

Spectro-Temporal Receptive Fields of Auditory Neurons in the Grassfrog

I. Characterization of 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. In order to investigate the spectro-temporal selectivity of single auditory neurons in the auditory midbrain of the grassfrog under different stimulus conditions a stimulus ensemble consisting of *tonal* and *natural* stimuli has been constructed. The selection of characteristics of the tonal stimuli was inspired by the general properties of vocalizations of the male grassfrog. The natural stimuli contained a variety of vocalizations of various anuran species and other sounds, taken from the natural environment of the grassfrog (i.e. its biotope). The main point of interest in the present paper is the *spectro-temporal characterization* of acoustic stimuli. For narrowband signals a parametric description in terms of *instantaneous amplitude* and *instantaneous frequency* appears to be attractive. A more general approach, also covering wideband signals, is a description by means of the *dynamic power spectrum*. Both types of analysis are described. Mathematically they can be based on the theory of *analytic signals*. The basic principles of this theory are summarized.

1. Introduction

The neural representation of the acoustic environment is investigated predominantly by the study of single neuron behaviour. Different types of stimuli (e.g. tones, clicks, noise) are presented to the animal and the pattern of action potentials, generated by the neuron, is studied as a function of one or more parameters of the stimulus.

Widely used methods to describe the single neuron behaviour are the tuning curves (Kiang et al., 1965), PST-histograms (Kiang et al., 1965), and the reverse correlation function (De Boer and De Jongh, 1978). Spectral selectivity, amplitude sensitivity and the time constants involved are the characteristics of prime importance.

On the other hand ethological investigations have revealed the behavioural importance of spectro-temporal properties of (species-specific) vocalizations. The investigation of the underlying neural mechanism has received considerable attention (e.g. Worden and Galambos, 1972; Bullock, 1977).

The question to what extent the single neuron response to one type of stimuli, say natural sounds, can be understood from the response to another type of stimuli, say tones, has often been considered on a rather qualitative level (e.g. Smolders et al., 1979). The present investigation was undertaken to study this question in more detail. This paper (Part I) will be concerned with a description of the stimulus ensembles which were used: tonal stimuli and natural stimuli. In a companion paper (Part II, Aertsen et al., 1980) we introduce procedures to analyse the spectro-temporal properties of stimuli which influence the neural firing pattern: the neuron's *spectro-temporal receptive field*. A description of the procedures will be given, as well as the results of their application to a number of single unit responses to an ensemble of tonal stimuli. In a forthcoming paper (Part III) we will use one of these procedures to investigate the spectro-temporal receptive fields for natural stimuli.

The animal under investigation is the grassfrog, *Rana temporaria* L., which is quite common in Western Europe. For several anuran species both the spectral (e.g. Capranica, 1965; Gerhardt, 1974) and temporal (e.g. Van Gelder et al., 1978; Gerhardt, 1978) properties of species-specific vocalizations have been shown to be of behavioural significance. These findings, amongst others, suggest that the approach to investigate the frog's spectro-temporal receptive field properties under different stimulus conditions may provide information regarding the neural basis of behaviour in the natural acoustic environment.

A necessary prerequisite for this approach is a detailed description of the *spectro-temporal* properties

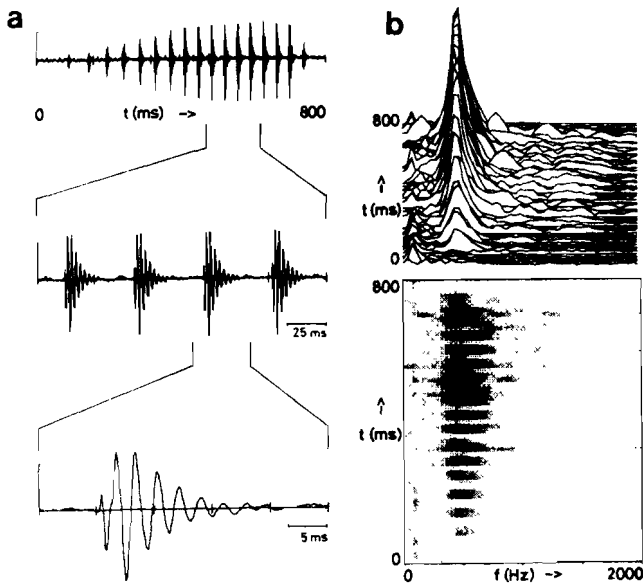


Fig. 1a and b. The “B-call” of the male grassfrog (*Rana temporaria* L.). **a** Time course of vocalization on different time scales, indicated by horizontal bars. Vertical bars refer to identical, arbitrary amplitude values. Note the quasiperiodic, pulsating character of the call. **b** Sonogram (dynamic amplitude spectrum) of vocalization in different representations, showing the spectro-temporal structure of the call: a dominant frequency band around 500 Hz and less intensive bands around 1100 Hz and 1500 Hz. Frequency resolution 39.1 Hz, time resolution 12.8 ms

of the stimulus ensemble. This analysis can be implemented in different ways, depending on the type of stimuli (narrowband vs. wideband). A description of the procedures which have been used to represent the stimuli of interest, *tonal* and *natural* sounds, is the main point of interest in the present paper. A comparison is made between on the one hand a *parametric* description in terms of temporal envelope and instantaneous frequency, which can be applied to narrowband stimuli, and on the other hand a spectro-temporal representation in the form of a *dynamic power spectrum*, applicable to an arbitrary signal.

2. The Tonal Stimulus Ensemble

2.1. Gamma-Tones

The selection of tonal stimuli, used in the present investigation, was inspired by the vocal repertoire of the male grassfrog. A typical example of a vocalization is shown in Fig. 1a. In general these vocalizations can be considered as quasi-periodic sequences of short similar sound elements (Van Gelder et al., 1978). Spectral analysis of the calls reveals a dominant fre-

quency band in the range of 300–600 Hz, with smaller contributions from frequency bands in the ranges of 1000–1400 Hz and 1400–1900 Hz (Brzoska et al., 1977). These figures provide only a rough indication; they vary for the different types of calls. The characteristics of female calls as well as more complex patterns appearing in some male calls (e.g. frequency modulation) are not considered here. Significant correlations between temperature and call duration, number of sound elements in the calls as well as the repetition rate of sound elements have been demonstrated (Schneider, 1973; Van Gelder et al., 1978).

The male vocalization shown in Fig. 1a is classified as a “B-call” on the base of its physical characteristics: duration of the call and the number of sound elements (Van Gelder et al., 1978). The ethological context of this call leads to its classification as a “mating call” (Schneider, 1973; Brzoska, 1977) or a “chorus call” (Van Gelder et al., 1978). The dynamic spectrum or sonogram of this call, shown in Fig. 1b in two representations, shows the dominant frequency band around 500 Hz, with additional higher frequency bands at about 1100 Hz and 1500 Hz, which are due mainly to the “transient” onset of the sound elements.

An approximative formal description of a single sound element from the vocalizations is given by the amplitude modulated waveform

$$e(t) = Am(t) \cos(\omega t + \phi), \quad (2.1)$$

where A = amplitude factor; $m(t)$ = amplitude modulator; ω = angular carrier frequency; ϕ = phase angle.

