

Quantitative characterisation procedure for auditory neurons based on the spectro-temporal receptive field

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

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

(Received 9 March 1982; accepted 1 December 1982)

Studies dealing with auditory information processing often present the dynamic spectrum of the sound stimulus (sonogram) in addition to the stimulus waveform. The sonogram, presenting the spectral and temporal properties of the sound in a combined way, reflects properties that are assumed relevant in central information processing.

For 12 neurons recorded from the midbrain of the grass frog the sonogram of a Gaussian wide-band noise stimulus was correlated with the output of the neuron to that noise. From this input–output correlogram the spectro-temporal receptive field (STRF) was calculated. The STRF reflects those spectral and temporal properties of the stimulus that influence the firing probability of the neuron.

A quantitative procedure was developed to calculate the neuron's response as far as it could be derived from the STRF. This procedure basically consisted of a convolution between STRF and the sonogram of the stimulus followed by a summation over the various frequency bands. In this way it proved possible to estimate to what extent the STRF characterised the neuron's firing behaviour.

Heuristic approaches, in which the neuron was modelled to a parallel series of band-pass filters, a summator and a static nonlinearity, representing a spike-generating mechanism, resulted in a considerable improvement of the characterisation.

Key words: auditory information processing; cross-correlation; sonogram; spectro-temporal receptive field; frog; torus semicircularis.

Introduction

The importance of complex sound signals for the investigation of the auditory central nervous system has been emphasised by several authors [1,19,21,23,26]. These complex sounds are preferably chosen from the natural environment of the animal, the acoustic biotope [1], in which species-specific vocalisations play a substantial role. When synthetic complex sounds are used, they are often inspired by the frequency–time patterns found in the sonograms of these species-specific vocalisations. It is thereby implicitly assumed that the sonogram, a spectro-temporal representation of sound, presents those frequency–time patterns that are relevant for the animal. