

A Comparison of the Spectro-Temporal Sensitivity of Auditory Neurons to Tonal and Natural Stimuli

A. M. H. J. Aertsen and P. I. M. Johannesma

Department of Medical Physics and Biophysics, University of Nijmegen, Nijmegen, The Netherlands

Abstract. The spectro-temporal sensitivity of auditory neurons has been investigated experimentally by averaging the spectrograms of stimuli preceding the occurrence of action potentials or neural events (the APES: Aertsen et al., 1980, 1981). The properties of the stimulus ensemble are contained in this measure of neural selectivity. The spectro-temporal receptive field (STRF) has been proposed as a theoretical concept which should give a stimulus-invariant representation of the second order characteristics of the neuron's system function (Aertsen and Johannesma, 1981). The present paper investigates the relation between the experimental and the theoretical description of the neuron's spectro-temporal sensitivity for sound. The aim is to derive a formally based stimulus-normalization procedure for the results of the experimental averaging procedure.

Under particular assumptions, regarding both the neuron and the stimulus ensemble, an integral equation connecting the APES and the STRF is derived. This integral expression enables to calculate the APES from the STRF by taking into account the stimulus spectral composition and the characteristics of the spectrogram analysis. The inverse relation, i.e. starting from the experimental results and by application of a formal normalization procedure arriving at the theoretical STRF, is effectively hindered by the nature of the spectrogram analysis. An approximative "normalization" procedure, based on intuitive manipulation of the integral equation, has been applied to a number of single unit recordings from the grassfrog's auditory midbrain area to tonal and natural stimulus ensembles. The results indicate that spectrogram analysis, while being a useful real-time tool in investigating the spectro-temporal transfer properties of auditory neurons, shows fundamental shortcomings for a theoretical treatment of the questions of interest.

1. Introduction

The combined spectro-temporal characterization of the stimulus-response relation of auditory neurons can be given by the analysis of the spectro-temporal structure of stimuli preceding the occurrence of action potentials or neural events (Johannesma, 1972). This analysis can be performed by determination of the *Average Pre-Event Spectrogram* (APES; Aertsen et al., 1980, 1981). It has been applied to the responses of single neurons in the auditory midbrain of the anaesthetized grassfrog (*Rana temporaria* L.) to tonal (Aertsen et al., 1980) and natural (Aertsen et al., 1981) stimuli. In that case the spectro-temporal analysis had been implemented in hardware in the form of a real-time Dynamic Spectrum Analyser (DSA; Aertsen and Johannesma, 1980). A closely related approach has been used to study the responses of neurons in the lightly anaesthetized grassfrog to stationary wideband noise (Hermes et al., 1981), with the spectro-temporal representation of sound given in a more general and formalized manner (Johannesma and Aertsen, 1979; Johannesma et al., 1981). The formerly mentioned neurons which were investigated with tonal and natural stimuli did not respond to stationary noise, probably due to the higher level of anaesthesia used in those experiments.

The form of the APES is obviously highly influenced by the spectro-temporal characteristics of the stimulus ensemble. In order to arrive at a proper representation of the neuron's spectro-temporal sensitivity, irrespective of the way it has been measured, the result of averaging should be normalized for the a priori known stimulus properties. A somewhat intuitive approach has been described as the "stimulus-filtering" procedure: the APES was divided by the power spectral density of the overall stimulus ensemble (Aertsen et al., 1981). The goal of a stimulus-normalization procedure is to investigate whether the

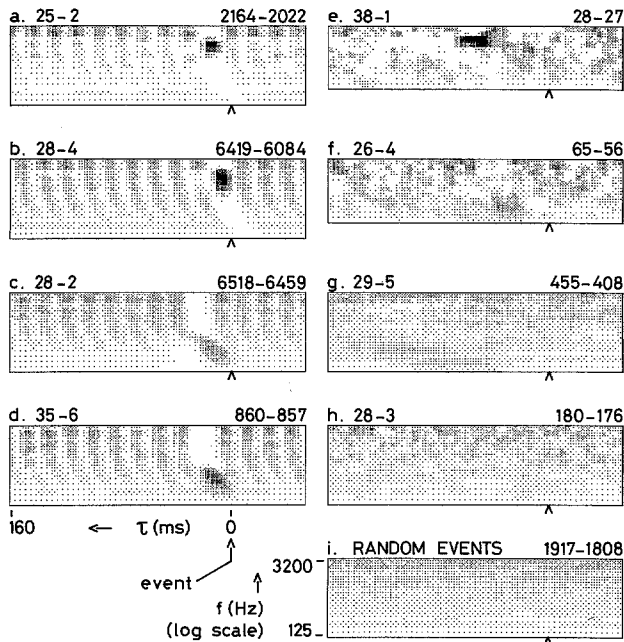


Fig. 1a-i. Average Pre-Event Spectrogram (APES) of eight torus semicircularis units **a-h**, determined with a tonal stimulus ensemble. The a priori expected result for this stimulus ensemble, the Average Spectrogram (AS), is shown under **i**. The intensity scale in all Figs. refers to arbitrary, identical units. High values are indicated by darker, low values by lighter grey-values. The numbers N_1 and N_2 in the Figs. indicate the number of events presented to the DSA (N_1), resp. the number of events actually considered in averaging (N_2). Further explanation in the text

normalized APES for such various stimulus ensembles like tones and natural sounds can be reconciled into one model of the neuron's spectro-temporal sensitivity.

In the present paper the relation between the APES for tonal and natural stimuli will be investigated, using a more formal approach, based on the concept of the *Spectro-Temporal Receptive Field* (STRF; Aertsen and Johannesma, 1981). The STRF has been formally identified with the second order Volterra kernel of the neuron's system function. It has been shown that under certain conditions on both the neuron and the stimulus ensemble a stimulus-normalization procedure can be given which, when applied to the second order cross correlation function of the neuron's in- and output leads to the stimulus-invariant STRF. Although in the present case these conditions are far from being fulfilled, the use of such an approach may nevertheless provide information which can be relevant in tackling the long-standing issue of reconciling the neural responses to "simple" and "complex" stimuli (e.g. Worden and Galambos, 1972; Bullock, 1977). In this context it will be necessary to consider the formal relation between the dynamic spectrum analysis (DSA), leading to the APES and the evaluation of second order input-output cross correlation.

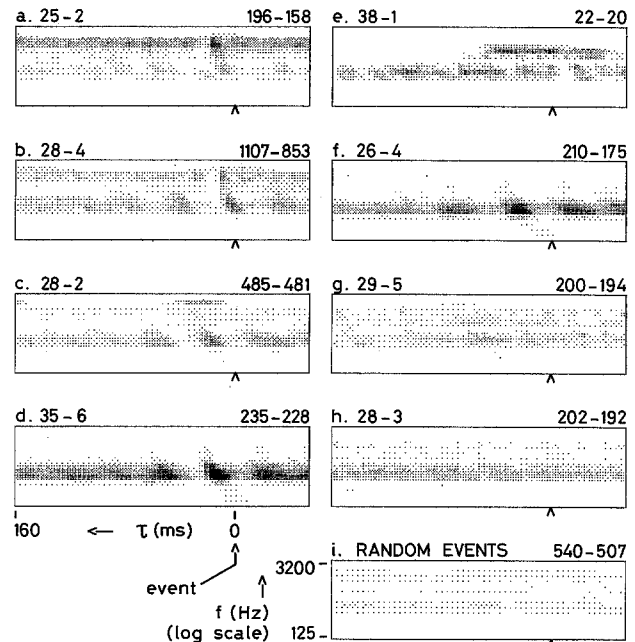


Fig. 2a-i. Average Pre-Event Spectrogram (APES) of eight torus semicircularis units **a-h**, determined with a natural stimulus ensemble. The a priori expected result for this stimulus ensemble, the Average Spectrogram (AS), is shown under **i**. These results refer to the same neurons as have been shown in Fig. 1 for a tonal stimulus ensemble. Further details as in Fig. 1

2. Methods

The experimental data used in the present investigation were obtained from extra-cellular single unit recordings in the torus semicircularis (the auditory midbrain area) of the grassfrog, *Rana temporaria* L., described in earlier papers (Aertsen et al., 1980, 1981).