For the amplitude modulator $m(t)$ we propose

$$m(t) = c \left(\frac{t}{\beta} \right)^{\gamma-1} \exp\left(-\frac{t}{\beta} \right) \quad t \geq 0 \quad \beta > 0 \\ = 0 \quad t < 0 \quad \gamma \geq 1, \quad (2.2)$$

where β = duration parameter; γ = form parameter; $c = \left(\frac{2^{2\gamma-1}}{\beta \Gamma(2\gamma-1)} \right)^{1/2}$ normalization constant, to give $m^2(t)$ unit area; $\Gamma(z)$ = Gamma function (e.g. Abramowitz and Stegun, 1972).

The form $m(t)$ appears both as the integrand in the definition of the Gamma function $\Gamma(\gamma)$ and as the density function of the Gamma distribution (e.g. Cox and Lewis, 1966), therefore we propose to use for (2.1) the term “Gamma-tone” or “ γ -tone”. The choice of the particular waveform (2.1) with $m(t)$ specified by (2.2) was inspired, amongst others, by its use in a parametric description of impulse responses for single units in the cat’s auditory nerve (De Boer, 1973) and cochlear nucleus (Johannesma, 1972; Grashuis, 1974), as obtained by the “reverse correlation” method (De Boer

and De Jongh, 1978). The introduction of this type of waveform goes back to Flanagan (1965) who used it in modelling the impulse response for the human basilar membrane.

Numerical values of the parameters which provide a reasonable approximation for the calls of the male grassfrog are

$$\begin{aligned}\beta &= 1.45 \text{ ms} \\ \gamma &= 3 \\ \omega/2\pi &= 500 \text{ Hz} \\ \phi &= -\pi/2.\end{aligned}\quad (2.3)$$

The stimulus ensemble in the present investigation consisted of sequences of γ -tones with constant duration parameter β and form parameter γ [see (2.3)] and different values of frequency ω and amplitude factor A . The γ -tone sequences are described in detail in Sect. 4. A segment of such a sequence is shown in Fig. 2a.

2.2. Complex Representation of Signals

In the analysis of single unit responses to the γ -tone ensemble we will use a description of the stimulus ensemble in terms of nonlinear functionals of the signal $x(t)$: the *instantaneous amplitude or temporal envelope* $a(t)$, the *instantaneous phase* $\phi(t)$ and the *instantaneous angular frequency* $\omega(t)$ [or the instantaneous frequency

$$f(t) = \frac{\omega(t)}{2\pi}].$$

These functionals can be defined elegantly by means of the concept of the *analytic signal* (Gabor, 1946; Ville, 1948). Using this convention a signal representation in terms of instantaneous amplitude $a(t)$ and phase $\phi(t)$ bears a one-to-one relation to a representation in terms of signal values $x(t)$. A signal representation in terms of instantaneous amplitude $a(t)$ and frequency $\omega(t) = 2\pi f(t)$ determines the signal apart from its absolute phase. In the case of a tonal stimulus ensemble the amplitude-frequency representation provides us with a framework intuitively suited to the analysis of the auditory system, where amplitude sensitivity and frequency selectivity are considered as important characteristic parameters (e.g. Kiang et al., 1965).

The *analytic signal* $\xi(t)$ associated with the real signal $x(t)$ is constructed by adding an imaginary component $i\tilde{x}(t)$ to the given signal $x(t)$, where $i = \sqrt{-1}$. As a consequence the analytic signal is a complex valued function of time, defined by

$$\xi(t) = x(t) + i\tilde{x}(t). \quad (2.4)$$

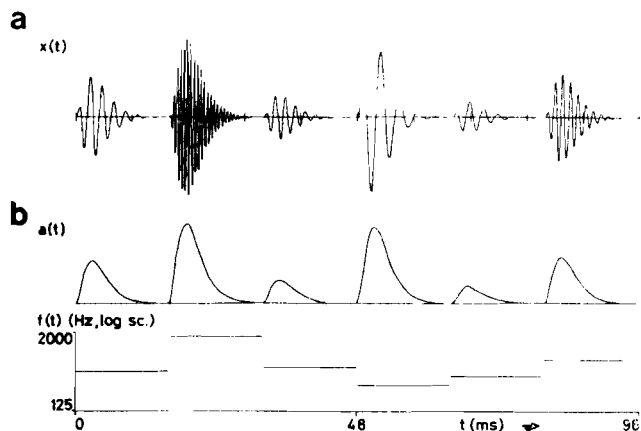


Fig. 2a and b. Segment from a sequence of γ -tones. **a** Signal $x(t)$ as a function of time. **b** Upper trace: instantaneous amplitude $a(t)$. Lower trace: instantaneous frequency $f(t) = \frac{1}{2\pi}\omega(t)$. Further explanation in text

The real function $\tilde{x}(t)$, associated with $x(t)$, is defined as the Cauchy principle value of the integral

$$\tilde{x}(t) = (\mathcal{H}x)(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} ds \frac{x(s)}{t-s}. \quad (2.5)$$

This linear integral transform, in fact the convolution of $x(t)$ and the integral kernel $\frac{1}{\pi t}$ is known as the *Hilbert transform* (Thitchmarsh, 1937). There is a discrepancy in the sign of the definition of $\tilde{x}(t)$ between Gabor on the one hand and Ville, Thitchmarsh on the other hand. In (2.5) we conform to the latter.

The analytic signal (2.4) is also referred to as the *pre-envelope* (Dugundji, 1958). The function $\tilde{x}(t)$ is referred to as the signal *in quadrature* to $x(t)$: x and \tilde{x} are mutually orthogonal.

The reciprocal relation holds

$$x(t) = (\mathcal{H}^{-1}\tilde{x})(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} ds \frac{\tilde{x}(s)}{t-s} \quad (2.6)$$

from which it follows that

$$\tilde{\tilde{x}}(t) = -x(t) \quad (2.7a)$$

or, equivalently,

$$(i\mathcal{H})^2 = \mathbf{1}, \quad (2.7b)$$

where $\mathbf{1}$ denotes the identity operator.

For the harmonic oscillation

$$x(t) = A_0 \cos(\omega_0 t + \phi_0) \quad (2.8a)$$

the quadrature signal $\tilde{x}(t)$ is

$$\tilde{x}(t) = A_0 \sin(\omega_0 t + \phi_0) \quad (2.8b)$$

and the analytic signal is

$$\xi(t) = A_0 \exp i(\omega_0 t + \phi_0). \quad (2.8c)$$

In the frequency domain the following relation holds

$$\hat{\xi}(\omega) = (1 + \text{sgn}(\omega)) \hat{x}(\omega), \quad (2.9)$$

where

$$\text{sgn}(\omega) = \begin{cases} +1 & \omega > 0 \\ 0 & \omega = 0 \\ -1 & \omega < 0 \end{cases} \quad (2.10)$$

and $\hat{x}(\omega)$ denotes the Fourier transform of $x(t)$.

For negative frequencies the spectrum of the analytic signal $\xi(t)$ equals zero, for positive frequencies it has the same form as the spectrum of the original signal $x(t)$.

Since the analytic signal $\xi(t)$ is a complex function it can be written in the form

$$\xi(t) = a(t) \exp i\phi(t) = \exp(\alpha(t) + i\phi(t)), \quad (2.11)$$

where the *instantaneous amplitude* or *temporal envelope* is

$$a(t) = |\xi(t)| = (\xi^*(t) \xi(t))^{1/2}, \quad (2.12)$$

the *logarithm of the instantaneous amplitude* is

$$\alpha(t) = \ln a(t) = \text{Re} \ln \xi(t) = \frac{1}{2} \ln(\xi^*(t) \xi(t)) \quad (2.13)$$

and the *instantaneous phase* is

$$\phi(t) = \arg \xi(t) = \text{Im} \ln \xi(t) = \frac{1}{2i} \ln(\xi^{*-1}(t) \xi(t)). \quad (2.14)$$

In these equations $\xi^*(t)$ denotes the complex conjugate of $\xi(t)$. The *instantaneous angular frequency* $\omega(t)$ is defined as the time derivative of the instantaneous phase

$$\omega(t) = \frac{d}{dt} \phi(t) = \text{Im} \frac{d}{dt} \ln \xi(t). \quad (2.15)$$

Analogously the *relative amplitude modulation* $\beta(t)$ is given by the time derivative of the logarithm of the instantaneous amplitude

$$\beta(t) = \frac{d}{dt} \alpha(t) = \text{Re} \frac{d}{dt} \ln \xi(t). \quad (2.16)$$

The *temporal intensity (density)* $I(t)$ is defined as

$$I(t) = a^2(t) = \xi^*(t) \xi(t), \quad (2.17)$$

the *spectral intensity (density)* is defined as

$$J(\omega) = \hat{\xi}^*(\omega) \hat{\xi}(\omega). \quad (2.18)$$

In this terminology we conform to Middleton (1960).

Application of these definitions to the example of the harmonic oscillation (2.8) leads to

$$\begin{aligned} x(t) &= A_0 \cos(\omega_0 t + \phi_0) \\ a(t) &= A_0 \quad \phi(t) = \omega_0 t + \phi_0 \\ \alpha(t) &= \ln A_0 \quad \omega(t) = \omega_0 \\ \beta(t) &= 0 \\ I(t) &= A_0^2. \end{aligned} \quad (2.19)$$

We conclude that for the harmonic oscillation the formal generalized definitions lead to expected results.

2.3. Parametric Representation of the Gamma-Tone Ensemble

In this section the notions as introduced in Sect. 2.2 will be applied to the γ -tone ensemble. This approach is simplified by a product theorem for Hilbert transforms (Tricomi, 1951; Bedrosian, 1963; Cain, 1973), which states

$$(\mathcal{H}g_1 g_2)(t) = g_1(t) (\mathcal{H}g_2)(t) \quad (2.20)$$

provided that at least one of the following two conditions is fulfilled:

1. the spectral intensity of $g_2(t)$ is confined to a frequency range which is located entirely above that for $g_1(t)$,

2. both $g_1(t)$ and $g_2(t)$ are analytic signals.

Amplitude modulated tones can be represented by

$$x(t) = Am(t) \cos(\omega_0 t + \phi_0). \quad (2.21)$$

If the carrier frequency ω_0 is larger than the maximum frequency present in $m(t)$, condition 1 is fulfilled and as a consequence the analytic signal is given by

$$\xi(t) = Am(t) \exp i(\omega_0 t + \phi_0). \quad (2.22)$$

If this is not strictly the case, but still ω_0 is large with respect to the bandwidth of $m(t)$, (2.22) may be used as an approximation with a correction term which decreases in magnitude as ω_0 becomes higher (Lerner, 1960; Rubin and Di Franco, 1963). For narrowband AM-signals (2.22) approximates fairly well the analytic signal associated with (2.21). This result can be generalized to include slow phase- or frequency modulation (Lerner, 1960; Rihaczek, 1966).

From (2.22) it follows that the instantaneous amplitude $a(t)$ and angular frequency $\omega(t)$ are given by

$$\begin{aligned} a(t) &= Am(t) \\ \omega(t) &= \omega_0. \end{aligned} \quad (2.23)$$

The γ -tone ensemble used in the present study is described by

$$x(t) = \sum_{k=0}^{K-1} A_k U_T(t-kT) m(t-kT) \cos(\omega_k(t-kT) + \phi). \quad (2.24)$$

where K = total number of γ -tones in the stimulus ensemble; A_k = amplitude factor of k -th γ -tone; $U_T(t)$ = rectangular time window; $U_T(t) = 1, 0 \leq t \leq T$; $U_T(t) = 0$, elsewhere; $m(t)$... amplitude modulator as specified by (2.2); ω_k = carrier frequency of k -th γ -tone; T = time interval between onsets of γ -tones.

It can be verified that, to a good approximation, this ensemble can be considered as an ensemble of narrowband signals for all parameter values involved (cf. Sect. 4.1 and Appendix 1). As a consequence the *instantaneous amplitude* or *temporal envelope* $a(t)$ is given approximately by

$$a(t) = \sum_{k=0}^{K-1} A_k U_T(t-kT) m(t-kT) \quad (2.25)$$

and the *instantaneous frequency* $\omega(t)$ can be approximated by

$$\omega(t) = \sum_{k=0}^{K-1} \omega_k U_T(t-kT). \quad (2.26)$$

These two parameters, both a function of time, are shown for a segment of the γ -tone ensemble in Fig. 2b. The formal definitions lead to results which agree to our intuitive notions. However, the intuitive definitions of instantaneous amplitude and frequency only apply to signals of a simple tonal character, whereas these formally generalized definitions of amplitude and frequency [(2.12) resp. (2.15)] are applicable to any type of signal.

The parametric analysis described in this paper is based on the concepts of instantaneous frequency and instantaneous amplitude, defined as functionals of the analytic signal. The Hilbert transform, and as a consequence the analytic signal and all functions derived from it, are functionals of the real signal $x(t)$, where both the past and future of $x(t)$ have to be evaluated [see (2.5)]. Especially the definition of "instantaneous frequency", because of this "non-causal" nature, has given rise to considerable debate in the literature (e.g. Mandel, 1974; Gupta, 1975, and references therein). These particular definitions have been used here because of their mathematical elegance and because they have proven to be useful in signal and system theory, both linear and nonlinear (e.g. Rihaczek, 1968; Deutsch, 1969). The notion of *instantaneous frequency* $\omega(t)$ on one hand and *Fourier-frequency* ω on the other hand should not be confused. The instantaneous frequency $\omega(t)$ is a time dependent functional of a signal, whereas the spectral variable ω is the independent variable in the domain, complementary to the temporal domain, associated with it by means of the Fourier transform.

3. The Natural Stimulus Ensemble

The natural environment in which an animal lives, as far as functionally related to the animal, is defined as its *habitat* or *biotope* (Dahl, 1921). It encompasses both the biotic and abiotic environment. The *Sensory Biotope* may be defined as the total of sensory aspects of the biotope. It consists of several components, each one corresponding to a specific sensory modality of the animal under investigation. The acoustical component is of special interest here: the *Acoustic Biotope* (Aertsen et al., 1979; Smolders et al., 1979). The Acoustic Biotope contains different kinds of sounds, each with an associated probability of occurrence. In general it will not be possible to give an explicit mathematical expression for the probability density of the Acoustic Biotope, as can be done for a technically based stimulus ensemble like Gaussian noise. For any particular realization of the Acoustic Biotope, however, it is possible to evaluate numerically the relative frequency of occurrence of the various sounds. Such a realization may be considered as a wide selection of natural sounds, the probabilistic structure of which more or less truly portrays the Acoustic Biotope. The Acoustic Biotope thus is considered as a probabilistic generator of stimulus ensembles as particular realizations, comparable to the Gaussian probability density function and the particular noise signals generated by it.

For the present investigation a realization of a wide stimulus ensemble, inspired by ethological and ecological observations was made. It consisted of a selection of sounds, recorded from the natural environment of the animal under investigation, i.e. the grassfrog. The realization was composed according to the scheme as shown in Diagram 1.

It consisted of 32 different sound segments, each one with a duration of about 8 s. The figures in Diagram 1 specify the numbers of sounds taken from different categories. Next to the "A-, B-, and C-calls" from the *male grassfrog* (typology according to Van Gelder et al., 1978), the stimulus ensemble contained such various sounds as the vocalizations of several *other anuran species* [e.g. the green frog (*Rana esculenta* L.), the moorfrog (*Rana arvalis* Nilsson), the European treefrog (*Hyla arborea* L.) and the common toad (*Bufo bufo* L.)], sounds from *predators* (e.g. the heron, the gull), *prey* (the fly, the cricket) and *other animals* (e.g. the crow) as well as sounds from *abiotic origin* (e.g. water, wind, rain). The 32 segments of sound were concatenated in a random order on analog tape, thus constituting a stimulus ensemble lasting about 4 min, its bandwidth being 5 kHz. In the following this realization will be denoted briefly as the "acoustic biotope". Different sounds from this stimulus ensemble are shown in Fig. 3.

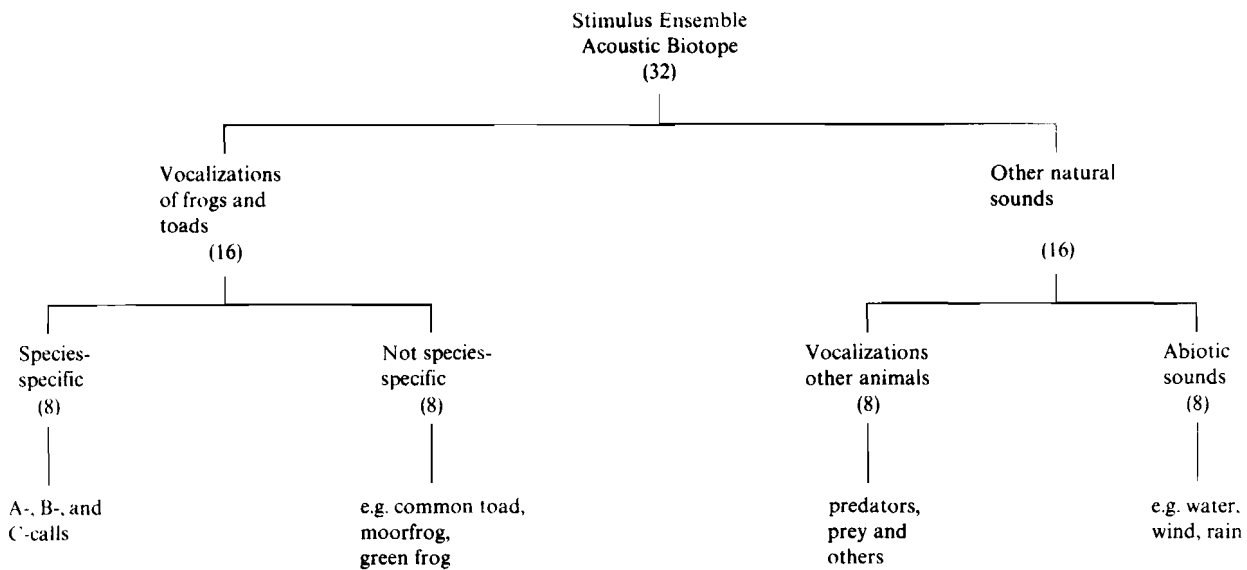


Diagram 1. Construction scheme for the realization of a stimulus ensemble inspired by the Acoustic Biotope of the grassfrog (*Rana temporaria* L.)

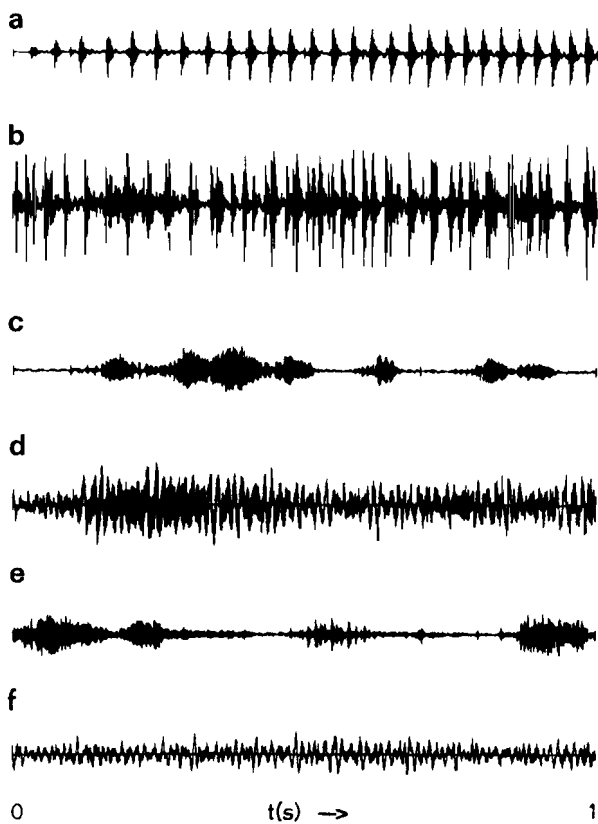


Fig. 3a-f. Different segments of sound, 1 s each, from the natural stimulus ensemble, the acoustic biotope, of the grassfrog (*Rana temporaria* L.). Vertical scale is identical in all cases. **a** Initial part of a C-call of the male grassfrog (total duration approx. 1.9 s). **b** Chorus of grassfrog vocalizations. **c** Vocalizations of the moorfrog (*Rana arvalis* Nilsson). **d** Vocalizations of gulls and chorus of green frogs (*Rana esculenta* L.) against noisy background (wind). **e** Vocalizations of heron and other birds. **f** Sound of a rippling brook

The construction scheme, shown in Diagram 1, takes into account several characteristics of the Acoustic Biotope, albeit in a rather rude way. Ethological and ecological considerations have led to the various categories as they appear in the scheme (Smolders et al., 1979). The probabilistic structure of the Acoustic Biotope is reflected in the numbers of sounds taken from each category. In the present context, however, special attention should be given to the physical structure of the signals in the stimulus ensemble, more notably their spectro-temporal intensity density.

Formally one might proceed along the lines as sketched in Sect. 2, resulting in a parametric description of natural sounds in terms of amplitude and frequency as functions of time. This approach, however, is most fruitful and at the same time agrees to our intuitive notions in the case of narrowband signals. In the case of a wideband signal the parametric approach lacks the intuitive appeal and results in parameters, the behaviour of which as functions of time comes close in complexity to the behaviour of the signal itself. A more general spectro-temporal description of signals is given by the *Dynamic Power Spectrum* (e.g. Flanagan, 1965) or *spectrogram* (e.g. Koenig et al., 1946; Potter et al., 1947). The dynamic power spectrum $P(f, t)$ of a signal gives the intensity as a function of time for the different frequency components of the signal. It can be measured by means of a set of bandpass filters and envelope detectors. Appendix 2 gives a description of the mathematical background and the principles of operation of an instrument, built for this purpose, the *Dynamic Spectrum Analyser* (DSA). In the present

investigation the filters used were 1/3 octave filters. As a consequence the dynamic spectrum is a set of functions of time representing the instantaneous intensity of the signal in a set of neighbouring 1/3 octave bands. Figure 4 may serve as an example: it shows the dynamic power spectrum of a *B*-call of the grassfrog as measured by the DSA (cf. Fig. 1).

