frequency accounts for the part $invz + P(nv) + \log A + i\theta$. In the structure equation it only participates with time independent terms. But the significant part is the contribution that varies in time with angular frequency v (recall that $z = t - a\beta + ia\gamma$).

The qualitative picture remains unchanged, if several neighboring harmonics are considered. Near the nth harmonic the dominant term in the function $F(z)$ is expected to be $invz + P(nv) + \log c_n$.

$$F(z) = \log\left\{ \sum_{m=1}^{\infty} c_m e^{imvz + P(mv)} \right\}$$

$$= invz + P(nv) + \log c_n + \log\left\{ 1 + \sum_{m \neq n} \frac{c_m}{c_n} e^{i(m-n)vz + P(\frac{m}{n})} \right\}.$$

This shows the presence of the carrier frequency. The remainder term is approximated by

$$R_n(z) = \sum_{m \neq n} \frac{c_m}{c_n} e^{i(m-n)vz + P(\frac{m}{n})}$$

$$= e^{ivz} \sum_{m > n} \frac{c_m}{c_n} e^{i(m-n-1)vz + P(\frac{m}{n})} + e^{-ivz} \sum_{m < n} \frac{c_m}{c_n} e^{i(m-n+1)vz + P(\frac{m}{n})}.$$

This function is $\frac{2\pi}{v}$-periodic in time, with leading term

$$e^{ivz} \frac{c_{n+1}}{c_n} e^{P(\frac{n+1}{n})} + e^{-ivz} \frac{c_{n-1}}{c_n} e^{P(\frac{n-1}{n})}.$$

In conclusion it can be said that locally around $a = \frac{1}{nv}$ the function $F(z)$ is of the form

$$F(z) = invz + P(nv) + \log c_n + R_n(z), \tag{94}$$

with a well defined remainder term that is $\frac{2\pi}{v}$-periodic in time. The term $invz + P(nv) + \log c_n$ shows the presence of the carrier frequency. The relevant information about the violin sound is however contained in the fact that the $\frac{2\pi}{v}$-periodicity extends over an interval along the cochlea that comprises more than three octaves in the tonal range. The nature of this contribution is similar all along the interval covered by the frequency spectrum of the violin sound. The exceptions are the low harmonics (essentially the first and second) at which the influence of the neighboring harmonics is very small. Furthermore, the amplitude spectrum of a violin sound very often fails to display monotonicity for the first few harmonics.

With regard to the violins, it should be mentioned that there are considerable differences between different (good quality) instruments. It is believed that the distribution in the first few harmonics very much contributes to the individuality of the violin.

2.6 The impulse response

The response to the impulse function (the dirac function at the origin) is up to rescaling the inverse Fourier transform $\check{h}_c$ of the extremal function

$$h_c(\omega) = k e^{i\varepsilon \operatorname{sgn}(\omega) - i\alpha \operatorname{sgn}(\omega) \log|\omega| - i\beta\omega} e^{-\frac{\kappa}{c}|\omega|^c} |\omega|^{\kappa - \frac{1}{2}}.$$

For simplicity it is assumed that $k e^{i\varepsilon \operatorname{sgn}(\omega)} = 1$. Since the Fourier transform of the dirac function is the constant function $\frac{1}{\sqrt{2\pi}}$, the impulse response is

$$u(x,t) = u(\log K + \log a, t) = \frac{1}{\sqrt{a}} \int_{\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \hat{\delta}_a h_c(\omega) e^{i\omega t} \, d\omega \tag{95}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\infty} h_c(a\omega) e^{i\omega t} \, d\omega \tag{96}$$

$$= \frac{1}{a} \check{h}_c\left(\frac{t}{a}\right). \tag{97}$$

Therefore $au(\log K + \log a, t) = \check{h}_c(\frac{t}{a})$ is a function of the single variable $s = \frac{t}{a}$. This could also be expressed by saying that $tu(x,t)$ only depends on the single variable s - the usual way to formulate the invariance statement.

Figure 5 shows the impulse response for different values of c. The impulse response must have its support on the positive half of the time axis. The membrane cannot show a reaction before the impulse arrives. The numerical calculations show that this is almost satisfied. At this point attention should be drawn to a deficiency of the approach. The basic difficulty lies in the concept of using the uncertainty principle. The appropriate thing would be to restrict the class of functions in the uncertainty principle to functions that in the time domain are supported on the positive half axis. However, in the restricted class there are no extremal functions. This can be seen

Fig. 5 Impulse response. The impulse response on a relative scale as a function of time (in seconds), for $\kappa = 12$ and for various values of c. The position along the cochlea corresponds to the frequency 200 Hz.
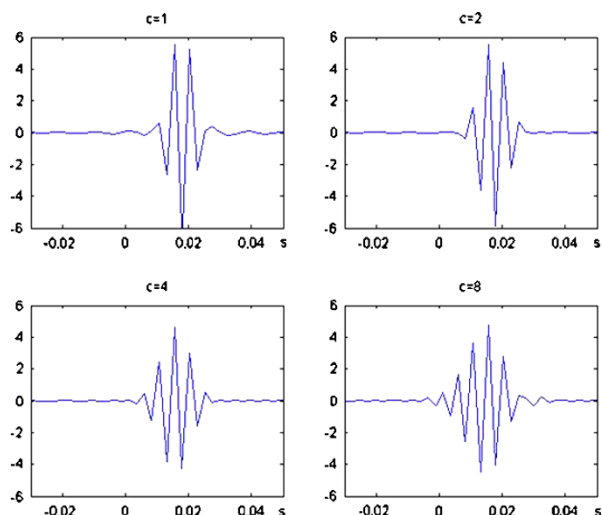
from the fact that the class of extremal functions is translation invariant in the time domain. With the present setting it is not strictly true that the impulse response has its support on the positive half axis. The extremal functions have to be interpreted as a first approximation. They have to be modified slightly such that they really vanish for negative values of t. The numerical calculations show that only small modifications are necessary.

The invariance property of the impulse response mentioned above in combination with the structure equations allow for an explicit approximate calculation of $\check{h}_c$. In the case $c = 1$ this procedure gives the precise value up to a multiplicative constant.

Recall that

$$Zf(a,t) = u(k + \log a, t) + i Hu(k + \log a, t) = r(a,t)e^{i\varphi(a,t)}$$

was the analytic response. The functions f and g are then defined starting from $au(k + \log a, t)$ as

$$g(s) = g\left(\frac{t}{a}\right) = \log a + \log r(a,t),$$

$$f(s) = f\left(\frac{t}{a}\right) = \varphi(a,t).$$

The derivatives appearing in the structure equations can be expressed in terms of f' and g':

$$a\frac{\partial}{\partial a}\log r(a,t) = a\frac{\partial}{\partial a}\left(-\log a + g\left(\frac{t}{a}\right)\right) = -\frac{t}{a}g'\left(\frac{t}{a}\right) - 1,$$

$$\frac{\partial}{\partial t}\log r(a,t) = \frac{\partial}{\partial t}g\left(\frac{t}{a}\right) = \frac{1}{a}g'\left(\frac{t}{a}\right),$$

$$a\frac{\partial}{\partial a}\varphi(a,t) = -\frac{t}{a}f'\left(\frac{t}{a}\right),$$

$$\frac{\partial}{\partial t}\varphi(a,t) = \frac{1}{a}f'\left(\frac{t}{a}\right).$$

They are inserted in the structure equations. The result is a system of differential equations for the functions f' and g' of the variable s.

$$-sg'(s) - 1 \cong \left(\gamma - \frac{1}{2}\right) - \beta g'(s) - \gamma f'(s),$$

$$-sf'(s) \cong -\alpha + \gamma g'(s) - \beta f'(s)$$

or equivalently

$$(s - \beta)g'(s) - \gamma f'(s) \cong -\left(\gamma + \frac{1}{2}\right),$$

$$\gamma g'(s) + (s - \beta)f'(s) \cong \alpha.$$

The complex calculation is quickly done.

$$(s - \beta + i\gamma)\big(g'(s) + if'(s)\big) \cong -\left(\gamma + \frac{1}{2}\right) + i\alpha,$$

$$g(s) + if(s) \cong \left(-\gamma - \frac{1}{2} + i\alpha\right)\log(s - \beta + i\gamma) + const.,$$

$$e^{g(s)+if(s)} \cong K(s - \beta + i\gamma)^{(-\gamma - \frac{1}{2} + i\alpha)}.$$

All that remains is to determine the integration constant K. The inverse Fourier transform $\check{h}_c$ of h_c is then approximated as

$$\check{h}_c(s) \cong \mathrm{Re}\big\{ K(s - \beta + i\gamma)^{(-\gamma - \frac{1}{2} + i\alpha)} \big\}.$$

This is an exact result if $c = 1$.

The function $\log(s - \beta + i\gamma)$ is well defined and holomorphic for $\mathrm{Im}\, s > -\gamma$. In this range

$$\vartheta = \arg(s - \beta + i\gamma)$$

can be chosen in the interval $(0, \pi)$. The modulus of the holomorphic function

$$(s - \beta + i\gamma)^{(-\gamma - \frac{1}{2} + i\alpha)} = e^{(-\gamma - \frac{1}{2} + i\alpha)\log(s - \beta + i\gamma)}$$

is then bounded in the upper half plane. Therefore the function $F(z)$ associated to the impulse response is holomorphic and can be calculated from it. Since z and s are related by $z = t - a\beta + ia\gamma = a(s - \beta + i\gamma)$, it follows that

$$\log r(a,t)e^{i\varphi(a,t)} = \log K - \log a + \left(-\gamma - \frac{1}{2} + i\alpha\right)\log(s - \beta + i\gamma)$$

$$= \log K + \left(\gamma - \frac{1}{2} - i\alpha\right)\log a - \left(\gamma + \frac{1}{2} - i\alpha\right)\log z$$

$$= \log K + P_\gamma(a) - \left(\gamma + \frac{1}{2} - i\alpha\right)\log z.$$

Therefore

$$F(z) = \log K - \left(\gamma + \frac{1}{2} - i\alpha\right)\log z, \qquad (98)$$

and this function is indeed holomorphic in the upper half plane.

The determination of the integration constant is left open. There is however some information that can be obtained directly from the differential equations for f and g. The above system is solved for f' and g':

$$g'(s) \cong \frac{\alpha\gamma - (s - \beta)(\gamma + \frac{1}{2})}{(s - \beta)^2 + \gamma^2},$$

$$f'(s) \cong \frac{\gamma^2 + \frac{\gamma}{2} + \alpha(s - \beta)}{(s - \beta)^2 + \gamma^2} := \frac{N}{D}.$$

From the equation $\frac{\partial}{\partial t} g(\frac{t}{a}) = \frac{1}{a} g'(\frac{t}{a})$ an estimate for the peak of $\log r$ at the position given by the angular frequency $\frac{1}{a}$ can be obtained. The peak is determined by $g'(s) = 0$:

$$0 = \alpha\gamma - (s - \beta)\left(\gamma + \frac{1}{2}\right),$$

$$t = a\left(\beta + \alpha - \frac{\alpha}{2\gamma + 1}\right).$$

This equation tells us how the peak arising from a click travels along the cochlea.

Information on the derivative of the phase is obtained from $\frac{\partial}{\partial t}\varphi(a, t) = \frac{1}{a} f'(\frac{t}{a})$. For this the second derivative of f is calculated:

$$f''(s) \cong \frac{\alpha D - 2(s - \beta)N}{D^2}.$$

Near the peak of $\log r$ the second derivative is small and it changes sign between $\beta + \alpha$ and β (recall that $\alpha < 0$):

$$f''(\beta + \alpha) \cong -\alpha \frac{\gamma^2 + \gamma + \alpha^2}{(\gamma^2 + \alpha^2)^2},$$

$$f''(\beta) \cong \frac{\alpha}{\gamma^2}.$$

Therefore f' is a slightly increasing function at $s = (\alpha + \beta)$ and it is slightly decreasing at $s = \beta$. The function f' itself takes values close to 1:

$$f'(\beta + \alpha) \cong \frac{\gamma^2 + \frac{\gamma}{2} + \alpha^2}{\gamma^2 + \alpha^2} = 1 + \frac{\gamma}{2(\gamma^2 + \alpha^2)},$$

$$f'(\beta) \cong 1 + \frac{1}{2\gamma}.$$

The graph of $f'(s)$ is shown in Figure 6. The parameter values are $\alpha = -5.3\pi$, $\beta = 16\pi$, furthermore, $\gamma = 32\pi$.

According to Shera [23] frequency glides in click responses of the basilar membrane have their origin in the dispersion properties of the slow traveling wave. The variable $s = \frac{t}{a} = t\xi(x)$ is the scale invariant variable denoted by $2\pi\tau = t2\pi f_{CF}(x)$ in [23], p. 2025. Above, the frequency is expressed in the variable s as $\varphi(a, t) = f(s)$. This gives

$$\frac{1}{2\pi}\frac{\partial\varphi}{\partial t}(a, t) = \frac{1}{2\pi a} f'(s)$$

for the instantaneous frequency and hence $f'(s) = \frac{1}{\xi(x)}\frac{\partial\varphi}{\partial t}(a, t)$ for the normalized instantaneous frequency. In [23], p. 2025, Shera denotes it by $\beta_{in}(\tau)$ and pictures its
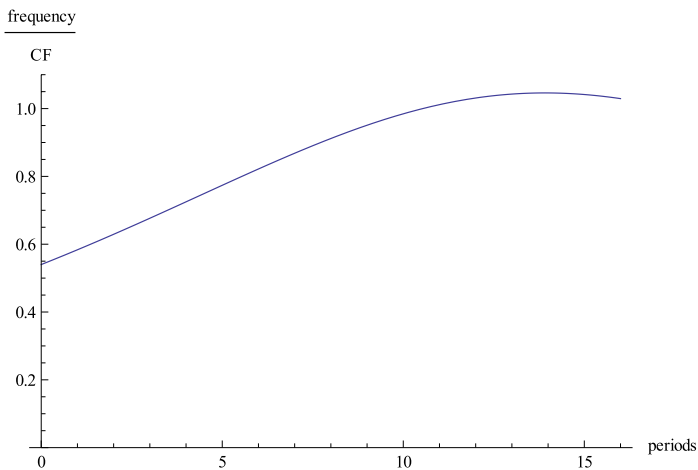
Fig. 6 Frequency glides. The ratio of momentaneous frequency versus CF as a function of time, measured in periods of CF. The parameters are $\alpha = -5.3\pi$, $\beta = 16\pi$ and $c = 32\pi$.

graph in Figure 2b. The above calculations show that $f'(s)$ is increasing for values below $\alpha + \beta$. This can be interpreted as a frequency glide. Note that $f'(s)$ starts to decrease near β. Hence Figure 2b in [23] would roughly confirm the present calculations provided that $\alpha + \beta \approx 12\pi$. There is however a difference in that the present calculation exhibits a dependence of f' on the sound level (as represented by c) with maximal values for f' that are slightly bigger than one.

2.7 General invariance groups

Scale invariance as considered in the previous sections is based on the dilation group δ_a and on the assumption that the tonotopic axis is given by the exponential law $\xi(x) = Ke^{-x}$. From the experimental data it should however rather be concluded that the symmetry hypotheses are satisfied only locally and in a first approximation. The question therefore arises whether the results subsist qualitatively when these basic assumptions are modified. To answer this question the setup of an abstract model is being presented.

The basic hypothesis is still that the symmetry in cochlear mechanics is given by a one parameter transformation group $\hat{\lambda}_a$ acting in phase space. This action can be taken in a quite general form. Ideally it should be possible to adapt it individually to each species. The specific form of the tonotopic axis is to a certain extent independent of the action of the one parameter group. It will be discussed as a separate issue and for the time being the exponential law will be retained.

The transformation group $\hat{\lambda}_a$ thus stands at the outset of a general framework for an abstract description of cochlear mechanics. The one parameter group will be enlarged to a bigger group. At first this will be done on the infinitesimal level by defining a multiplier operator $\hat{M}$ that plays the role of $\hat{B} = -i\omega$ in the previous sections. Together with the infinitesimal generator $\hat{L}$ for the one parameter group $\hat{\lambda}_a$ and together with $\hat{H}$, the action of the Lie algebra of the abstract symmetry group is then

completely determined. As an abstract group, the symmetry group is still $\Gamma \times S$, yet the action in phase space will have changed. It will in fact be conjugate to the standard action. The conjugation mapping typically maps a bounded symmetric range in frequency space onto the whole frequency axis. In the application the bounded range will be the interval $(-R, R)$. The upper bound R appears as an absolute frequency bound. In such a model, the inner ear is completely indifferent to signals whose frequency content lies beyond this limit.

As a first issue the wavelet transform for general one parameter groups is being discussed. Next the action of the one parameter group is determined in dependence of the parameter c that relates to the overall sound level of the signal. This action will then be extended on the infinitesimal level to all of $\Gamma \times S$. Along with this, the conjugation mapping will be defined.