The stimulus-event (s-e) relation was investigated for two different stimulus ensembles (Aertsen and Johannesma, 1980). The first one consisted of sequences of short (16 ms) tonal elements, the γ -tones, the form of which was inspired by the short sound segments in the male grassfrog vocalizations. The carrier frequency and the amplitude factor of the γ -tones were varied in a pseudo-random way throughout the sequence. The second stimulus ensemble consisted of a wide selection of sounds, occurring in the natural environment of the grassfrog: the Acoustic Biotope.

The properties of the s-e relation have been investigated by analysing the ensemble of stimuli preceding the occurrence of action potentials, the Pre-Event Stimulus Ensemble (PESE; Johannesma, 1972). The spectro-temporal characteristics of the PESE were evaluated by determining the ensemble average of the dynamic power spectra of individual pre-event stimuli, measured by a real-time Dynamic Spectrum Analyser (DSA; Aertsen and Johannesma, 1980). This average measure obviously is heavily influenced by the overall spectro-temporal properties of the stimulus ensemble. The average pre-event dynamic power spectra of eight different torus semicircularis units are shown in Fig. 1 for the tonal stimulus ensemble and in Fig. 2 for the natural stimulus ensemble. A few of these results have been published earlier (Aertsen et al., 1980, 1981). Results are given in a grey-scale representation where higher function values are indicated by darker grey-values. The results for the neural event sequences are shown in the figure elements a-h,

whereas element i shows the outcome of averaging for a random sequence of events. The latter result provides an indication of the overall spectro-temporal structure of the stimulus ensemble. In each figure the results were scaled using the same intensity scale to allow a comparison. The numbers of events, involved in the analysis, are indicated in the figures: the first number (N_1) indicates the number of events recorded during the presentation of the stimulus ensemble; the second number (N_2) denotes the number of pre-event stimuli actually considered in averaging. Due to a "dead-time" in the averaging processor of the DSA of 29.5 ms (Aertsen and Johannesma, 1980) it generally appears that $N_2 < N_1$.

The frequency range of the dynamic spectrum analysis ran from 125 Hz to 3200 Hz on a log-scale, i.e. $15 \frac{1}{3}$ -octave filters. The tonal stimulus sequences covered either one of two frequency ranges: 125–2000 Hz or 250–4000 Hz. For 6 out of 8 units (elements a–d, g, and h in Fig. 1) the average dynamic spectrum for tonal stimuli was obtained by combining the two separate averages corresponding to the respective tonal sequences in the following way: in the overlapping frequency band a weighed average was determined, while in the non-overlapping frequency bands only the average corresponding to one sequence of tonal stimuli was considered. The same procedure was applied to the results for the random event sequence (Fig. 1i). For the remaining 2 units (elements e and f in Fig. 1) the average dynamic spectrum for only the upper one of the two tonal sequences was available, as can be seen from the corresponding results.

3. From the APES to a Neural Characteristic

The Spectro-Temporal Receptive Field (STRF) of an auditory neuron has been formally defined as the frequency-time representation of the second order Volterra kernel of the neuron's nonlinear system function, transforming the acoustic stimulus $x(t)$ into the firing probability $y(t)$ (Aertsen and Johannesma, 1981). As such it constitutes a stimulus-invariant functional characteristic of the neuron, it may be viewed as a kind of spectro-temporal transfer function. It has been shown that for (1) a Gaussian stimulus ensemble and (2) a neuron's system function with no even order contributions of order higher than two, the second order cross correlation of stimulus and response, normalized with respect to the spectral contents of the stimulus ensemble, leads to the STRF. The normalization-procedure in this case is given by

$$\hat{v}_2(\omega, \nu) = \frac{\hat{R}_{xxy'}(\omega, \nu)}{2\hat{R}_{xx}(\omega)\hat{R}_{xx}(\nu)} \quad (1a)$$

where \hat{v}_2 denotes the double Fourier transform of the system's second order Volterra kernel $v_2(\sigma, \tau)$, \hat{R}_{xx} the power spectral density of the stimulus ensemble and $\hat{R}_{xxy'}$ the second order cross spectrum of stimulus x and the deviation of the response y from its average value during the stimulation: $y'(t) = y(t) - \bar{y}$. An equivalent formulation of (1) is given by

$$\hat{v}_2(\omega, \nu) = \frac{\hat{R}_{xxy'}(\omega, \nu)}{2\hat{R}_{xx}(\omega)\hat{R}_{xx}(\nu)} - \frac{\hat{R}_{xxy}(\omega, \nu)}{2\hat{R}_{xx}(\omega)\hat{R}_{xx}(\nu)} \quad (1b)$$

where the second term on the right-hand side differs only from zero for $\omega + \nu = 0$: for a Gaussian stimulus ensemble it holds that

$$\hat{R}_{xxy}(\omega, \nu) = \bar{y}\hat{R}_{xx}(\omega)\delta(\omega + \nu).$$

The APES has been introduced to provide an experimental spectro-temporal characterization of the neuron's stimulus-response relation. For every stimulus segment, preceding a neural event, the spectro-temporal structure is measured by a dynamic spectrum analyser and the resulting spectrograms are averaged for all events occurring during the presentation of the stimulus ensemble (Aertsen et al., 1980). The derivation of a mathematical relation between the STRF and the APES necessitates a more formal description of the latter concept. The goal is to arrive at a stimulus-normalization for the APES, inspired by the relation given in (1).

3.1. Formal Description of the APES

The Spectrogram. The operation of determining the dynamic power spectrum or spectrogram of a signal $x(t)$ by a DSA-type analyser can be formalised as follows. The output of the DSA consists of a set of non-negative functions of time $\{P_l(t); l=1, L\}$, shortly $P(l, t)$, representing the time-dependent intensity of the sound in neighbouring $1/3$ -octave bands with center frequencies ω_l . For each separate time function $P(l, t)$, which forms a particular cross section of the dynamic power spectrum $P(\omega, t)$ it holds that (cf. Appendix 2 in Aertsen and Johannesma, 1980)

$$P(l, t) = (f_l * x)^2(t) + (\widetilde{f_l * x})^2(t) \quad (2)$$

with

$f_l(t)$ = impulse response of l -th bandpass filter in the DSA

$*$ = convolution operator

\sim = Hilbert transform (e.g. Papoulis, 1962).

By introduction of the "analytic signal" $\phi_l(t) = f_l(t) + i\tilde{f}_l(t)$, with $i = \sqrt{-1}$ (Gabor, 1946; Ville, 1948) and using the relation (e.g. Deutsch, 1969)

$$\widetilde{f_l * x} = \tilde{f}_l * x = f_l * \tilde{x} \quad (3)$$

the relation (2) can be rewritten as

$$P(l, t) = \int_{-\infty}^{\infty} ds \int_{-\infty}^{\infty} ds' \phi_l(s) \phi_l^*(s') x(t-s) x(t-s') \quad (4a)$$

which expresses the spectrogram as a quadratic functional of the signal $x(t)$. Taking the Fourier transform of (4a) we arrive at its spectral equivalent:

$$\hat{P}(l, \mu) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda \hat{\phi}_l(\lambda) \hat{\phi}_l^*(\lambda - \mu) \hat{x}(\lambda) \hat{x}^*(\lambda - \mu) \quad (4b)$$

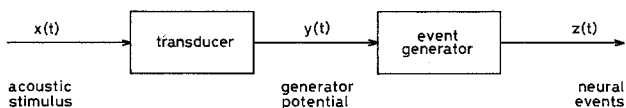


Fig. 3. Model of a primary auditory neuron (from Johannesma, 1980)

where $\hat{\cdot}$ indicates the Fourier transform and $\hat{\phi}_i^*$ denotes the complex conjugate of $\hat{\phi}_i$. The spectrum $\hat{\phi}_i(\lambda)$, associated with the analytic filter impulse response $\phi_i(t)$ equals (Gabor, 1946)

$$\hat{\phi}_i(\lambda) = 2\hat{f}_i(\lambda) \quad (\lambda > 0) \\ = 0 \quad (\lambda < 0). \quad (5)$$

The frequency integral in (4b) should be taken from $-\infty$ to $+\infty$, the properties of $\hat{\phi}_i(\lambda)$ imply that for $\mu \geq 0$ the result is equivalent to an integration from μ to $+\infty$. Furthermore, since $P(l, t)$ is real, it holds that $\hat{P}(l, -\mu) = \hat{P}^*(l, \mu)$.

Averaging of Spectrograms and Cross Correlation. The operation of averaging spectrograms is a particular form of statistical analysis of the stimulus-response relation of a neuron. The stimulus ensemble SE is presented to the animal and the electrical activity of a single neuron is measured extracellularly. The result is a sequence of action potentials, the neural events. Repetition of the experiment leads to another sequence of events, generally not identical to the first one. For our purposes the neural activity $z(t)$ can be modelled as a single realization of a stochastic point process:

$$z(t) = \sum_{n=1}^N \delta(t - t_n) \quad (6)$$

with

$\delta(t)$ = Dirac delta function

$\{t_n\}$ = moments of occurrence of action potentials

N = number of action potentials during presentation of the stimulus ensemble (duration T).

A simple model of the single neuron system is depicted in Fig. 3: a transducer followed by an event generator (De Jongh, 1978; Johannesma, 1971, 1980). The acoustic stimulus $x(t)$ is transformed by the transducer into the generator potential $y(t)$. The transducer mechanism can be described by a deterministic, time-invariant, continuous, dynamic nonlinear system S , i.e. a Volterra system. All stochastic aspects of the neural signal transformation are comprised in the event generator: the neural response function or firing probability $g(t)$ is a continuous, non-negative, algebraic function of the generator potential. Although the model, strictly speaking, only applies to primary auditory neurons, its formalism will be adopted here.

The event-density or PSTH $n(t)$, defined as the ensemble average $\langle z(t) \rangle$ of the neuron's firing pattern over a number of presentations of the same stimulus repertoire, is used as an experimental estimate of the firing probability density $g(t)$. It has been shown that the cross correlation function $R_{\Psi n}$ of any stimulus functional $\Psi(x)$ and the event density $n(t)$ equals the expected value of the stimulus-event correlation $\langle R_{\Psi z} \rangle$ (Johannesma, 1972, 1980):

$$\langle R_{\Psi z} \rangle = R_{\Psi n} = \text{experimental estimate of } R_{\Psi g}. \quad (7)$$

When we adopt the simple model of a *linear* event generator $g(y)$ (which implies that the transducer should be such that $y(t) \geq 0$ for all t , or, the possible nonlinearity in the firing probability has been absorbed in the transducer mechanism) this leads to

$$\langle R_{\Psi z} \rangle = R_{\Psi y}. \quad (8)$$

The "stochastic" stimulus-event correlation $R_{\Psi z}$ thus has been related to the "deterministic" Volterra system correlation $R_{\Psi y}$. A related approach has been used for an exponential event generator (Johannesma, 1980).

In case a long and wide stimulus ensemble is used, i.e. long and wide as compared to the range of $\langle R_{\Psi z} \rangle$, this effectively implies an averaging over the stochastics of the event generator. In that case the stimulus-event correlation over a single presentation of the SE forms a reasonable estimate of its expected value over a number of presentations:

$$R_{\Psi z} = \langle R_{\Psi z} \rangle. \quad (9)$$

Combining (7), (8), and (9) we have

$$R_{\Psi z} = R_{\Psi y}. \quad (10)$$

Or: under the assumption of a linear event generator the experimental result of stimulus-event correlation over a long and wide stimulus ensemble can be used as an estimate of the theoretical stimulus-response correlation of the deterministic nonlinearity S .

When applied to the case of the dynamic power spectrum $P(\omega, t)$ as stimulus functional Ψ this leads to the stimulus-event correlation R_{Pz} as the experimental estimate of the cross correlation of stimulus spectrogram and generator potential. In connection to (9) it may be noted that for a long ensemble of various natural stimuli the result of stimulus-event correlation R_{Pz} appeared to be highly reproducible upon repeated presentation of the stimulus ensemble (Aertsen et al., 1981). The stimulus ensemble SE, considered as a single time-function $x(t)$ with large duration T , gives for the cross correlation R_{Py}

$$R_{Py}(\omega, \tau) = \frac{1}{T} \int_0^T dt P(\omega, t - \tau) y(t). \quad (11)$$

In that case the SE is thought to be a long, single realization of a Gaussian process and furthermore the system S is time-invariant, the time-integral in (11) can be considered as an estimate of the expected value $E[P(\omega, t - \tau) y(t)]$ for the stochastic input process x and the corresponding output process y .

The experimentally obtained stimulus-event correlation $R_{Pz}(\omega, \tau)$, in analogy to (11), is given by

$$R_{Pz}(\omega, \tau) = \frac{1}{T} \int_0^T dt P(\omega, t - \tau) z(t) \quad (12)$$

with $z(t)$ as described in (6). The event character of $z(t)$ causes this integral to result in an averaging operation:

$$R_{Pz}(\omega, \tau) = \frac{1}{T} \sum_{n=1}^N P(\omega, t_n - \tau) \quad (13)$$

with t_n the moment of occurrence of the n -th action potential. This result leads to a more general interpretation of stimulus-event correlation.

The Average Pre-Event Spectrogram. The *Pre-Event Stimulus Ensemble* (PESE) is defined as the collection of stimuli preceding the neural events (Johannesma, 1972; Aertsen et al., 1980, 1981). As a consequence the PESE forms a subset of the SE, its structure reflecting the neuron's selective properties. We have investigated the spectro-temporal structure of the PESE by measuring the ensemble average of dynamic power spectra, associated with the individual pre-event stimuli: the *Average Pre-Event Spectrogram* (APES) $p_e(\omega, \tau)$ given by

$$p_e(\omega, \tau) = \langle P(\omega, \tau) \rangle_{\text{PESE}} = \frac{1}{N} \sum_{n=1}^N P_n(\omega, \tau) \quad (14)$$

where $\tau > 0$ indicates time *before* the neural event and $P_n(\omega, \tau)$ denotes the spectrogram of the n -th pre-event stimulus $x_n(\tau) = x(t_n - \tau)$. The comparison of (10), (13), and (14) immediately shows the following proportionality

$$\frac{N}{T} p_e(\omega, \tau) = R_{Py}(\omega, \tau). \quad (15)$$

The relations (6)–(15) thus lead to the conclusion that the product of the average firing rate N/T and the average pre-event spectrogram is equal to the cross correlation across the deterministic system S of the stimulus spectrogram and the generator potential.