The DSA also offers the possibility of triggered averaging of dynamic power spectra. The temporal relation between the trigger and the signal of interest can be varied between completely post- and completely pre-trigger position. This possibility of determining the average spectro-temporal characteristics of an ensemble of signals will be used in the investigation of the stimulus-event relation of auditory neurons (cf. the accompanying paper, Part II, Aertsen et al., 1980 and a forthcoming paper, Part III).

The parametric description of signals (cf. Sect. 2) and the description by means of the dynamic power spectrum can be formally connected by introduction of a general *Complex Spectro-Temporal Intensity Density* (CoSTID) of signals, defined on the analytic signal (Rihaczek, 1968; Johannesma and Aertsen, 1979).

4. Practical Realization of the Stimulus Ensembles

4.1. The Tonal Stimulus Ensemble

The tonal stimulus ensemble consisted of sequences of γ -tones with constant duration parameter β and form parameter γ [see (2.3)] and different values of frequency ω and amplitude factor A . The phase ϕ was either set to $-\pi/2$ or was not controlled. The frequency range was 4 octaves: either from 125–2000 Hz or from 250–4000 Hz, depending on the frequency characteristics of the neuron; values in this range were taken at 255 positions, equidistant on log frequency scale (64 intervals per octave). The amplitude factor resolution was 127 values at equal intervals ΔA (on linear scale) and minimum value ΔA , thus covering a range of about 42 dB, apart from the dynamic range within the individual γ -tones. The different values of ω and A were presented in a pseudo-random order to avoid possible systematic effects (e.g. hysteresis) of monotonous variation of these parameters (details of the pseudo-random sequences are described in the companion paper, Part II). The complete sequence consists of $255 \times 127 = 32,385$ γ -tones. The interval between onsets of successive γ -tones was 16 ms. Although a γ -tone as function of time theoretically stretches to plus infinity, the modulator $m(t)$, with β and γ as specified in (2.3), has decreased to less than 0.5% of its maximum value after 16 ms. This enables a truncation of the γ -tone after 16 ms without much

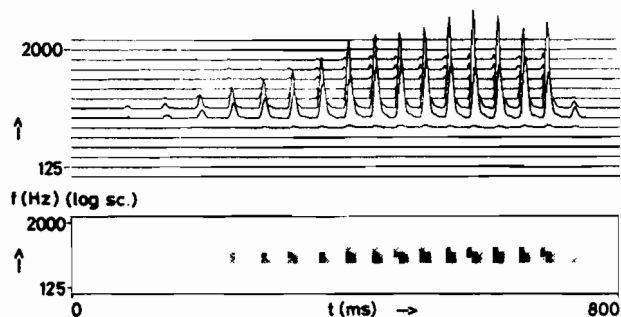


Fig. 4. Different representations of the Dynamic Power Spectrum of the "*B*-call" of the male grassfrog (cf. Fig. 1), determined by the Dynamic Spectrum Analyser (DSA), using a set of 1/3 octave filters

distortion. The result is a sequence of γ -tones with at any time only one tonal component present (cf. Fig. 2).

The tonal stimuli were generated by a Programmable Stimulus Generator (PSG), containing the following main components: the signal source was a programmable frequency synthesizer (Rockland 5100), the output of which was amplitude modulated by 2 serial 12-bits multiplying DA-converters (Hybrid Systems DAC 316-12); stimulus control was performed by a PDP 11/10. The generated waveform was lowpass filtered (2 cascaded -24 dB/oct Bessel filters, each -3 dB at 5 kHz) before it was amplified and presented to the animal.

4.2. The Natural Stimulus Ensemble

Most of the recordings of frog and toad vocalizations were made in the vicinity of Nijmegen and were obtained from the Department of Animal Ecology at the University of Nijmegen, the Netherlands. Details of the recording techniques were described by Van Gelder et al. (1978). The remainder of the natural sounds constituting the stimulus ensemble, was also recorded in the vicinity of Nijmegen (recorder Uher 4200, tape speed 19 cm/s, microphones Sennheiser MD 211 N). A stimulus tape containing a number of these sounds was composed according to the scheme in Diagram 1. This tape was played back to the animal during the experiment from an AM-recorder (Otari MX 5050 QXH, tape speed 38 cm/s). Before amplification and presentation to the grassfrog the signal was lowpass filtered according to the same specifications as the tonal stimulus ensemble.

4.3. Application

The stimulus ensembles as described here, together with the signal representations associated with them, have been used in an investigation into the spectro-temporal properties of the stimulus-event relation of single auditory neurons in the auditory midbrain area of the

grassfrog. The spectro-temporal intensity density of the particular stimuli which precede the occurrence of action potentials, the *Pre-Event Stimulus Ensemble* (PESE), is compared to the intensity density of the overall stimulus ensemble. The result of this comparison leads to the "spectro-temporal receptive field" of an auditory neuron. The results of this investigation, together with a discussion on the feasibility of the approach as given here, are described in separate papers. The companion paper (Part II, Aertsen et al., 1980) is concerned with the analysis of neural responses to the tonal stimulus ensemble. The response to the natural stimulus ensemble will be the subject of a forthcoming paper (Part III).

Appendix 1

Some Properties of the γ -Tones

A1.1. The Amplitude Modulation. The amplitude modulator of the γ -tone is given by [see (2.2)]

$$m(t) = \begin{cases} c \left(\frac{t}{\beta}\right)^{\gamma-1} \exp\left(-\frac{t}{\beta}\right) & t \geq 0 \quad \beta > 0 \\ 0 & t < 0 \quad \gamma \geq 1. \end{cases} \quad (\text{A1.1})$$

In this expression c is a *normalization* constant, to give $m^2(t)$ unit area:

$$c = \left(\frac{2^{2\gamma-1}}{\beta\Gamma(2\gamma-1)}\right)^{1/2}, \quad (\text{A1.2})$$

where $\Gamma(x)$ denotes the Gamma function (Abramowitz and Stegun, 1972).

The parameter β influences the *duration* of $m(t)$, it can be considered as a time scaling factor. The parameter γ controls the *form* of $m(t)$. The larger γ , the more symmetrical $m(t)$: for $\gamma = 1$ $m(t)$ describes pure exponential decay, for $\gamma \rightarrow \infty$ $m(t)$ approaches a Gaussian form. The temporal intensity $I(t) = m^2(t)$ is represented by the same function, with $\beta_I = \frac{\beta}{2}$ and $\gamma_I = 2\gamma - 1$.

The spectrum $\hat{m}(\omega)$ of $m(t)$ is

$$\hat{m}(\omega) = \frac{c\beta\Gamma(\gamma)}{(1+i\beta\omega)^\gamma}. \quad (\text{A1.3})$$

The spectral intensity density $J(\omega)$ is given by

$$J(\omega) = \hat{m}^*(\omega)\hat{m}(\omega) = \frac{c^2\beta^2\Gamma(\gamma)}{(1+\beta^2\omega^2)^\gamma}. \quad (\text{A1.4})$$

Several characteristics of the modulator $m(t)$ can now be calculated, e.g. the maximum value and the first and second order moment, both in the time and frequency domain (Gabor, 1946; Grashuis, 1974).