2.7.1 The wavelet transform for general one parameter groups

One parameter transformation groups on $\mathbf{R}^+ = (0, \infty)$ are generated by vector fields $v : \mathbf{R}^+ \to \mathbf{R}$. Assume that v extends to a continuous odd function on $\mathbf{R}$ and that the solutions $\tau_t(\omega)$ to the differential equation

$$\frac{dx}{dt} = v(x) \tag{99}$$

with initial condition $\tau_0(\omega) = \omega$ are uniquely determined and depend smoothly on the initial condition ω. Then τ_t is a one parameter transformation group of $\mathbf{R}^+$. It extends in an antisymmetric way to all of $\mathbf{R}$. At the point $x = 0$ the vector field vanishes and $x = 0$ is a stationary solution of the differential equation. Let us transform the time parameter t and set

$$a = e^t,$$

$$\lambda_a = \lambda_{e^t} = \tau_t.$$

The one parameter group $\{\lambda_a\}$ is then written with a multiplicative parameter $a \in \mathbf{R}^+$ (the multiplicative group of positive real numbers). It then follows that

$$\frac{d\lambda_a}{da}(\omega) = \frac{d\tau_t}{da}(\omega) = \frac{d\tau_t}{dt}(\omega)\frac{1}{a} = v(\tau_t\omega)\frac{1}{a}$$

$$= \frac{1}{a}v(\lambda_a\omega).$$

Example

$$v(x) = x \log \frac{|x|}{R}.$$

The differential equation has the solutions

$$\tau_t(\omega) = \text{sgn}(\omega) R \left| \frac{\omega}{R} \right|^{e^t},$$

$$\lambda_a(\omega) = \text{sgn}(\omega) R \left| \frac{\omega}{R} \right|^a.$$

Assume that the vector field v has a finite number of zeroes $\pm x_i$, labeled in ascending order

$$x_0 = 0 < x_1 < x_2 < \cdots < x_k = \infty$$

(the point $x_k = \infty$ is included though this is not necessarily a zero of v). If

$$\omega \in (x_{i-1}, x_i) := I_i,$$

then $\lambda_a \omega$ stays in I_i for all $a \in \mathbf{R}^*$. If in addition $v > 0$ in I_i, then

$$\lim_{a \to \infty} \lambda_a \omega = x_i$$

and

$$\lim_{a \to 0} \lambda_a \omega = x_{i-1}.$$

The variable transform $\zeta = \lambda_a \omega$,

$$\frac{d\zeta}{da} = \frac{d}{da}(\lambda_a \omega) = \frac{1}{a} v(\lambda_a \omega) = \frac{1}{a} v(\zeta) \tag{100}$$

gives

$$\int_0^\infty g(\lambda_a \omega) \frac{da}{a} = \int_{x_{i-1}}^{x_i} g(\zeta) \frac{d\zeta}{v(\zeta)}$$

whenever the integrals exist. If however $v < 0$ in I_i then the sign changes. Hence in both cases

$$\int_0^\infty g(\lambda_a \omega) \frac{da}{a} = \int_{I_i} g(\zeta) \frac{d\zeta}{|v(\zeta)|}.$$

If $\tau_t(\omega)$ is the solution of the differential equation $\frac{dx}{dt} = v(x)$ with initial condition $\tau_0(\omega) = \omega$ then the variational equation is

$$\frac{d}{dt} \frac{d\tau_t}{d\omega} = v'(\tau_t \omega) \frac{d\tau_t}{d\omega}.$$

Integrating with respect to t gives

$$\ln \left| \frac{d\tau_t}{d\omega} \right| - \ln \left| \frac{d\tau_0}{d\omega} \right| = \int_0^t \frac{d}{dt} \ln \left| \frac{d\tau_t}{d\omega} \right| dt$$

$$= \int_0^t v'(\tau_t \omega) \, dt = \int_\omega^{\tau_t \omega} \frac{v'(x)}{v(x)} \, dx$$

$$= \ln \left| \frac{v(\tau_t \omega)}{v(\omega)} \right|,$$

$$\left|\frac{d\tau_t}{d\omega}\right| = \frac{|v(\tau_t\omega)|}{|v(\omega)|} = \frac{v(\tau_t\omega)}{v(\omega)}.$$

This can be written as

$$\frac{|d\omega|}{|v(\omega)|} = \frac{|d\lambda_a|}{|v(\lambda_a\omega)|}. \tag{101}$$

The formula expresses that $\frac{|d\omega|}{|v(\omega)|}$ is the invariant measure for the group action.

The transformation group $\{\lambda_a\}$ induces a unitary representation λ on $L^2(\mathbf{R}, \mathbf{C})$:

$$\lambda_a h(\omega) = \left|\frac{d\lambda_a^{-1}}{d\omega}(\omega)\right|^{\frac{1}{2}} h\left(\lambda_a^{-1}(\omega)\right),$$

$$\|\lambda_a h\|^2 = \int_{\mathbf{R}} \left|\frac{d\lambda_a^{-1}}{d\omega}(\omega)\right| \left|h\left(\lambda_a^{-1}(\omega)\right)\right|^2 d\omega = \|h\|^2. \tag{102}$$

(The same notation λ_a is used for both the group and its unitary representation. To be consistent with the previous notation the group should actually be denoted by $\hat{\lambda}_a$, since it will be taken as a group that acts in phase space.)

Note that

$$\left|\frac{d\lambda_a}{d\omega}(\omega)\right| = \frac{|v(\lambda_a\omega)|}{|v(\omega)|}$$

is a cocycle for the transformation group $\{\lambda_a\}$.

Given the transformation group $\{\lambda_a\}$ in phase space and a function $\psi \in L^2(\mathbf{R}, \mathbf{C})$, define the one parameter family of wavelets by

$$\psi(a, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \lambda_a \hat{\psi}(\omega) e^{it\omega} d\omega. \tag{103}$$

The wavelet transform with wavelet ψ and transformation group $\{\lambda_a\}$ is

$$W_\psi f(a, b) = \int_{-\infty}^{\infty} f(t)\bar{\psi}(a, t - b) dt \tag{104}$$

$$= \int_{-\infty}^{\infty} \hat{f}(\omega)\overline{\lambda_a \hat{\psi}(\omega)} e^{ib\omega} d\omega. \tag{105}$$

The isometry properties of the wavelet transform have their origin in the isometry property of the Fourier transform. Applied to W_ψ they give

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \left|W_\psi f(a, b)\right|^2 db = \int_{-\infty}^{\infty} \left|\hat{f}(\omega)\right|^2 \left|\lambda_a \hat{\psi}(\omega)\right|^2 d\omega.$$

This equality is integrated with respect to the measure $\frac{da}{a}$ and the order of integration is interchanged:

$$\frac{1}{2\pi} \int_0^{\infty} \int_{-\infty}^{\infty} \left|W_\psi f(a, b)\right|^2 \frac{da\, db}{a}$$

$$= \int_{-\infty}^{\infty} |\hat{f}(\omega)|^2 \left(\int_0^\infty \frac{da}{a} |\lambda_a \hat{\psi}(\omega)|^2 \right) d\omega$$

$$= \int_{-\infty}^{\infty} |\hat{f}(\omega)|^2 \left(\int_0^\infty \frac{da}{a} \left| \frac{v(\lambda_a^{-1}\omega)}{v(\omega)} \right| |\hat{\psi}(\lambda_a^{-1}\omega)|^2 \right) d\omega.$$

With the change of variables $\omega = \lambda_a(\sigma)$ the integration with respect to $\frac{da}{a}$ can be transformed into an integration over the interval $I_i = (x_i, x_{i+1})$ that contains ω:

$$\int_0^\infty \frac{da}{a} |v(\lambda_a^{-1}\omega)| |\hat{\psi}(\lambda_a^{-1}\omega)|^2 = \int_{I_i} |\hat{\psi}(\sigma)|^2 \, d\sigma := C_i.$$

Assume now that ψ is a real valued wavelet. Then

$$\frac{1}{2\pi} \int_0^\infty \int_{-\infty}^\infty |W_\psi f(a,b)|^2 \frac{da\, db}{a} = \sum_i C_i \int_{-I_i \cup I_i} |\hat{f}(\omega)|^2 \frac{d\omega}{|v(\omega)|}. \tag{106}$$

From the formula it can be concluded that the full information on f is contained in the wavelet transform, provided the integrals are finite and the constants C_i different from zero for all indices. A formal reconstruction can be obtained in terms of the wavelets

$$\chi(a,t) := \frac{1}{C_i \sqrt{2\pi}} \int_{-\infty}^\infty \left| \frac{d\lambda_a^{-1}\omega}{d\omega} \right|^{-\frac{1}{2}} \hat{\psi}(\lambda_a^{-1}\omega) e^{it\omega} \, d\omega.$$

(Notice the negative sign in the exponent!)