For the practical implementation of the APES we return to the DSA-equation given in (4a) and (4b). The combination of (4) and (14) leads to the APES $p_e(l, \tau)$, given by

$$p_e(l, \tau) = \int ds \int ds' \phi_l(s) \phi_l^*(s') \cdot \langle x(t_n - \tau - s) x(t_n - \tau - s') \rangle_{\text{PESE}} \quad (16a)$$

or, in the frequency domain,

$$\hat{p}_e(l, \mu) = \frac{1}{2\pi} \int d\lambda \hat{\phi}_l^*(\lambda) \hat{\phi}_l(\lambda - \mu) \cdot \langle x_n^*(\lambda) x_n(\lambda - \mu) \rangle_{\text{PESE}}. \quad (16b)$$

Following (15) the APES forms an estimate of R_{Py} . As a consequence we have the theoretical relation

$$\frac{N}{T} p_e(l, \tau) = \frac{1}{T} \int_0^T dt \int ds \int ds' \phi_l(s) \phi_l^*(s') \cdot x(t - \tau - s) x(t - \tau - s') y(t). \quad (17)$$

Changing order of integration this leads to

$$\frac{N}{T} p_e(l, \tau) = \int ds \int ds' \phi_l(s) \phi_l^*(s') \cdot \frac{1}{T} \int_0^T dt x(t - \tau - s) x(t - \tau - s') y(t) \quad (18)$$

or, equivalently,

$$\frac{N}{T} p_e(l, \tau) = \int ds \int ds' \phi_l(s) \phi_l^*(s') R_{xxy}(\tau + s, \tau + s'). \quad (19a)$$

In the frequency domain this reads

$$\frac{N}{T} \hat{p}_e(l, \mu) = \frac{1}{2\pi} \int d\lambda \hat{\phi}_l^*(\lambda) \hat{\phi}_l(\lambda - \mu) \hat{R}_{xxy}(\lambda, \mu - \lambda). \quad (19b)$$

This implies that the APES $p_e(l, \tau)$ is a linear functional of the second order input-output cross correlation function: for every l -th DSA-channel (19a) describes the 2-dimensional convolution of the time-inverted cross correlation function R_{xxy} and the matrix $\phi_l \phi_l^*$ associated with the filter impulse response ϕ_l . The corresponding relation (19b) in the frequency domain thus may be viewed as a 2-dimensional DSA-filtering of the second order cross spectrum.

It should be noted that the transformation from R_{xxy} to the APES cannot be inverted, due to the projection-like integration in (19). This is quite a general characteristic of the sonograph-type of spectrum analysers, described by (4), of which the DSA forms an example: phase relations of different spectro-temporal components in the signal are not preserved.

3.2. Relation Between APES and STRF

For a Gaussian white stimulus ensemble the second order input-output correlation R_{xxy} is tightly connected to the system's second order Wiener kernel w_2 (Lee and Schetzen, 1965). In case the system does not contain even order nonlinearities of order higher than two the Wiener kernel w_2 is identical to the system's quadratic Volterra kernel v_2 ; in the general case w_2 can be considered as the second order "Wiener-

contribution" to the system kernel v_2 which forms the base of the formally defined STRF (Aertsen and Johannesma, 1981). The relation (19) implies that for a Gaussian white stimulus ensemble the APES $p_e(l, \tau)$ minus Average Spectrogram (AS) $p(l, \tau)$ of the overall stimulus ensemble, by its connection to the difference $R_{xxy} - R_{xxxy}$, represents the "DSA-mediated" second order Wiener contribution to the neuron's STRF. For the special class of systems with $v_2 = w_2$ the measurement of the *Difference Spectrogram* (DS) $q_e(l, \tau)$, defined by

$$q_e(l, \tau) = p_e(l, \tau) - p(l, \tau) \quad (20)$$

directly represents the DSA-filtered STRF. The magnitude of the higher even order nonlinearities in the neuron's system function determines the degree in which the DS, for a Gaussian white stimulus ensemble, represents a faithful image of the STRF. The distortion introduced by the use of stimulus ensembles which deviate from Gaussian white noise will now be considered.

The stimulus-normalization of the APES will be considered on the base of the normalization procedure of the STRF as given in (1), i.e. for linearly filtered Gaussian noise. A more general approach for nonlinearly transformed Gaussian noise, involving the use of higher order correlation functions, as indicated by Aertsen and Johannesma (1981) is beyond the scope of the present discussion.

The correlation function R_{xxy} can be expressed, by using (1), in terms of the system kernel v_2 and the stimulus autocorrelation R_{xx} . This results in a formal relation between the APES and the STRF. When stated in the frequency domain this leads to

$$\frac{N}{T} (\hat{p}_e(l, \mu) - \hat{p}(l, \mu)) = \frac{1}{\pi} \int d\lambda \hat{\phi}_l^*(\lambda) \phi_l(\lambda - \mu) \cdot \hat{R}_{xx}(\lambda) \hat{R}_{xx}(\mu - \lambda) \hat{v}_2(\lambda, \mu - \lambda). \quad (21)$$

The kernel $\hat{v}_2(\lambda, \mu - \lambda)$ appearing in this relation equals the double Fourier transform of $v_2(\tau + \sigma, \tau)$ with respect to τ and σ . It may be noted here that for Gaussian noise the Fourier transformed AS $\hat{p}(l, \mu)$ in fact represents a filter-dependent constant $p_l \delta(\mu)$, measuring the contribution to the power spectrum of the overall stimulus ensemble which falls within the passband region of the l -th DSA-filter. It can experimentally be estimated by measuring the average spectrogram of stimuli preceding a random sequence of events (cf. Figs. 1i and 2i for the tonal respectively the natural SE).

The left-hand side of (21) represents an experimental measure: the product of 1) the average firing rate N/T and 2) the deviation of the average pre-event spectrogram p_e from the a priori expected stimulus

spectrogram p . The right-hand side shows a frequency integral in which the following factors can be discerned: 1) the quadratic system kernel $\hat{v}_2(\lambda, \mu - \lambda)$, which is filtered 2-dimensionally by 2) the bi-spectral properties of the stimulus ensemble $\hat{R}_{xx}(\lambda) \hat{R}_{xx}(\mu - \lambda)$, the result of which in turn is filtered again by 3) the DSA-filter operator $\hat{\phi}_l^*(\lambda) \hat{\phi}_l(\lambda - \mu)$.

If in some way v_2 is known, for instance from correlation analysis of Gaussian white noise experiments, and if all the other v_{2n} are zero, the relation (21) offers the possibility to calculate the form of the APES for a non-white Gaussian stimulus ensemble. This result then may be compared to the experimental result of stimulus-event correlation in order to test the stimulus-invariance of the neuron's STRF. Another point of interest is the stimulus amplitude-dependence of the system characteristics. The expression (21) predicts that scaling of the SE-amplitude by a factor A results in a scaling of the product $N/T q_e(l, \tau)$ by a factor of A^4 . When the experimentally determined product of average firing rate and difference spectrogram does not show this scaling property the system may either be not second order in its even components and/or the STRF is inherently amplitude-dependent, for instance by an amplitude-dependent gain factor ("long-term adaptation").