The *maximum value* is:

$$\text{in the time domain: } \begin{aligned} m_{\max} &= c(\gamma-1)^{\gamma-1} e^{-(\gamma-1)} \\ &\text{at } t = \beta(\gamma-1) \end{aligned} \quad (\text{A1.5})$$

$$\text{in the frequency domain: } \begin{aligned} |\hat{m}|_{\max} &= c\beta\Gamma(\gamma) \\ &\text{at } \omega = 0. \end{aligned}$$

The *mean value* μ_t of the "epoch" t of $m(t)$ is given by

$$\mu_t = \frac{\int_{-\infty}^{\infty} dt m(t) t m(t)}{\int_{-\infty}^{\infty} dt m(t) m(t)} = \frac{\beta}{2} (2\gamma - 1). \quad (\text{A1.6})$$

The *standard deviation* σ_t , which describes the *temporal width* or *duration* of $m(t)$, is given by

$$\sigma_t = \left(\frac{\int_{-\infty}^{\infty} dt m(t) (t - \mu_t)^2 m(t)}{\int_{-\infty}^{\infty} dt m(t) m(t)} \right)^{1/2} = \frac{\beta}{2} (2\gamma - 1)^{1/2}. \quad (\text{A1.7})$$

Similar definitions can be given in the frequency domain. The *mean value* μ_ω of the angular frequency is given by

$$\mu_\omega = \frac{\int_{-\infty}^{\infty} d\omega \hat{m}^*(\omega) \omega \hat{m}(\omega)}{\int_{-\infty}^{\infty} d\omega \hat{m}^*(\omega) \hat{m}(\omega)} = 0. \quad (\text{A1.8})$$

The *standard deviation* σ_ω , describing the *spectral width* is given by

$$\sigma_\omega = \left(\frac{\int_{-\infty}^{\infty} d\omega \hat{m}^*(\omega) (\omega - \mu_\omega)^2 \hat{m}(\omega)}{\int_{-\infty}^{\infty} d\omega \hat{m}^*(\omega) \hat{m}(\omega)} \right)^{1/2} = \frac{(2\gamma - 3)^{-1/2}}{\beta} \quad (\gamma > \frac{3}{2}) \quad (\text{A1.9})$$

or, equivalently,

$$\sigma_f = \frac{\sigma_\omega}{2\pi} = \frac{(2\gamma - 3)^{-1/2}}{2\pi\beta}. \quad (\text{A1.10})$$

Combining the measures for width in time and frequency domain we obtain the *spectro-temporal resolution* of the signal or its "uncertainty product" (Gabor, 1946)

$$\Delta = \sigma_\omega \sigma_t = \frac{1}{2} \left(\frac{2\gamma - 1}{2\gamma - 3} \right)^{1/2} \quad (\gamma > \frac{3}{2}) \quad (\text{A1.11})$$

which, for a Gaussian waveform ($\gamma \rightarrow \infty$), reaches a minimum of $\frac{1}{2}$. The parameters of the γ -modulator

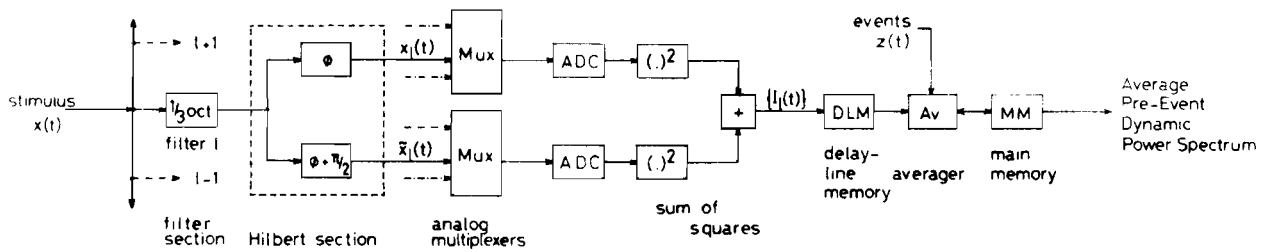


Fig. 5. Functional diagram of Dynamic Spectrum Analyser (DSA). The various control functions have been omitted for reasons of clarity

used in the present investigation are $\beta = 1.45$ ms and $\gamma = 3$ [see (2.3)]. This gives numerical values for the *maximum*:

$$\begin{aligned} &\text{in the time domain at } t = 2.90 \text{ ms} \\ &\text{in the frequency domain at } \omega = 0 \text{ rad/s} \end{aligned} \quad (\text{A1.12})$$

for the *moments*:

$$\begin{aligned} &\text{-- in the time domain: } \mu_t \approx 3.63 \text{ ms,} \\ &\quad \sigma_t \approx 1.62 \text{ ms} \\ &\text{-- in the frequency domain: } \mu_\omega = 0 \text{ rad/s,} \\ &\quad \sigma_\omega \approx 398.2 \text{ rad/s} \\ \text{or:} &\quad \mu_f = 0 \text{ Hz,} \\ &\quad \sigma_f \approx 63.4 \text{ Hz} \end{aligned} \quad (\text{A1.13})$$

and for the *uncertainty product*:

$$\Delta \approx 0.645, \quad (\text{A1.14})$$

which is only slightly above the theoretical limit of $\frac{1}{2}$.

A1.2. The γ -tone. The γ -tone $e(t)$ is defined by [see (2.1)]

$$e(t) = Am(t) \cos(\omega_\gamma t + \phi) \quad (\text{A1.15})$$

with the spectrum

$$\hat{e}(\omega) = \frac{1}{2} A (\hat{m}(\omega - \omega_\gamma) + \hat{m}(\omega + \omega_\gamma)). \quad (\text{A1.16})$$

The Hilbert transform of $e(t)$ is

$$e(t) = Am(t) \sin(\omega_\gamma t + \phi) \quad (\text{A1.17})$$

provided that the highest frequency component in $m(t)$ is below ω_γ [see (2.20)]. This condition is not fulfilled by the spectrum $\hat{m}(\omega)$, however it is useful to regard (A1.17) as an approximation with an error determined by the spectral tail of $m(t)$ (Rubin and Di Franco, 1963). The integral

$$F(\nu) = \frac{\int_{-\nu}^{\nu} d\omega \hat{m}^*(\omega) \hat{m}(\omega)}{\int_{-\infty}^{\infty} d\omega \hat{m}^*(\omega) \hat{m}(\omega)} \quad (\text{A1.18})$$

gives the *fraction* of spectral intensity in the interval $(-\nu, \nu)$. Substitution of (A1.3), using $\gamma = 3$, leads to (Gradshteyn and Ryzhik, 1965)

$$F(\nu) = \frac{1}{2\pi} c^2 \beta \Gamma^2(\gamma) \left(\frac{u}{4(1+u^2)^2} + \frac{3u}{8(1+u^2)} + \frac{3}{8} \text{arctg} u \right) \Big|_{u=-\beta\nu}^{u=\beta\nu}. \quad (\text{A1.19})$$

Numerical evaluation of $F(\nu)$ gives

$$\begin{aligned} F(\sigma_\omega) &= 0.75 \\ F(2\sigma_\omega) &= 0.95 \\ F(3\sigma_\omega) &= 0.99. \end{aligned} \quad (\text{A1.20})$$

From the results (A1.13) and (A1.20) we conclude that for the γ -tones as used in the present investigation, with f_γ in the range of 125 Hz to 4000 Hz, the analytic signal representation is, to a fair approximation, given by

$$\xi(t) \approx Am(t) \exp i(\omega_\gamma t + \phi) \quad (\text{A1.21})$$

which gives, for the instantaneous amplitude or temporal envelope

$$a(t) \approx Am(t) \quad (\text{A1.22})$$

and for the instantaneous frequency

$$\omega(t) \approx \omega_\gamma. \quad (\text{A1.23})$$

Obviously the approximations (A1.21–23) are better for higher carrier frequency ω_γ .