The corresponding wavelet transform is

$$V_\chi g(a,b) = \int_{-\infty}^\infty g(t) \bar{\chi}(a, t-b) \, dt.$$

A similar calculation as before then gives

$$\frac{1}{2\pi} \int_0^\infty \int_{-\infty}^\infty W_\psi f(a,b) \overline{V_\chi g(a,b)} \frac{da\, db}{a} = \int_{-\infty}^\infty \hat{f}(\omega) \bar{\hat{g}}(\omega) \, d\omega = \int_{-\infty}^\infty f(t) \bar{g}(t) \, dt.$$

The function is thus recovered weakly in the sense of L^2-duality by

$$\frac{1}{2\pi} \int_0^\infty \int_{-\infty}^\infty W_\psi f(a,b) \chi(a, t-b) \frac{da\, db}{a} = f(t). \tag{107}$$

For the description of cochlear mechanics the reconstruction of the signal from its output at the cochlear level (that is, from its wavelet transform) is not an issue. No reconstruction is taking place in the auditory pathway. However from the point of information processing it is of relevance to know whether the wavelet transform contains the full information of the original signal.

In the application the wavelet transform will be described by a wavelet with frequency support in $[-R, R]$. The above reconstruction process would then give the projection of the signal onto the subspace of band limited signals.

2.7.2 Extension of the group action

In the case of the dilation group when λ_a is $\hat{\delta}_a$ and the generating vector field is $v(x) = -x$, the operator $\hat{B} = -i\omega$ together with the infinitesimal generator $\hat{A} = \frac{1}{2} + \omega\frac{d}{d\omega}$ for the action of the dilation group satisfy the commutator relation

$$[\hat{A}, \hat{B}] = \hat{B}$$

of the Lie algebra of the affine group Γ. For general vector fields v, the infinitesimal generator for the group action is

$$Lh(\omega) = \frac{d}{da}\bigg|_{a=1} \lambda_a h(\omega) \tag{108}$$

$$= -\frac{1}{2}v'(\omega)h(\omega) - v(\omega)\frac{d}{d\omega}h(\omega). \tag{109}$$

Together with $\hat{B}$ it will in general fail to span a finite dimensional Lie algebra of operators. However there exists a skew hermitian multiplier operator M such that

$$[L, M] = M. \tag{110}$$

The multiplier of M can be taken in the form

$$M = -i\,\mathrm{sgn}(\omega)s(\omega). \tag{111}$$

The symmetric real valued function $s(\omega)$ is then determined by the differential equations that results by applying the commutator relation

$$-v\frac{d}{d\omega}s = s. \tag{112}$$

In any interval void of zeros of the vector field, the function s is determined up to a multiplicative factor by

$$s(\omega) = e^{-\int \frac{d\omega}{v(\omega)}}.$$

If v is a smooth vector field with zeros at $\pm x_i$ then $s(\omega)$ will map any interval $I_i = (x_{i-1}, x_i)$ onto the positive real half axis $\mathbf{R}^+$. Furthermore, it will conjugate the action of the transformation group λ_a (restricted to I_i) with the action of the dilation group $\hat{\delta}_a$:

$$\lambda_a = s^{-1} \circ \hat{\delta}_a \circ s. \tag{113}$$

This can be verified by calculating the infinitesimal generator of the conjugate group:

$$\frac{d}{da}\bigg|_{a=1} s^{-1} \circ \hat{\delta}_a \circ s(\omega) = -\left(s^{-1}\right)'\left(s(\omega)\right)s(\omega) = -\frac{s(\omega)}{s'(\omega)}$$

$$= v(\omega).$$

The conjugation map s induces an isometry s^* between L^2 and the subspace L_I^2 of L^2-functions restricted to $-I_i \cup I_i$:

$$s^*h(\omega) := \left| \frac{ds}{d\omega} \right|^{\frac{1}{2}} h\big(s(\omega)\big) = \left| \frac{s(\omega)}{v(\omega)} \right|^{\frac{1}{2}} h\big(s(\omega)\big). \tag{114}$$

In the application, the interval on which λ_a acts will be $I = (0, R)$ and the vector field will take negative values. The function $\text{sgn}(\omega)s(\omega)$ then maps $(-R, R)$ onto $\mathbf{R}$ and conjugates λ_a to $\hat{\delta}_a$.

The operators L and M satisfy the commutator relation of the Lie algebra of the affine group Γ. The action of the group λ_a is thus extended on the infinitesimal level to a Lie algebra action of the affine group.

It can be lifted to $\Gamma = \{(a, t) : a > 0, t \in \mathbf{R}\}$ by setting

$$\mu_t h(\omega) = e^{-it\,\text{sgn}(\omega)s(\omega)} h(\omega).$$

The circle group action $\hat{\varepsilon}_\varphi$ with infinitesimal generator $\hat{H}$ remains unchanged and commutes with the Γ-action

$$[L, \hat{H}] = [M, \hat{H}] = 0. \tag{115}$$

The operators corresponding to $\hat{B}_c$ are

$$M_c = -i\,\text{sgn}(\omega)s^c(\omega), \quad a > 0.$$

The only non trivial commutator relation is

$$\left[\frac{1}{c}L, M_c \right] = M_c. \tag{116}$$

For later use note that λ_a commutes with L, whereas

$$\lambda_a M = aM\lambda_a,$$
$$\lambda_a M_c = a^c M_c \lambda_a.$$

2.7.3 The uncertainty inequality

For a given vector field generating the transformation group λ_a, an interval $I_i = (x_{i-1}, x_i)$ is fixed. The signal space $L^2(\mathbf{R}, \mathbf{R})$ is then restricted to the subspace L_I^2 of band limited signals with frequencies in $-I_i \cup I_i$. In this space, the general uncertainty inequality for $\Gamma \times S$ can be stated as

$$v\|h\|^2 \le \frac{2}{c} \big\|(L - \alpha\hat{H} - \beta M)h\big\| \big\|(\hat{H}M_c + v)h\big\|, \tag{117}$$

with

$$v = v_c = -\frac{(h, \hat{H}M_c h)}{\|h\|^2} = \frac{1}{\|h\|^2} \int_{-\infty}^{\infty} |s(\omega)|^c |h(\omega)|^2 \, d\omega \tag{118}$$

and

$$\alpha \|h\|^2 - \beta(h, \hat{H}Mh) + (h, \hat{H}Lh) = 0,$$
$$-\alpha(h, \hat{H}Mh) + \beta \|Mh\|^2 - \mathrm{Re}(Lh, Mh) = 0.$$

The quantities $v(h)$, $\alpha(h)$ and $\beta(h)$ are defined for all functions in L_I^2 with Lh and $M_c h$ in L^2. Under λ_a they transform as

$$v(\lambda_a h) = -\frac{(\lambda_a h, \hat{H}M_c\lambda_a h)}{\|h\|^2} = -a^{-c}\frac{(\lambda_a h, \hat{H}M_c\lambda_a h)}{\|h\|^2} = a^{-c}v(h), \quad (119)$$

$$\alpha(\lambda_a h) = \alpha(h), \quad (120)$$

$$\beta(\lambda_a h) = a\beta(h). \quad (121)$$

For every parameter $c > 0$ define the frequency localization of h by

$$\sigma_c(h) = s^{-1}\left(|v_c(h)|^{\frac{1}{c}}\right). \quad (122)$$

Since s conjugates λ_a to $\hat{\delta}_a$, it follows that

$$\sigma_c(\lambda_a h) = \lambda_a\big(\sigma_c(h)\big). \quad (123)$$

Similarly, the phase expectation can be defined by

$$\varrho(h) = s^{-1}\big(\beta(h)\big).$$

It satisfies

$$\varrho(\lambda_a h) = \lambda_a^{-1}\big(\varrho(h)\big).$$

The equation for the extremal functions is

$$(L - \alpha\hat{H} - \beta M)h = \kappa(\hat{H}M_c + v)h. \quad (124)$$

It has explicit solutions in terms of $s = s(\omega)$:

$$h_c(\omega) = |v(\omega)|^{-\frac{1}{2}} k e^{i\varepsilon\,\mathrm{sgn}(\omega) - i\alpha\,\mathrm{sgn}(\omega)\log|s(\omega)| - i\beta s(\omega)} e^{-\frac{\kappa}{c}|s(\omega)|^c} |s(\omega)|^{\kappa v}. \quad (125)$$

These solutions can either be calculated directly from the differential equation for the extremal solutions or they can be determined from h_c by applying the conjugation mapping s^*.