As already noted the integral in (21) cannot simply be inverted to yield an explicit expression of the system kernel v_2 in terms of the spectrogram, the stimulus properties and the DSA-filter characteristics. Further simplification has to be obtained in order to acquire a practically useful stimulus-normalization procedure of the APES. This will involve some quite rigorous approximations which will be discussed in the following section.

4. Approximative "Stimulus-Normalization" of the APES

Under the given assumptions, the formal relations given in Sect. 3, which result in the integral equation (21), are exact. The "approximations" which have been made were "probabilistic" approximations: experimental estimators have been identified with theoretically expected values. In this section a "normalization" procedure will be developed, based on (21). The goal is to arrive at a procedure which can be implemented for practical use, e.g. in the form of a computer program. Several strong approximations will be made, which implies that it will not be possible to attain mathematical rigour at the various steps. We nevertheless consider it to be useful to reach an applicable "normalization" procedure, in order to explore the boundaries of useful application of stimulus-event characterization procedures as have been described here.

The quantity of basic interest in (21) is the system kernel v_2 . The first step thus is to reach an approximation where v_2 appears before the integral sign instead of being part of the integrand. If now 1) v_2 , in the bi-frequency domain, varies smoothly compared to the remainder of the integrand and if 2) the absolute value of this remainder, for each l , obtains a single sharp maximum, then an approximation is obtained by evaluating v_2 at the bi-spectral location of this maximum. Using the 1/3-octave shape of the DSA-filters and furthermore assuming that also the stimulus power spectrum \hat{R}_{xx} is behaving smoothly as compared to the filter characteristic $|\hat{\phi}_l|$ it follows that for the l -th DSA-channel the maximum is obtained for $\lambda_l = \omega_l + \mu$. As a consequence, under these subsequent smoothness assumptions relation (21) may be approximated by

$$\frac{N}{T} \hat{q}_e(l, \mu) = \hat{v}_2(\omega_l + \mu, -\omega_l) \frac{1}{\pi} \int d\lambda \hat{\phi}_l^*(\lambda) \hat{\phi}_l(\lambda - \mu) \hat{R}_{xx}(\lambda) \hat{R}_{xx}(\mu - \lambda) \quad (22)$$

or, in the form of a normalization procedure,

$$\hat{v}_2(\omega_l + \mu, -\omega_l) = \frac{\frac{N}{T} (\hat{p}_e(l, \mu) - \hat{p}(l, \mu))}{\frac{1}{\pi} \int d\lambda \hat{\phi}_l^*(\lambda) \hat{\phi}_l(\lambda - \mu) \hat{R}_{xx}(\lambda) \hat{R}_{xx}(\mu - \lambda)} \quad (23)$$

A mathematical problem which has been ignored here is the fact that both \hat{v}_2 and $\hat{\phi}_l$ are complex and oscillating functions of frequency. The main argument for still making the approximation (22) is the assumption that $|\hat{\phi}_l|$ is very sharp compared to the other functions, so the most prominent contribution to the integral comes from a narrow frequency band around the maximum of $|\hat{\phi}_l^*(\lambda)| |\hat{\phi}_l(\lambda - \mu)|$, i.e. $\lambda_l = \omega_l + \mu$. This point however, should be studied more carefully.

From (23) it follows that the Fourier transformed difference spectrogram DS should be divided by an integral expression which in principle can be evaluated, since all quantities in the integrand are amenable for experimental determination. For every l and every μ we have a different integral with gradually shifting filters and power spectra. Since, however, the numerical evaluation of all these integrals requires a large amount of computing, it is worthwhile to find out whether the integral in (23) can be expressed in terms of quantities which can simply be measured. First of all, the comparison of (4b) and (23) shows that the denominator in (23) can be interpreted as the bi-frequency representation of the DSA-spectrogram of the stimulus autocorrelation function $R_{xx}(\tau)$, considered as a signal. In other words, following (23) the

Difference Spectrogram has to be inversely filtered with a 2-dimensional filter composed of the DSA-filters and the "stimulus filter". Guided by intuition we have made a further approximation, leading from (23) to

$$\hat{v}_2(\omega_l + \mu, -\omega_l) = \frac{\frac{N}{T} (\hat{p}_e(l, \mu) - \hat{p}(l, \mu))}{\frac{1}{\pi} \hat{R}_{xx}(\omega_l + \mu) \int d\lambda \hat{\phi}_l^*(\lambda) \hat{\phi}_l(\lambda) \hat{R}_{xx}(\lambda)} \quad (24)$$

which implies that the spectrogram in the denominator of (23) has been factorized into a purely spectral component and a component representing temporal features of $R_{xx}(\tau)$. The spectral component

$$I(\omega_l) = \int d\lambda \hat{\phi}_l^*(\lambda) \hat{\phi}_l(\lambda) \hat{R}_{xx}(\lambda) \quad (25)$$

is simply the overall stimulus power spectrum as measured by the DSA at the filter center frequencies ω_l and thus is readily obtained experimentally [compare the Average Spectrogram $p(l, \tau)$]. The second component $\hat{R}_{xx}(\omega_l + \mu)$, contrarily to $I(\omega_l)$, is a function of μ and should be interpreted as the Fourier transform

$$\hat{R}_{xx}(\omega_l + \mu) = \mathcal{F}_{\tau \rightarrow \mu} R_{xx}(\tau) e^{-i\omega_l \tau} \quad (26)$$

or: it is the frequency representation of the ω_l -component of $R_{xx}(\tau)$. The combination of the two factors (25) and (26) can loosely be written as

$$I(\omega_l) \hat{R}_{xx}(\omega_l + \mu) = \mathcal{F}_{\tau \rightarrow \mu} I(\omega_l) e^{-i\omega_l \tau} R_{xx}(\tau) \quad (27)$$

which, when we consider $I(\omega_l)$ as the DSA-version of $\hat{R}_{xx}(\omega)$ in ω_l , shows some reminiscences of the complex spectro-temporal intensity density function of the autocorrelation $R_{xx}(\tau)$ (Rihaczek, 1968; Johannesma and Aertsen, 1979; Johannesma et al., 1981).

By a heuristic type of reasoning the integral relation between APES and system kernel, expressed in (21), has been transformed into the approximative relation (24), which can be considered to define an approximative "normalization"-procedure, based on experimentally obtainable quantities like the AS and its 1-dimensional equivalent, $I(\omega_l)$. The "stimulus-normalization" of the APES, described by (24), leading to the DSA-mediated STRF, can be considered to be composed of three different stages:

1. The subtraction from the APES of the a priori expected average spectrogram AS of the overall stimulus ensemble, leading to the *Difference Spectrogram* DS = APES - AS.