Appendix 2

Dynamic Spectrum Analyser (DSA)

A2.1. Spectro-Temporal Analysis. A spectro-temporal analysis of sound can be made by measuring the *Dynamic Power Spectrum* $P(f, t)$, also referred to as *short-time power spectrum* (Fano, 1950; Schroeder and Atal, 1962; Flanagan, 1965; Gambardella, 1971). The dynamic power spectral analysis has been implemented in hardware, thereby operating in real time, on the principle of simultaneous, multiple bandpass filtering and envelope detection (Hoeks and Krijt, 1976). Figure 5 shows a schematic diagram of the *Dynamic Spectrum Analyser* (DSA).

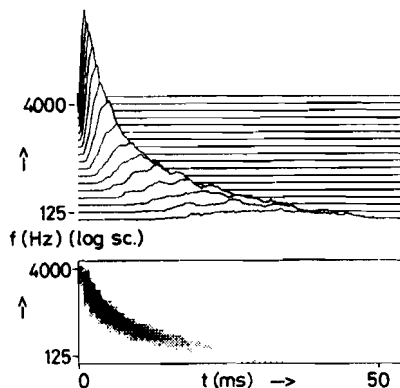


Fig. 6. Dynamic Power Spectrum of an impulse, as determined by the DSA, using a set of 1/3 octave filters. Note the frequency dependent time shift in the spectrogram

The frequency analysis is performed by a set of 1/3 octave filters (HP8065A), each one followed by a Hilbert transform section, producing two signals with a phase difference of $\pi/2$ rad. These two signals correspond to the real and imaginary part of the analytic bandfiltered signal [cf. (2.4)]. Summation of the squares of both signals gives the temporal intensity of the bandfiltered signal [cf. (2.17)]. The result of this section is $P(f, t)$ at $f = \mu_{f,l}$ ($l = 1, L$), where $\mu_{f,l}$ is the center frequency of the l -th bandpass filter. An example is shown in Fig. 4: the dynamic power spectrum of a "B-call" of the grassfrog.

The filter section consists of 44 1/3 octave filters between 2 Hz and 50 kHz, with center frequencies equidistant on log scale. From this range a frequency interval of maximally 8 octaves can be selected. The overall behaviour of each filter is that of a 6-th order Chebychev filter with a ripple in the passband region of less than 0.5 dB. For each of the filters the effective bandwidth σ_f [cf. (A1.10)] equals about $0.082 \mu_f$, the effective duration σ_t of the impulse response [cf. (A1.7)] equals about $1.52/\mu_f$, as a consequence the uncertainty product $\Delta = 2\pi\sigma_f\sigma_t$ [cf. (A1.11)] for each of the filters is 0.78. The attenuation outside the passband region is 20 dB at $0.79 \mu_f$ and $1.26 \mu_f$, 50 dB at $\mu_f/2$ and $2\mu_f$ and amounts to 70 dB at $\mu_f/4$ and $4\mu_f$. Due to the phase characteristics of the bandpass filters and the Hilbert transformers the filtered signals have a delay, depending on the filter center frequency μ_f . This effect is demonstrated by the dynamic power spectrum of an impulse, shown in Fig. 6. The delay caused by the bandpass filter is the group delay μ_t of the filter impulse response [cf. (A1.6)], where for each of the filters μ_t is inversely proportional to the center frequency: $\mu_t\mu_f \approx 3.45$. The frequency dependent delay of the Hilbert transformers added to μ_t gives a total delay μ'_t for which it holds $\mu'_t\mu_f \approx 3.8$. Obviously this delay

should be taken into account in the interpretation of results, especially the estimation of latencies etc.

The differing absolute bandwidth of the 1/3 octave filters causes the spectral intensity density to be transformed: a $1/f$ -spectrum turns "white" and a white spectrum would look like an " f -spectrum". An intensity correction option, consisting of filter-frequency dependent amplifiers combined with the Hilbert transformers, is available. Further spectral manipulation is possible by means of the variable gain of each bandfilter, covering a range of 40 dB.

Both the set of "real" and the set of "imaginary" filtered signals from the Hilbert transformers are multiplexed, using 2 simultaneously operating analog time multiplexers. Both multiplexers are followed by 9-bits AD-converters (maximum rate 100 ksamples/s). The differential time jitter of sampling moments is less than 10 ns. Multiplexing and sampling can be performed in 2 modes: 1. *sequential* sampling in which the center frequency of the highest filter selected determines the sampling rate (Shannon), which is then used to sample all filter outputs and 2. *proportional* sampling in which the sampling rate for the 3 filters in each separate octave is determined by the upper filter frequency from that octave [compare the "logons" introduced by Gabor (1946)]. In the latter sampling mode the total number of samples is strongly reduced. The outputs from the Hilbert sections are AC-coupled before the multiplexer to eliminate possible DC-offsets, causing undesired cross products in the squarers. Up to this level of analysis the dynamic range is 54 dB.

The operations of squaring and summation are performed digitally, using 18 bits resolution, and result in 18-bits digitized temporal intensities of the band-filtered signals.

A2.2. Averaging of Dynamic Power Spectra. In the DSA the dynamic power spectra of an ensemble of signals can be averaged. In the present investigation the ensemble of interest is the ensemble of signals *prior* to the neural events (the PESE, see the companion paper, Part II).

In order to have access to the dynamic power spectrum prior to an event, the data from the "sum of squares"-units are continuously stored into a delay line memory (4096 18-bits words). The maximum temporal interval to be contained in this memory is determined by the selection of filters and the sampling mode. If necessary this interval can be reduced by powers of 2. Upon a trigger the contents of the delay-line memory is retrieved and averaged with the contents of the main memory (4096 18-bits words). Data transfer time is 7.0–7.5 μ s/word. The retrieval of data from the delay line memory is performed asynchronously with the storage of data. Triggering can be done in internal,

manual and external mode, the latter both on negative and positive slope of the event. The temporal relation between the trigger and the dynamic power spectrum to be transferred is set by the temporal trigger mode. The 5 possibilities are: 0%, 25%, 50%, 75%, and 100% pre-event dynamic spectrum, combined with the complementary fraction of the post-event dynamic spectrum. During transfer of data to main memory triggering is inhibited. The resulting "dead time" equals the maximum of 1. the post-trigger time interval and 2. the time taken to transfer a complete dynamic power spectrum to the main memory (about 29.5 ms for full memory). This "dead time" may result in a missing of triggers and a consecutive departure from the true ensemble average to an amount determined by the rate of (neural) events and the total number of samples per dynamic spectrum.

Averaging is performed on a running base: the weight factor of individual spectra is continuously adjusted to the number of averages already executed. Two running average algorithms are implemented. In the present investigation we have used throughout

$$q_n = \frac{q_{n-1}}{2^{k_n - k_{n-1}}} + \frac{p_n}{2^{k_n}}, \quad (\text{A2.1})$$

where q_n , q_{n-1} = running average of n resp. $n-1$ dynamic power spectra; p_n = dynamic power spectrum of n -th ensemble element; k_n = integer number, with $2^{k_{n-1}} < n \leq 2^{k_n}$.