For later purpose note the explicit formula for $\lambda_a h_c(\omega)$. Since s conjugates $\hat{\delta}_a$ to λ_a, one has $s(\lambda_a^{-1}(\omega)) = as(\omega)$ and therefore

$$\lambda_a h_c(\omega) = \left|\frac{d\lambda_a^{-1}}{d\omega}(\omega)\right|^{\frac{1}{2}} h_c\big(\lambda_a^{-1}(\omega)\big) \quad (126)$$

$$= |v(\omega)|^{-\frac{1}{2}} k e^{i\varepsilon\,\mathrm{sgn}(\omega) - i\alpha\,\mathrm{sgn}(\omega)\log|as(\omega)| - i\beta as(\omega)} e^{-\frac{\kappa}{c}|as(\omega)|^c} |as(\omega)|^{\kappa v}. \quad (127)$$

Apart from the factor $|v(\omega)|^{-\frac{1}{2}}$ this is a function of the single variable $as(\omega)$.

2.7.4 Structure equations in the general setting

As in the previous section, assume that v is a vector field with zeros at $\pm x_i$ that generates a one parameter group λ_a of transformations on $-I \cup I = (-x_i, -x_{i-1}) \cup (x_{i-1}, x_i)$. Signals $f \in L_I^2$ with frequency content in $-I \cup I$ can then be analyzed with the group wavelet transform

$$W_\psi f(a, t) = \int_{-\infty}^{\infty} \hat{f}(\omega) \overline{\lambda_a \hat{\psi}(\omega)} e^{it\omega} \, d\omega.$$

The Fourier transform $\hat{\psi}$ of the wavelet with frequency support in $-I \cup I$ is taken to be a normalized (with $v = 1$) extremal function for the uncertainty inequality for the group $\Gamma \times S$. The Lie algebra action is determined by the operators L, M_c and $\hat{H}$. If h is an extremal function with coefficients α, β and v, then $\bar{h}$ is also an extremal function. Its coefficients are $-\alpha$, $-\beta$ and v. It is thus possible to take $\bar{\hat{\psi}} = h_c$:

$$W f(a, t) = \int_{-\infty}^{\infty} \hat{f}(\omega) \lambda_a h_c(\omega) e^{it\omega} \, d\omega. \tag{128}$$

The structure equations have their origin in the differential equations

$$(L - \alpha \hat{H} - \beta M)h = \kappa(\hat{H} M_c + v)h \tag{129}$$

for the extremal functions. In combination with the differentiated form of the wavelet transform

$$a \frac{d}{da} W f(a, b) = \int_{-\infty}^{\infty} \hat{f}(\omega) L \lambda_a h_c(\omega) e^{ib\omega} \, d\omega$$

this gives the basic formula

$$a \frac{d}{da} W f(a, t) = \int_{-\infty}^{\infty} \hat{f}(\omega) \lambda_a (\alpha \hat{H} + \beta M + \kappa \hat{H} M_c + \kappa) h_c(\omega) e^{it\omega} \, d\omega \tag{130}$$

$$= \int_{-\infty}^{\infty} \hat{f}(\omega) (\alpha \hat{H} + a\beta M + a^c \kappa \hat{H} M_c + \kappa) \lambda_a h_c(\omega) e^{it\omega} \, d\omega. \tag{131}$$

In this situation, a linear approximation for the multiplier operator M_c is inserted. The frequency localization of $\lambda_a h_c$ is at $\sigma = \sigma_c(\lambda_a h_c)$. At this position s^c is approximated by

$$s^c(\omega) \cong s^c(\sigma) + cs^{c-1}(\sigma)s'(\sigma)(\omega - \sigma) \tag{132}$$

$$= s^c(\sigma) - c \frac{s^c(\sigma)}{v(\sigma)}(\omega - \sigma). \tag{133}$$

Note that

$$s^c(\sigma) = s^c(\sigma_c(\lambda_a h_c)) = v_c(\lambda_a h_c) = a^{-c} v_c(h_c) = a^{-c} \tag{134}$$

with the chosen normalization. Therefore the multiplier operators M and M_c - when applied to $\lambda_a h_c$ - are approximated by

$$M_c \cong a^{-c}\left(1 + \frac{c\sigma}{v(\sigma)}\right)\hat{H} - a^{-c}\frac{c}{v(\sigma)}\hat{B}, \tag{135}$$

$$M \cong a^{-1}\left(1 + \frac{\sigma}{v(\sigma)}\right)\hat{H} - a^{-1}\frac{1}{v(\sigma)}\hat{B}. \tag{136}$$

Altogether, when applied to $\lambda_a h_c$, the expression in brackets in the basic formula above is replaced by

$$\left(\alpha + \beta\left(1 + \frac{\sigma}{v(\sigma)}\right)\right)\hat{H} - \frac{\beta}{v(\sigma)}\hat{B} - \frac{c\kappa\sigma}{v(\sigma)} - \frac{c\kappa}{v(\sigma)}\hat{H}\hat{B}.$$

The derivative of the wavelet transform is then expressed in terms of Wf, $\frac{d}{dt}Wf$ and their Hilbert transforms

$$\begin{aligned}
a\frac{d}{da}Wf &\cong -\frac{\gamma\sigma}{v(\sigma)}Wf + \left(\alpha + \beta\left(1 + \frac{\sigma}{v(\sigma)}\right)\right)\hat{H}Wf \\
&\quad + \frac{\beta}{v(\sigma)}\frac{d}{dt}Wf + \frac{\gamma}{v(\sigma)}\frac{d}{dt}W\hat{H}f
\end{aligned} \tag{137}$$

(again, the notation $c\kappa = \gamma$ has been used).

In order to identify the Fourier transform of the wavelet with the cochlear filter, it must be suitably normalized. In case of the dilation group, the normalization $\frac{1}{\sqrt{a}}\hat{\delta}_a h(\omega)$ has been used. The function $\hat{\delta}_a h$ is localized at $\xi(x) = \frac{1}{a}$. The appropriate normalization for $\lambda_a h$ is thus $\sqrt{\sigma(\lambda_a h)}\lambda_a h$. Note that $\frac{d}{da}\sigma(\lambda_a h) = \frac{d}{da}\lambda_a(\sigma(h)) = v(\sigma(\lambda_a h))$. As in the section on the structure equation, the analytic wavelet transform of f is then defined by

$$Zf(a,t) = \sqrt{\sigma}Wf(a,t) + i\sqrt{\sigma}HWf(a,t), \tag{138}$$

with $\sigma = \sigma(\lambda_a h)$. The polar coordinates of Zf are

$$r(a,t) = |Zf(a,t)| = \left(\sigma|Wf(a,t)|^2 + \sigma|HWf(a,t)|^2\right)^{\frac{1}{2}},$$

$$\varphi(a,t) = \arg Zf(a,t) = \operatorname{arc tg}\frac{HWf(a,t)}{Wf(a,t)}.$$

The calculation of the general structure equations now proceeds as before. Taken the shift $\frac{v(\sigma)}{2\sigma}$ into account that is caused by the normalizing factor $\sqrt{\sigma}$, the coefficients in the above equation are abbreviated by

$$A = -\frac{\gamma\sigma}{v(\sigma)} + \frac{v(\sigma)}{2\sigma}, \tag{139}$$

$$B = \alpha + \beta\left(1 + \frac{\sigma}{v(\sigma)}\right), \tag{140}$$

$$C = \frac{\beta}{v(\sigma)}, \tag{141}$$

$$D = \frac{\gamma}{v(\sigma)}. \tag{142}$$

The general structure equations are

$$a\frac{\partial}{\partial a}\log r \cong A + C\frac{\partial}{\partial t}\log r + D\frac{\partial \varphi}{\partial t}, \tag{143}$$

$$a\frac{\partial \varphi}{\partial a} \cong -B - D\frac{\partial}{\partial t}\log r + C\frac{\partial \varphi}{\partial t} \tag{144}$$

and in their complex form

$$a\frac{\partial}{\partial a}\log Zf \cong A - iB + (C - iD)\frac{\partial}{\partial t}\log Zf. \tag{145}$$

In the special case that $v(x) = -x$ the coefficients reduce to $A = \gamma - \frac{1}{2}$, $B = \alpha$, $C = -a\beta$ and $D = -a\gamma$.

The previous structure equations are thus recovered.

The solutions of the complex equation can again be determined. The t-independent particular solution P, giving rise to the multiplicative factor e^P, is obtained by solving the equation

$$a\frac{\partial}{\partial a}P = A - iB. \tag{146}$$

If the variable is changed:

$$\sigma = \sigma(\lambda_a h) = \lambda_a\big(\sigma(h)\big) := \lambda_a(\sigma_1),$$

$$\frac{d\sigma}{da} = \frac{1}{a}v(\sigma),$$

then

$$\tilde{P}(\sigma) := P(a) = \int (A - iB)\frac{da}{a} = \int (A - iB)\frac{d\sigma}{v(\sigma)} \tag{147}$$

gives the t-independent factor as a function of the variable σ.