2. The division of each separate channel in the DS by the filter-dependent constant $I(\omega_l)$. This correction, in fact an "equalization" for the spectral composition

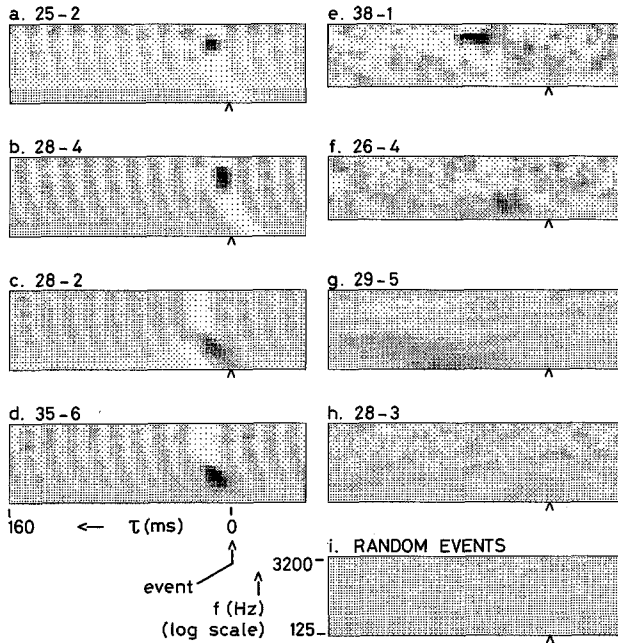


Fig. 4a-i. Equalized Difference Spectrogram (EDS) for a tonal stimulus ensemble of eight torus semicircularis units **a-h**, vs. the a priori expected results for this ensemble **i**. These Figs. represent the result of a partial “stimulus-normalization” procedure of the tonal APES’s in Fig. 1. Further explanation in the text

of the overall stimulus ensemble, was introduced earlier on purely intuitive grounds: the “stimulus-filtering” procedure (Aertsen et al., 1981). The result at this stage may be called the *Equalized Difference Spectrogram* (EDS).

3. The third stage, the division (in the frequency domain) of the EDS by the shifted spectrum $\hat{R}_{xx}(\omega_l + \mu)$ sort of demodulates the various channels in the EDS with respect to the spectral composition of the SE at the various frequencies within the passband region of the corresponding DSA-filter. The result after the third stage of the procedure will be called “*Normalized*” *Pre-Event Spectrogram*.

The acutes (') surrounding the adjective “normalized” should emphasize that the proposed procedure does not follow from straightforward theoretical considerations. Only the first stage, subtraction of the AS, follows from formal theory. The other two stages cannot lay a claim to mathematical rigour, rather they represent an intuitively based attempt to escape from the inherent limitations, imposed by an analysis based on DSA-type signal representations.

5. Results

In order to illustrate the effect of the “stimulus-normalization” procedure (24), developed in Sect. 4, we present the results of its application to the APES’s of 8

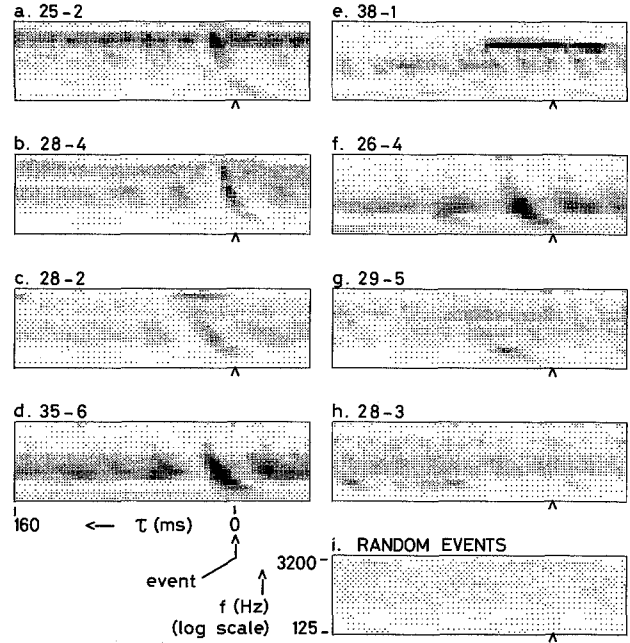


Fig. 5a-i. Equalized Difference Spectrogram (EDS) for a natural stimulus ensemble of eight torus semicircularis units **a-h**, vs. the a priori expected result for this ensemble **i**. These Figs. represent the result of a partial “stimulus-normalization” procedure of the natural APES’s in Fig. 2. Further explanation in the text

neurons from the torus semicircularis of the grassfrog, obtained for tonal and natural stimuli as shown in Figs. 1 respectively 2. The results after the second stage of “correction”, the Equalized Difference Spectrograms, are shown in Fig. 4 for tonal stimuli and Fig. 5 for natural stimuli. The final results, the “Normalized” Pre-Event Spectrograms are shown in Figs. 6 (tonal) and 7 (natural).

Analogously to the APES-results in Figs. 1 and 2 the elements a-h refer to various single neuron event sequences, whereas the element i shows the result for a random sequence of events. The latter results may give an indication regarding the overall effect of the procedure and the reliability of its results. Within each figure the results in the various elements were scaled using the same intensity scale; between figures the intensity scale generally differed to obtain a maximal resolution in the grey-scale representation. For a general discussion on the various phenomena encountered in these (ω, τ) -representations and their neurophysiological interpretation the reader is referred to Aertsen et al. (1980, 1981).

The comparison of APES (Fig. 1) and EDS (Fig. 4) for tonal stimuli shows that the effect of subtraction of the AS and subsequent equalizing in that case is not very drastic, as should be expected for a stimulus ensemble with a rather smooth spectral composition (cf. Fig. 1i). The main effect is an enhancement of low- ω_l regions as becomes evident from the flattening of

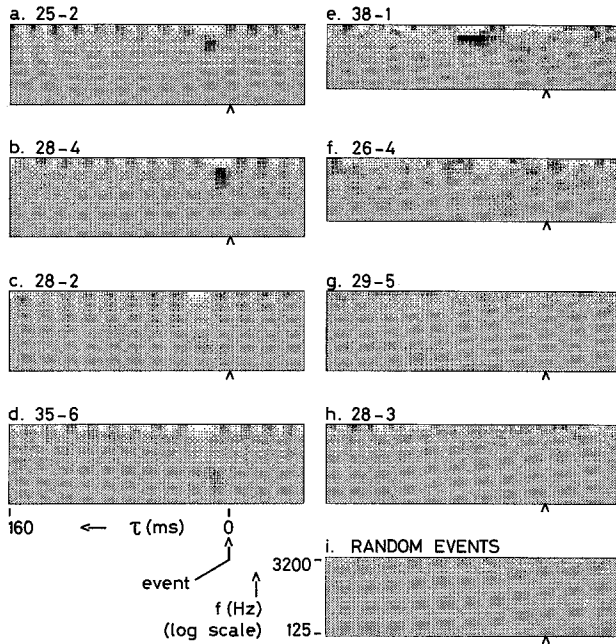


Fig. 6a-i. “Normalized” Pre-Event Spectrogram for a tonal stimulus ensemble of eight torus semicircularis units **a-h**, vs. the a priori expected result for this ensemble **i**. These Figs. represent the result of an approximative “stimulus-normalization” procedure of the tonal APES’s in Fig. 1. Compare also the corresponding EDS’s in Fig. 4. Further explanation in the text

the result for random events (compare Figs. 1i and 4i), which reflects the properties of the stimulus power spectrum as determined with the DSA. As a consequence the “activation areas”, i.e. areas in the (ω, τ) -domain with conspicuously higher values in the spectrogram than expected a priori, indicated by darker regions in the figures, for low-frequency units are enhanced somewhat, compared to the high-frequency units [e.g. units 35-6 (d) and 26-4 (f)].

The effect of equalizing for natural stimuli (compare Figs. 2 and 5) is more pronounced, as should be expected from a stimulus ensemble with an overall power spectrum which quite strongly varies with frequency [cf. Fig. 2i and also Fig. 6 in Aertsen et al. (1981)]. The strong contribution of certain frequency components in the natural SE, notably the 500 Hz component due to the grassfrog vocalizations, which dominates a number of APES’s [e.g. units 35-6 (d) and 28-3 (h)], is weakened. The same effect can be observed in the result for random events (Figs. 2i and 5i). This results in natural EDS’s which present, in their spectral aspects, a more balanced and smooth impression of the neuron’s selectivity.