In this algorithm q_n is a true average, however, it has correct scaling only for $n = 2^k$, $k = 0, 1, 2, \dots$. The other algorithm

$$q'_n = q'_{n-1} + \frac{p_n - p_{n-1}}{2^{k_n}} \quad (\text{A2.2})$$

has correct amplitude scaling but does not provide a true average. Averaged data from the main memory can be displayed on an oscilloscope: intensity as function of time and filter frequency using a quasi-3D display and/or intensity modulation. Read-out of the main memory is also possible to XY-recorder and computer. In the latter case a restricted portion of the main memory, containing the various selected control functions and preset values of the DSA, is transferred as well.

Acknowledgements. This investigation forms part of the research program of the Workgroup Neurophysics at the Department of Medical Physics and Biophysics, University of Nijmegen (The Netherlands) and was supported by the Netherlands Organization for Advancement of Pure Research (Z.W.O.). The authors wish to thank Jos Eggermont and Ton Vendrik for helpful discussions on this manuscript. Pieter Elst, Arnold Hoeks, Hans Krijt, and Jan Lammers made essential contributions to the development of the Dynamic Spectrum Analyser and actually realized it. Tape recordings of vocalizations of various species of anurans were kindly provided by Jan van Gelder (Department of Animal Ecology, University of Nijmegen).

References

- Abramowitz, M., Stegun, I.A. (eds.) Handbook of mathematical functions. New York: Dover 1972
- Aertsen, A.M.H.J., Smolders, J.W.T., Johannesma, P.I.M.: Neural representation of the Acoustic Biotope. on the existence of stimulus-event relations for sensory neurons. *Biol. Cybernetics* **32**, 175-185 (1979)
- 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. Cybernetics* **38**, 235-248 (1980)
- Bedrosian, E.: A product theorem for Hilbert transforms. *Proc. IEEE* **51**, 868-869 (1963)
- Boer, E.de: On the principle of specific coding. *J. Dyn. Syst. Meas. Control* **95G**, 265-273 (1973)
- Boer, E.de, Jongh, H.R.de: On cochlear encoding: potentialities and limitations of the reverse correlation technique. *J. Acoust. Soc. Am.* **63**, 115-135 (1978)
- Brzoska, J., Walkowiak, W., Schneider, H.: Acoustic communication in the grassfrog (*Rana temporaria* L.): calls, auditory thresholds, and behavioral responses. *J. Comp. Physiol.* **118**, 173-186 (1977)
- Bullock, T.H.(ed.): Recognition of complex acoustic signals (Life Sciences Res. Rep., Vol. 5). Berlin, Heidelberg, New York: Springer 1977
- Cain, G.D.: Hilbert transform relations for products. *Proc. IEEE* **61**, 663-664 (1973)
- Capranica, R.R.: The evoked vocal response of the bullfrog: a study of communication by sound. Cambridge, MA: MIT Press 1965
- Cox, D.R., Lewis, P.A.W.: The statistical analysis of series of events. London: Methuen 1966
- Dahl, F.: Grundlagen einer ökologischen Tiergeographie. Teil I. Jena: Gustav Fischer 1921
- Dugundji, J.: Envelopes and pre-envelopes of real waveforms. *IRE Trans. Inf. Theory* **IT-4**, 53-57 (1958)
- Fano, R.M.: Short-time autocorrelation functions and power spectra. *J. Acoust. Soc. Am.* **22**, 546-550 (1950)
- Flanagan, J.L.: Speech analysis, synthesis, and perception. Berlin, Heidelberg, New York: Springer 1965
- Gabor, D.: Theory of communication. *J. IEE (London) Part III*, **93**, 429-457 (1946)
- Gambardella, G.: A contribution to the theory of short-time spectral analysis with nonuniform bandwidth filters. *IEEE Trans. Circuit Theory CT-18*, 455-460 (1971)
- Gelder, J.J. van, Evers, P.M.G., Maagnus, G.J.M.: Calling and associated behaviour of the common frog, *Rana temporaria*, during breeding activity. *J. Anim. Ecol.* **47**, 667-676 (1978)
- Gerhardt, H.C.: The significance of some spectral features in mating call recognition in the green treefrog (*Hyla cinerea*). *J. Exp. Biol.* **61**, 229-241 (1974)
- Gerhardt, H.C.: Mating call recognition in the green treefrog (*Hyla cinerea*): the significance of some fine-temporal properties. *J. Exp. Biol.* **74**, 59-73 (1978)
- Gradshteyn, I.S., Ryzhik, I.M.: Table of integrals, series, and products. New York, London: Academic Press 1965
- Grashuis, J.L.: The pre-event stimulus ensemble: an analysis of the stimulus-response relation for complex stimuli applied to auditory neurons. Ph. D. Thesis, Nijmegen 1974
- Gupta, M.S.: Definition of instantaneous frequency and frequency measurability. *Am. J. Phys.* **43**, 1087-1088 (1975)
- Hoeks, A.P.G., Krijt, H.J.: Hardware processing of non-stationary signals. Internal report, Dept. of Medical Physics and Biophysics, University of Nijmegen 1976
- 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 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., Schleich, H., Schreiner, Chr. (eds.). *Exp. Brain Res.* [Suppl. II] (1979)
- Kiang, N.Y.-S., Watanabe, T., Thomas, E.C., Clark, L.F.: Discharge patterns of single fibers in the cat's auditory nerve. Cambridge, MA: MIT Press 1965
- Koenig, W., Dunn, H.K., Lacy, L.Y.: The sound spectrograph. *J. Acoust. Soc. Am.* **18**, 19–49 (1946)
- Lerner, R.M.: A matched filter detection system for complicated Doppler shifted signals. *IRE Trans. Inf. Theory* **IT-6**, 373–385 (1960)
- Mandel, L.: Interpretation of instantaneous frequencies. *Am. J. Phys.* **42**, 840–846 (1974)
- Middleton, D.: An introduction to statistical communication theory. New York, Toronto, London: McGraw-Hill 1960
- Potter, R.K., Kopp, G.A., Green, H.C.: Visible speech. New York: Van Nostrand 1947
- Rihaczek, A.W.: Hilbert transforms and the complex representation of real signals. *Proc. IEEE* **54**, 434–435 (1966)
- Rihaczek, A.W.: Signal energy distribution in time and frequency. *IEEE Trans. Inf. Theory* **IT-14**, 369–374 (1968)
- Rubin, W.L., DiFranco, J.V.: Analytic representation of wideband radio frequency signals. *J. Franklin Inst.* **275**, 197–204 (1963)
- Schneider, H.: Die Paarungsrufe einheimischer Ranidae (Anura, Amphibia). *Bonn. Zool. Beitr.* **24**, 51–61 (1973)
- Schroeder, M.R., Atal, B.S.: Generalized short-time power spectra and auto-correlation functions. *J. Acoust. Soc. Am.* **34**, 1679–1683 (1962)
- Smolders, J.W.T., Aertsen, A.M.H.J., Johannesma, P.I.M.: Neural representation of the Acoustic Biotope: a comparison of the response of auditory neurons to tonal and natural stimuli in the cat. *Biol. Cybernetics* **35**, 11–20 (1979)
- Thitchmarsh, E.C.: Introduction to the theory of Fourier integrals. London: Oxford University Press 1937
- Tricomi, F.G.: On the finite Hilbert transformation. *Q. J. Math. Oxford* **2**, 199–211 (1951)
- 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: May 11, 1980

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