Similarly, the homogeneous equation

$$a\frac{\partial}{\partial a}X(a,t) = (C - iD)\frac{\partial}{\partial t}X(a,t) \tag{148}$$

can be formulated in the variables σ and t:

$$X(a,t) := \tilde{X}(\sigma,t),$$

$$a\frac{\partial}{\partial a}X(a,t) = a\frac{d\sigma}{da}\frac{\partial}{\partial \sigma}\tilde{X}(\sigma,t) = v(\sigma)\frac{\partial}{\partial \sigma}\tilde{X}(\sigma,t).$$

The result is the $\bar{\partial}$-equation

$$v(\sigma)\frac{\partial}{\partial\sigma}\tilde{X}(\sigma,t) = (C - iD)\frac{\partial}{\partial t}\tilde{X}(\sigma,t) \tag{149}$$

for the complex variable

$$z = t + \int (C - iD)\frac{d\sigma}{v(\sigma)}. \tag{150}$$

Expressed in the variables σ and t, the solutions of the complex structure equation are thus of the form $\tilde{P} + \log G$ with $G(z) = \tilde{X}(\sigma,t)$ a holomorphic function in the variable z.

2.7.5 The abstract models

An abstract model for the cochlea is a model for the cochlea that is based on a one parameter group λ_a. Its cochlear filter $g(x,\omega)$ is described by the translates $\lambda_a h$ of a normalized function h. This function is an extremal for the uncertainty relation that goes with the action of the symmetry group $\Gamma \times S$ that is determined from the one parameter group λ_a. The frequency localization $\sigma(g(x,\cdot))$ will be used in place of the CF. It is independent of the parameter c that stands for the level of sound intensity and it satisfies

$$\sigma(\lambda_a h) = \lambda_a\big(\sigma(h)\big). \tag{151}$$

Assume that the extremal function $\lambda_a h$ with frequency location $\sigma(\lambda_a h)$ represents the cochlear filter at x:

$$\sigma(\lambda_a h) = \xi(x). \tag{152}$$

Then the point x' at which $\lambda_{a'} h$ represents the cochlear filter is given by

$$x' = \xi^{-1}\big(\sigma(\lambda_{a'}h)\big) = \xi^{-1}\big(\lambda_{a'}\sigma(h)\big) = \xi^{-1} \circ \lambda_{a'} \circ \xi(x). \tag{153}$$

The position-frequency map ξ conjugates the group λ_a to the transformation group $\tilde{\lambda}_a = \xi^{-1} \circ \lambda_a \circ \xi$ along the x-axis. Observe that the generating vector fields v and $\tilde{v}$ for the group λ_a and its conjugate $\tilde{\lambda}_a$ are related by

$$\tilde{v}(x) = \left(\frac{d\xi}{dx}(x)\right)^{-1} v\big(\xi(x)\big). \tag{154}$$

At the outset of the present studies the group action was given by the dilation group $\hat{\delta}_a$ and the tonotopic axis was defined by

$$x = \log K - \log\omega, \quad \omega > 0. \tag{155}$$

As a result, the group parameter was related to the position by

$$x = \log K + \log a \tag{156}$$

and the group conjugate to $\hat{\delta}_a$ under the tonotopic axis was the translation group, generated by the vector field $\tilde{v} = 1$.

Basically there are now two different ways in which the tonotopic axis can be built into the abstract model:

1. The starting point is the translation group along the x-axis (generated by the constant vector field). Under the tonotopic axis, this group is conjugated to the group λ_a in phase space. As the prime example, take the experimentally determined tonotopic axis

$$x = \log K - \log(\omega + S) \tag{157}$$

   with 'shift' S (the inverse to the position-frequency map, see the section 'Background').
2. The preassigned group λ_a in phase space is the point of departure. The position-frequency map ξ then conjugates this group to a group $\tilde{\lambda}_a = \xi \circ \lambda_a \circ \xi$ along the x-axis. In general, $\tilde{\lambda}_a$ will not be the translation group.

Examples for the two variants will be given in the next subsection. Here is an illustration, how the choice of the one parameter group λ_a can be motivated. There is a graduation of the physiological and geometrical properties along the cochlear duct. The cochlea is arranged in a spiral and the geometric quantities like, for example, the width of the cochlear duct change gradually. The gradation of the physiological data manifests itself, for example, in the change of the elasticity properties of the basilar membrane and in the increase in length of the hair cells and their cilia. The proposition is to take this into account by replacing the translation group along the x-axis (generated by the constant vector field $\tilde{v} = 1$) by the group $\tilde{\lambda}_a$ generated by an affine vector field $\tilde{v} = 1 + kx$ (k constant). This group is then conjugated under the 'rough' position-frequency map $\omega = \xi(x) = Ke^{-x}$ to the scaling group λ_a. The generating vector field for this group is then

$$v(\omega) = \frac{d\xi}{dx}(\xi^{-1}\omega)\tilde{v}(\xi^{-1}(\omega)) = -\omega\left(1 + k \log \frac{K}{\omega}\right) = k\omega \log \frac{\omega}{R}, \quad \omega > 0, \quad (158)$$

with

$$R = Ke^{\frac{1}{k}}. \tag{159}$$

2.7.6 Two examples

In this section the procedure of setting up a model according to invariance principles is summarized by explicitly calculating two specific models from the data derived in the previous sections. At the core is the one parameter transformation group λ_a acting in phase space. This action is extended to a group action of the affine group Γ, compatible with the natural action of the circle group S in phase space. The general uncertainty principle for $\Gamma \times S$ leads to families E_c of extremal functions that are invariant under the action of λ_a.

In the cochlea, incoming signals are transformed into neuronal impulses. This process is non linear - in particular with respect to changes in the level of sound

pressure of the incoming signal. The parameter c stands for an unspecified average level of sound pressure. The action of the cochlea can then be described in its linear approximation at the sound level captured by the parameter c. Linear actions are completely determined by the cochlear transfer function $g_c(x, \omega)$.

The models thus obtained will be called abstract models, since they are based on general principles. There are very few parameters in such a model, and these will have to be estimated on the basis of experimental results. Some flexibility lies in the choice of the one parameter group action that stands at the outset of all considerations. Yet the experimental results restrict the choice to group actions that are close to the action of the dilation group. Abstract models vary smoothly in dependence on the one parameter group λ_a. It can be said that the qualitative behavior of abstract models - in particular with respect to the structure equations - is affected very little by the choice of the group λ_a.

In the first example, the underlying one parameter group is the transformation group in phase space generated by the vector field

$$v(\omega) = \omega \log \frac{|\omega|}{R} \tag{160}$$

with flow

$$\lambda_a(\omega) = \operatorname{sgn}(\omega) R \left| \frac{\omega}{R} \right|^a \tag{161}$$

that has already been discussed as an example in the section on general invariance groups. This vector field is conjugate *via* the tonotopic axis to an affine vector field along the x-axis. The heuristic motivation for this choice is the gradation of all physical quantities along the cochlea. The affine vector field should take care of this aspect. The infinitesimal generator for the L^2-action is

$$Lh(\omega) = -\frac{1}{2}v'(\omega)h(\omega) - v(\omega)\frac{d}{d\omega}h(\omega)$$

$$= -\frac{1}{2}\left(1 + \log\frac{|\omega|}{R}\right)h(\omega) - \omega\log\frac{|\omega|}{R}\frac{d}{d\omega}h(\omega)$$

and the multiplier operator M that satisfies $[L, M] = M$ is given by

$$M = -i\,\operatorname{sgn}(\omega)s(\omega). \tag{162}$$

The symmetric real valued function $s(\omega)$ is determined up to a multiplicative constant by

$$s(\omega) = \frac{-1}{\log\frac{\omega}{R}}. \tag{163}$$

It maps the interval $(-R, R)$ onto the real axis. (The zeroes of the vector field are at 0 and at $\pm R$.) The value R appears as an absolute upper bound for the frequency range that is relevant in the hearing process. The extremal functions $h_c \in E_c$ for the

uncertainty relation have been calculated above (formula (125)). For the normalized function ($\nu = 1$)

$$h_c(\omega) = \left| v(\omega) \right|^{-\frac{1}{2}} k e^{i\varepsilon \, \text{sgn}(\omega) - i\alpha \, \text{sgn}(\omega) \log |s(\omega)| - i\beta s(\omega)} e^{-\frac{\kappa}{c}|s(\omega)|^c} \left| s(\omega) \right|^\kappa \tag{164}$$

the frequency localization is given by

$$\sigma_c(h_c) = s^{-1}\left(\left| v_c(h_c) \right|^{\frac{1}{c}}\right) = s^{-1}(1) = Re^{-1}, \tag{165}$$

since the inverse function to $\tau = s(\omega)$ is

$$\omega = s^{-1}(\tau) = Re^{-\frac{1}{\tau}}. \tag{166}$$

The frequency localization of $\lambda_a h_c$ is then

$$\sigma(a) := \sigma_c(\lambda_a h_c) = \lambda_a\big(\sigma_c(h_c)\big) = \lambda_a\big(Re^{-1}\big) = Re^{-a}. \tag{167}$$