On the whole these results are very similar to those already obtained by the “stimulus-filtering” procedure (Aertsen et al., 1981): the spectral preferences as reflected in the EDS’s for different stimulus ensembles

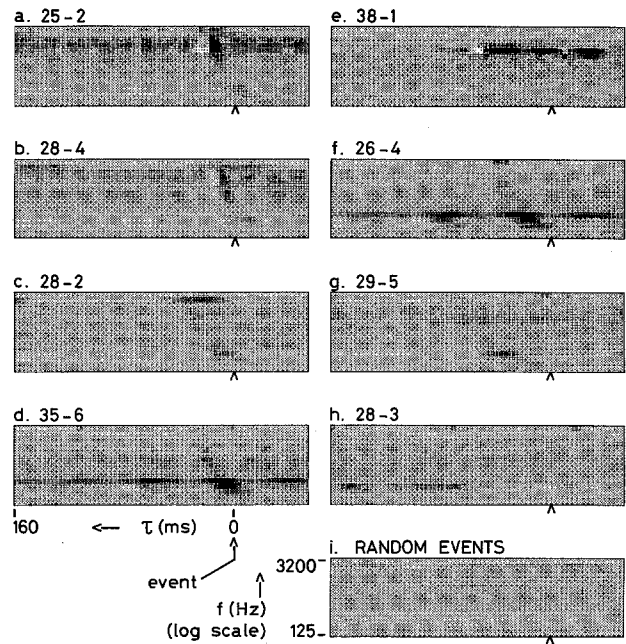


Fig. 7a-i. “Normalized” Pre-Event Spectrogram for a natural stimulus ensemble of eight torus semicircularis units **a-h**, vs. the a priori expected result for this ensemble **i**. These Figs. represent the result of an approximative “stimulus-normalization” procedure of the natural APES’s in Fig. 2. Compare also the corresponding EDS’s in Fig. 5. Further explanation in the text

(Figs. 4 and 5) show a better match than in the original APES’s (Figs. 1 and 2). As to be expected from a frequency equalizing procedure, it has no effect whatsoever on the temporal patterns in the APES, which are due to the specific temporal structure of both the tonal ensemble and the natural ensemble, combined with the spectro-temporal sensitivity of the various neurons.

The effect of the third stage of the “normalization”-procedure can be seen from a comparison of Figs. 4 and 6 for tonal stimuli, respectively Figs. 5 and 7 for natural stimuli. The general tendency here is that the high- ω_i regions in the EDS are enhanced compared to the low- ω_i regions. This is caused by the gradually decreasing values of $\hat{R}_{xx}(\omega_i + \mu)$ on a linear frequency scale with increasing ω_i . This decrease holds for both the tonal and the natural ensemble. This decrease of the denominator $\hat{R}_{xx}(\omega_i + \mu)$, which induces increasing numerical problems for increasing ω_i -channels in the EDS would not have occurred for stimulus ensembles with an overall “whiter” power spectral density. The general tendency in the results (Figs. 6 and 7), however, is consistent with a log-frequency representation (induced by the DSA) of a spectral density which should rather be viewed on a linear frequency scale (compare a $\log \omega$ representation of white noise). The “overrepresentation” of high- ω_i phenomena in the “normal-

ized" APES more or less suppresses the low-frequency activation areas, as becomes most apparent in the tonal results [e.g. units 28-2 (c), 35-6 (d), and 26-4 (f)], less prominent in the natural results although also there the effect is present. The effect of the third stage on the temporal structure in the EDS is hardly discernible and certainly not capable of removing the temporal stimulus properties. On the whole the results in Figs. 6 and 7 certainly cannot be called an improvement as compared to the equalization results. They much more tend to restore the gross characteristics of the original APES [e.g. the "natural" results of units 35-6 (d) and 26-4 (f)] than to join the spectro-temporal sensitivity as determined with tonal and natural stimuli into one single model.

6. Discussion

In the present paper we have investigated the formal relation between the Average Pre-Event Spectrogram (APES) and the Spectro-Temporal Receptive Field (STRF) of an auditory neuron. The APES has been introduced as an experimental measure of the neuron's spectro-temporal sensitivity for acoustic stimuli and in its implementation was based on the dynamic spectrum representation of sound (Aertsen and Johannesma, 1980; Aertsen et al., 1980, 1981). The STRF on the other hand forms a theoretical concept which aims to connect the general spectro-temporal representation of sound and the corresponding description of neural sensitivity, inspired by the work of Gabor (1946) and Ville (1948), with the general theory of nonlinear systems, as developed by Volterra, Wiener, and Lee (Aertsen and Johannesma, 1981).

The basic result is given by relation (21) which expresses the APES in the STRF in the form of a quadratic integral expression, which also contains the stimulus spectrum and the transfer function of the filters in the dynamic spectrum analyser (DSA). This relation is based on two assumptions: 1) the even order nonlinearity in the neuron's system function is of second order and 2) the stimulus ensemble is Gaussian. Under these assumptions the relation (21) enables to predict, for any Gaussian stimulus ensemble, the form of the APES, determined with a DSA-type analyser, from the knowledge of the STRF, the stimulus spectrum and the analyser's filter characteristics.

The inverse relation, a straightforward expression of the theoretical STRF in terms of the experimentally obtained APES is effectively rendered impossible, due to the "destructive" properties of the spectrogram analysis. The average pre-event spectrogram only presents partial information regarding the second order input-output cross correlation function, a necessary tool for the assessment of the quadratic system kernel,

the STRF. This implies that the use of DSA-type analysers to describe the spectro-temporal structure of sound and the neural sensitivity to it, does not allow the derivation of a rigorous stimulus-normalization procedure of the APES as was our original goal. The descriptive value of the APES for the characterization of neural sensitivity therefore is restricted to the type of stimuli it was determined with.

An approximative "stimulus-normalization" procedure has been proposed, based on the assumption of a DSA-analysis with bandpass filters which can be considered "sharp" as compared to the frequency dependence of the stimulus spectrum and the STRF. Quite apart from a number of mathematical difficulties associated with the "approximation" which we certainly do not consider to be solved, the results of application of this approximative normalization procedure were far from being impressive. In fact, the procedure could not be considered to give better results than an earlier, more intuitively based procedure of "equalizing" the APES for the overall power spectrum of the stimulus ensemble did give (Aertsen et al., 1981).

These results lead to the conclusion that in the context of a formal approach to the investigation of neuron characteristics like the spectro-temporal receptive field and its possible inherent stimulus-dependence, an instrument-based analysis like the APES shows severe shortcomings, which can only be overcome by using more refined and theoretically more firmly and elegantly based methods of spectro-temporal representation of sound. A possible example of such a method has been described as the complex spectro-temporal intensity density function CoSTID (Rihaczek, 1968; Johannesma and Aertsen, 1979; Johannesma et al., 1981). One of its essential characteristics is that, unlike the sonogram-analysis, it preserves phase-relations which are present in the signal. It has successfully been applied to the characterization of auditory neuron responses to Gaussian wideband noise (Hermes et al., 1981). Furthermore a stimulus-normalization procedure for the CoSTID has been developed, quite similar to relation (1), which appears to provide quite satisfactory results in neuron simulation studies, using as stimulus ensembles various mixtures of Gaussian wideband noise and natural sounds (in preparation).