Under the 'rough' position-frequency map $\xi(x) = Ke^{-x}$ the parameter a is related to the position x by the equation

$$Re^{-a} = \sigma_c(\lambda_a h_c) = \xi(x) = Ke^{-x} \tag{168}$$

and hence

$$a = x + \log \frac{R}{K}. \tag{169}$$

A parametric view of the modulus of the basilar membrane filter is shown in Figures 7 and 8. The first figure gives it as a function of the distance to the stapes and the second as a function of frequency (on a relative scale). The shape of the amplitude curves changes with CF from wide shallow tunings to sharp tunings at high CF. This is consistent with neuronal tuning curves (Kiang *et al.* 1965 [24]) and with basilar membrane data (Robles and Ruggero 2001 [18]). Nonlinear analysis techniques that build on Wiener kernels (Temchin *et al.* 2005 [25], Recio-Spinoso *et al.* 2005 [26]) or on 'zwuis-analysis' (van der Heijden and Joris 2003 [27], 2006 [28]) allow to recover
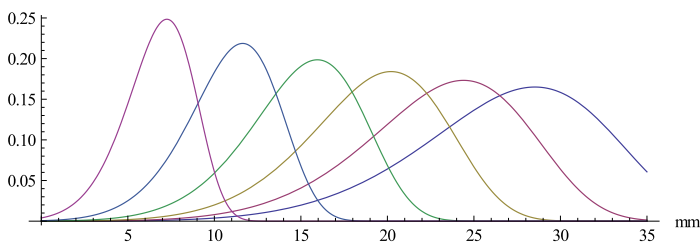
Fig. 7 The cochlear filter for an abstract model. The amplitude (on a relative scale) of the cochlear filter calculated for the invariance group λ generated by the vector field $v(\omega) = \omega \log \frac{\omega}{R}$. This is a panoramic view, showing the amplitude on a relative scale as a function of the distance d to the stapes for the frequencies $f = 200, 400, 800, 1,600, 3,200$ and $6,400$ Hz.
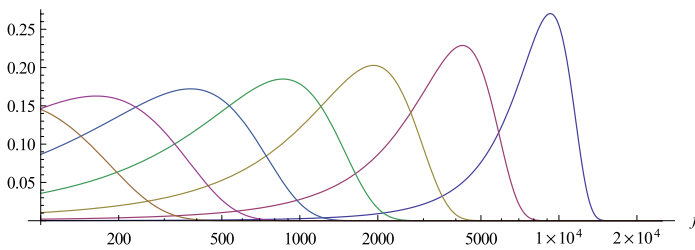
Fig. 8 The cochlear filter as a function of frequency. The amplitude (on a relative scale) of the same cochlear filter as in the previous figure, but as a function of frequency on a logarithmic scale. The places are taken at distances $d = 5, 10, 15, 20, 25, 30$ and 35 mm from the stapes.

the basilar membrane motion from measurements in the auditory nerve. The resulting panoramic graphs (for example, [28], Figure 5 or [25], Figures 1 and 2) should be compared with Figures 7 and 8. That the basilar movement in the apical region of the cochlea is difficult to come by with experimental techniques is discussed by Temchin and Ruggero 2010 [29].

The wavelet transform in the cochlea is given by

$$Wf(a, b) = \int_{-\infty}^{\infty} \hat{f}(\omega) \lambda_a h_c(\omega) e^{ib\omega} \, d\omega. \tag{170}$$

The function $\lambda_a h_c$ is given by formula (127). The expressions for v and s have to be inserted.

The structure equations are approximate differential equations, which are satisfied by

$$r(a, t)e^{\varphi(a,t)} = Zf(a, t) = \sqrt{\sigma(a)}Wf(a, t) + iH\sqrt{\sigma(a)}Wf(a, t). \tag{171}$$

The equations are

$$a\frac{\partial}{\partial a}\log r \cong A + C\frac{\partial}{\partial t}\log r + D\frac{\partial\varphi}{\partial t},$$

$$a\frac{\partial\varphi}{\partial a} \cong -B - D\frac{\partial}{\partial t}\log r + C\frac{\partial\varphi}{\partial t}.$$

The functions A, B, C and D of the parameter a are:

$$A = \frac{\gamma}{a} - \frac{a}{2},$$

$$B = \alpha + \beta\left(1 - \frac{1}{a}\right),$$

$$C = -\frac{\beta}{aR}e^a,$$

$$D = -\frac{\gamma}{aR}e^a.$$

The function $\log Zf(a,t) = \log r(a,t) + i\varphi(a,t)$ is of the form

$$\log Zf(a,t) \cong P(a) + \log G(z) = -\frac{\gamma}{a} - \frac{a}{2} + \frac{i\beta}{a} + i(\alpha+\beta)\log a + \log G(z), \quad (172)$$

with $G(z)$ a holomorphic function of the variable

$$z = t + \int (C - iD)\frac{da}{a} = \frac{-\beta + i\gamma}{R} \int e^a \frac{da}{a}. \quad (173)$$

The second example is based on the 'physiological' tonotopic axis

$$x = \log K - \log(\omega + S) \quad (174)$$

(see section 'Background') with inverse

$$\xi(x) = Ke^{-x} - S. \quad (175)$$

The one parameter group λ_a is taken to be conjugate to $\hat{\delta}_a$ under the shift mapping

$$s(\omega) = \omega + S \quad (176)$$

with shift S. Unfortunately, this mapping does not map $\mathbf{R}^+$ to itself. Hence it has to be modified near $\omega = 0$ such that the modified mapping $\tilde{s}$ is monotone, maps $\mathbf{R}^+$ onto itself and agrees with s for values $\omega \geq \varepsilon$. The positive constant ε can be chosen to be arbitrarily small. This difficulty will be suppressed and the notation s instead of $\tilde{s}$ will be used for the modified mapping. The vector field for the conjugate one parameter group $\lambda_a = s^{-1} \circ \hat{\delta}_a \circ s$ is

$$v(\omega) = -\omega - S. \quad (177)$$

The frequency localization of the extremal function $\lambda_a h$ is

$$\sigma(\lambda_a h) = \lambda_a(\sigma(h)) = s^{-1}\left(\frac{1}{a}\right) = \frac{1}{a} - S. \quad (178)$$

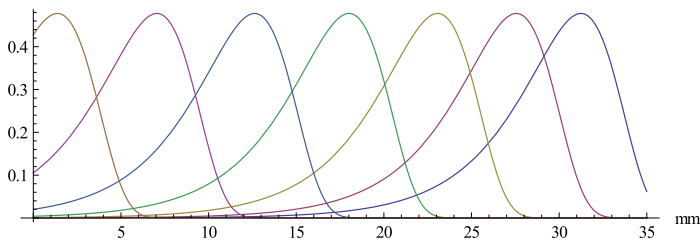The parameter a is related to the position by

$$\frac{1}{a} = \sigma(\lambda_a h) + S = \xi(x) + S. \quad (179)$$

The normalized function

$$\frac{1}{\sqrt{a}}\lambda_a h_c = \sqrt{\frac{\xi(x) + S}{\omega + S}} k e^{i\varepsilon \operatorname{sgn}(\omega) - i\alpha \operatorname{sgn}(\omega)\log|as(\omega)| - i\beta as(\omega)} \\ \times e^{-\frac{\kappa}{c}|as(\omega)|^c} |as(\omega)|^{\kappa\nu} \quad (180)$$

is thus a function of the 'scaling variable'

$$as(\omega) = \frac{\omega + S}{\xi(x) + S}. \quad (181)$$

Fig. 9 Adjusting for the position-frequency map. The amplitude (on a relative scale) of the cochlear filter shown for the physiological position-frequency map $f(d) = Ke^{-\frac{d}{l}} - S$ with $K = 20{,}000$, $S = 200$, $l = 6.6$ and parameter $c = 4$. The amplitude is drawn as function of the distance to the stapes for the frequencies $f = 250, 500, 1{,}000, 2{,}000, 4{,}000, 8{,}000$ and $16{,}000$ Hz.

This is in accordance with the generalized scaling variable defined by Shera (2007, [12], formula 5, p. 2740).

Note that the normalization slightly differs from the normalization $\sqrt{\xi(x)}\lambda_a h_c$ that was used above for the general setting.

In Figure 9 the effect of the physiological tonotopic axis is illustrated with a panoramic view of the modulus of the basilar membrane transfer function.

3 Conclusions

Signal processing in the cochlea is investigated from an abstract point of view. The time variable t and the place variable x along the cochlea provide the basic operators $\frac{d}{dt}$ and $\frac{d}{dx}$. These are translated into frequency space. For the place variable the tonotopic axis is used and for the time variable it is the Fourier transform that accomplishes the transfer. The corresponding operators $\hat{A}$ and $\hat{B}$ in frequency space do not commute. They satisfy the commutator relation $[A, B] = B$ that displays the relation for the Lie algebra of the affine group Γ. The image under the Fourier transform of the space of acoustic signals is identified as the subspace L_{sym} of the complex Hilbert space $L^2(\mathbf{R}, \mathbf{C})$. The only one-parameter group of transformations S on L_{sym} that commutes with the action of the affine group is generated by the (Fourier transform $\hat{H}$ of the) Hilbert transform H. The invariance group for the hearing process in the cochlea is thus the product $\Gamma \times S$ of the affine group Γ with the circle group S.

Signal processing in the cochlea is highly compressive and thus non-linear. The approach pursued here is to fix a parameter c for the general level of sound intensity and to specify a linear approximation at this level.

From a mathematical point of view this is done by choosing a different representation of $\Gamma \times S$ for every parameter c. The non-commutativity of the associated infinitesimal operators gives rise to a family of uncertainty inequalities depending on the parameter c. The extremal functions - the coherent states - for these uncertainty relations are expected to play a special role. It is known from the work of Daubechies (1992) and Yang, Wang and Shamma (1992) that the cochlea performs a wavelet transform. The wavelet is defined by the cochlear filter. In the present approach, the

linear approximation and with it the cochlear filter depend on the parameter c. Comparison with experimental results now suggests that the wavelet that determines the cochlear filter is an extremal function for the uncertainty relation with parameter c.

The abstract model as it is derived here from general mathematical concepts has very few parameters and thus gives a very concise picture of signal processing in the cochlea. On the other hand, with just a few parameters at diposition, it will be impossible to capture all the fine structure that has been established experimentally.

With the model at hand, signal processing in the cochlea can be understood in a global way. At the core of the analysis is a system of differential equations - called the structure equations - that hold for the processed signals, that is, for the wavelet transforms of the acoustic signals. The equations are formulated for physiological observables. In this context these are the derivatives of amplitude and phase. The equations provide us with qualitative and quantitative information on the structure of signal processing in the cochlea. Specifically they give insight into the delicate balance of phase versus amplitude. A global picture emerges since it is possible to present the solutions to the structure equations in terms of holomophic functions.

As examples, pure sounds, amplitude modulated sounds, clicks and sounds of the violin are subject to special scrutiny. The click response exhibits the wavelet in the time variable (Figure 5). A deficiency of the mathematical model becomes apparent, since the impulse response should have its support in the positive time axis. This would imply that the Fourier transform of the wavelet has a holomophic continuation to the half space - and this is not the case. In order to remedy this deficiency, the extremals for the uncertainty inequality should be determined within the class of these functions. Yet this class is not compact and such extremals do not exist. However looking at the impulse response, it is clear that the values for $t < 0$ are quite close to zero. The impulse responses at various intensities are thus close to functions with support in the positive time axis.

The analysis of violin sounds tells us that the pitch frequency is present in the movement of the cochlea on a range that covers several octaves. The symmetries observed experimentally are the most important constituents of the abstract model.

On a closer look, these symmetries are only local, that is they hold approximately on bounded time and frequency intervals. The question then arises whether the theory can be adapted to a wider concept of symmetry. In the last section such a theory is developed and it is argued that the principal features of the model can be preserved in this enlarged framework. In particular, the structure equations still remain approximately satisfied. It thus appears that there is some inherent stability in this model.

Competing interests

The author declares that he has no competing interests.

References

1. Daubechies, I.: Ten Lectures on Wavelets. SIAM, Philadelphia (1992)

2. Yang, X., Wang, K., Shamma, S.: Auditory representation of acoustic signals. IEEE Trans. Inf. Theory **38**, 824–839 (1992)

3. Zweig, G.: Basilar membrane motion. Cold Spring Harbor Symp. Quant. Biol. **40**, 619–633 (1976)

4. Siebert, W.M.: Stimulus transformations in the peripheral auditory system. In: Kolers, P.A., Eden, M. (eds.) Recognizing Patterns, pp. 104–133, MIT, Cambridge (1968)

5. von Békésy, G.: The variation of phase along the basilar membrane with sinusoidal vibrations. J. Acoust. Soc. Am. **19**, 452–460 (1947)

6. Rhode, W.S.: Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique. J. Acoust. Soc. Am. **49**, 1218–1231 (1971)

7. Kiang, N.Y.S., Moxon, E.C.: Tails of tuning curves of auditory-nerve fibers. J. Acoust. Soc. Am. **55**, 620–630 (1974)

8. Liberman, M.C.: Auditory-nerve response from cats raised in a low-noise chamber. J. Acoust. Soc. Am. **63**, 442–455 (1978)

9. Liberman, M.C.: The cochlear frequency map for the cat: labeling auditory nerve-fibers of known characteristic frequency. J. Acoust. Soc. Am. **72**, 1441–1449 (1982)

10. Eldredge, D.H., Miller, J.D., Bohne, B.A.: A frequency-position map for the chinchilla cochlea. J. Acoust. Soc. Am. **69**, 1091–1095 (1981)

11. Greenwood, D.D.: A cochlear frequency-position function for several species - 29 years later. J. Acoust. Soc. Am. **87**, 2592–2605 (1990)

12. Shera, C.A.: Laser amplification with a twist: traveling-wave propagation and gain functions from throughout the cochlea. J. Acoust. Soc. Am. **122**, 2738–2758 (2007)

13. Gabor, D.: Theory of communication. J. IEEE **93**, 429–457 (1946)

14. Cohen, L.: The scale representation. IEEE Trans. Signal Process. **41**, 3275–3292 (1993)

15. Irino, T.: An optimal auditory filter. In: Proc. IEEE Signal Processing Society. Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, 1995

16. Irino, T., Patterson, R.D.: A time-domain, level-dependent auditory filter: the gammachirp. J. Acoust. Soc. Am. **101**, 412–419 (1997)

17. Reimann, H.M.: Uncertainty principles for the affine group. Funct. Approx. Comment. Math. **40**, 45–67 (2009)

18. Robles, L., Ruggero, M.A.: Mechanics of the mammalian cochlea. Physiol. Rev. **81**, 1305–1352 (2001)

19. Rhode, W.S., Recio, A.: Study of mechanical motions in the basal region of the chinchilla cochlea. J. Acoust. Soc. Am. **107**, 317–3332 (2000)

20. Recio, A., Rhode, W.S.: Basilar membrane response to broadband stimuli. J. Acoust. Soc. Am. **108**, 2281–2298 (2000)

21. Shera, C.A.: Intensity-invariance of fine time structure in basilar-membrane click responses: implications for cochlear mechanics. J. Acoust. Soc. Am. **110**, 332–348 (2001)

22. Prisma-Realtime. http://www.prisma-music.ch

23. Shera, C.A.: Frequency glides in click responses of the basilar membrane and auditory nerve: their scaling behavior and origin in travelling-wave dispersion. J. Acoust. Soc. Am. **109**, 2023–2034 (2001)

24. Kiang, N.Y.S., Watanabe, T., Thomas, E.C., Clark, L.F.: Discharge Patterns of Single Fibers in the Cat's Auditory Nerve. MIT, Cambridge (1965)

25. Temchin, A., Recio-Spinoso, A., van Dijk, P., Ruggero, M.: Wiener kernels of Chinchilla auditory-nerv fibers: verification using responses to tones, clicks, and noise and comparision with basilar-membrane vibrations. J. Neurophysiol. **93**, 3635–3648 (2005)

26. Recio-Spinoso, A., Temchin, A., van Dijk, P., Fan, Y.H., Ruggero, M.: Wiener-kernel analysis of responses to noise of chinchilla auditory-nerve fibers. J. Neurophysiol. **93**, 3615–3634 (2005)

27. van der Heijden, M., Joris, P.X.: Cochlear phase and amplitude retrieved from auditory nerve at arbitrary frequencies. J. Neurosci. **23**(27), 9194–9198 (2003)

28. van der Heijden, M., Joris, P.X.: Panoramic measurements of the apex of the cochlea. J. Neurosci. **26**(44), 11462–11473 (2006)

29. Temchin, A., Ruggero, M.: Phase-locked responses to tones of chinchilla auditory nerve fibers: implications for apical cochlear mechanics. JARO **11**, 297–318 (2010)