The unsatisfactory results of the approximative "normalization"-procedure most probably have also causes in the two basic assumptions underlying a second order normalization procedure not being fulfilled. Both the tonal and the natural stimulus ensemble are certainly far from Gaussian, whereas a reliable statement regarding the order of the even part of the neuron's nonlinearity cannot be given. The natural

stimulus ensemble contains very specific spectro-temporal structures, is too specific and too complex to be adequately described by its overall power spectrum. The same holds, albeit to a lower extent, for the tonal ensemble. This leads to the conclusion that neuron characteristics determined for these types of stimulus ensembles have, strictly speaking, only descriptive value when remaining within the same type of stimulus ensemble. This is characteristic of the problems one faces when trying to reconcile the various approaches to the study of neural function, especially in the more central regions of the auditory system. From neuroethology we have the demands which lead to "complex" and (species-) specific natural stimuli. The "classical" approach in auditory neurophysiology leads to the use of "simple" stimuli like tones. A general theoretical framework is only provided for Gaussian noise stimuli. This creates a field of tension where adequate strategies have to be developed. Several ways to proceed are conceivable. One possible way is to remain within the relative safety of general theory as long as possible, i.e. to start from Gaussian white noise and gradually transform it in various possible ways e.g. by linear and/or nonlinear filtering or the progressive addition of non-Gaussian signals. This approach, in fact, sets out to explore the experimental boundaries of applicability of the general theory. An attractive alternative is to start from the inspiration of neuroethology, where relevant questions are formulated in terms of concepts like "information bearing elements" (Suga, 1978) or "bio-acoustic dimensions" (Scheich, 1977), leading to a type of "parameter variation" of the interesting aspects of natural sounds. The question whether this approach can be formalized into a general framework, making use of the various possibilities of signal and system theory available, certainly forms an interesting challenge for any neurophysiologist with a soft spot for the elegance of mathematics and biophysics. Whatever the approach one chooses, the chances for success will certainly be enlarged by using all the available neurophysiological evidence and turning it into possible models of neural function. The inherent weakness of a complete "black box" approach, both factually and conceptually, would thereby be reduced to a large extent. Our final conclusion would be that the dynamic spectrum analysis can play a useful role in the auditory neurophysiological experiments where rapid, preferably real-time, feedback to the experimenter may indicate what to look for and where, in the spectro-temporal domain, the questions should be focussed. Once the interesting regions have been located, the characteristics have to be investigated by more adequate and precise, at the same time more (computer) time consuming spectro-temporal procedures.

Acknowledgements. This investigation forms part of the research program of the Workgroup Neurophysics at the Department of Medical Physics and Biophysics of the University of Nijmegen (the Netherlands) and was supported by the Netherlands Organization for Advancement of Pure Research (ZWO). The authors wish to thank Ton Vendrik for helpful comments on the manuscript. Technical support has been given by Jan Bruijns and Wim van Deelen.

References

- Aertsen, A.M.H.J., Johannesma, P.I.M.: Spectro-temporal receptive fields of auditory neurons in the grassfrog. I. Characterization of tonal and natural stimuli. *Biol. Cybern.* **38**, 223-234 (1980)
- Aertsen, A.M.H.J., Johannesma, P.I.M., Hermes, D.J.: Spectro-temporal receptive fields of auditory neurons in the grassfrog. II. Analysis of the stimulus-event relation for tonal stimuli. *Biol. Cybern.* **38**, 235-248 (1980)
- Aertsen, A.M.H.J., Olders, J.H.J., Johannesma, P.I.M.: Spectro-temporal receptive fields of auditory neurons in the grassfrog. III. Analysis of the stimulus-event relation for natural stimuli. *Biol. Cybern.* **39**, 195-209 (1981)
- Aertsen, A.M.H.J., Johannesma, P.I.M.: The spectro-temporal receptive field: a functional characteristic of auditory neurons. *Biol. Cybern.* **42**, 133-143 (1981)
- Bullock, T.H. (ed.): Recognition of complex acoustic signals. *Life Sciences Res. Rep.*, Vol. 5. Berlin, Heidelberg, New York: Springer 1977
- Deutsch, R.: System analysis techniques. Englewood Cliffs, N.J.: Prentice Hall 1969
- Gabor, D.: Theory of communication. *J.I.E.E. (London)*, pt. III, **93**, 429-457 (1946)
- Hermes, D.J., Aertsen, A.M.H.J., Johannesma, P.I.M., Eggermont, J.J.: Spectro-temporal characteristics of single units in the auditory midbrain of the lightly anaesthetized grassfrog, *Rana temporaria* L., investigated with noise stimuli. *Hearing Research* (in press) (1981)
- Johannesma, P.I.M.: Dynamical aspects of the transmission of stochastic neural signals. In: *Proc. First European Biophysics Congress*, pp. 329-333. Broda, E., Locker, A., Springer-Lederer, H. (eds.). Vienna: Verlag der Wiener Medizinischen Akademie 1971
- Johannesma, P.I.M.: The pre-response stimulus ensemble of neurons in the cochlear nucleus. In: *Proc. of the IPO Symp. on Hearing Theory*, pp. 58-69. Cardozo, B.L. (ed.). Eindhoven, the Netherlands 1972
- Johannesma, P.I.M., Aertsen, A.M.H.J.: Neural image of sound in the grassfrog. In: *Hearing Mechanisms and Speech*, pp. 79-86. Creutzfeld, O., Scheich, H., Schreiner, Chr. (eds.). *Exp. Brain Res. Suppl.* II (1979)
- Johannesma, P.I.M.: Functional identification of auditory neurons based on stimulus-event correlation. In: *Psychophysical, physiological and behavioural studies in hearing*, pp. 77-84. Brink, G. van den, Bilsen, F.A. (eds.). Delft: University Press 1980
- Johannesma, P.I.M., Aertsen, A.M.H.J., Cranen, L., Erning, L.J.Th.O. van: The Phonochrome: a coherent spectro-temporal representation of sound. *Hearing Research* (in press) (1981)
- Jongh, H.R. de: Modelling the peripheral auditory system. Fitting model pulse patterns to neuronal impulse trains. Ph. D. Thesis, Amsterdam, the Netherlands 1978
- Lee, Y.W., Schetzen, M.: Measurement of the Wiener kernels of a nonlinear system by cross-correlation. *Int. J. Control.* **2**, 237-254 (1965)

- Papoulis, A.: The Fourier integral and its applications. New York: McGraw-Hill 1962
- Rihaczek, A.W.: Signal energy distribution in time and frequency. IEEE Trans. Inf. Theory IT-14, 369-374 (1968)
- Scheich, H.: Central processing of complex sounds and feature analysis. In: Recognition of complex acoustic signals, pp. 161-182. Bullock, T.H. (ed.). Life Sciences Res. Rep., Vol. 5. Berlin, Heidelberg, New York: Springer 1977
- Suga, N.: Specialization of the auditory system for reception and processing of species-specific sounds. Fed. Proc. Fed. Am. Soc. Exp. Biol. 37, 2342-2354 (1978)
- Ville, J.: Théorie et applications de la notion de signal analytique. Cables Transm. 2, 61-74 (1948)
- Worden, F.G., Galambos, R. (eds.): Auditory processing of biologically significant sounds. Neurosci. Res. Prog. Bull., Vol. 10. Brookline, MA.: Neurosciences Research Program 1972

Received: June 19, 1981

Dr. A. M. H. J. Aertsen
Department of Medical Physics and Biophysics
University of Nijmegen
Geert Grooteplein Noord 21
NL-6525 EZ Nijmegen
The Netherlands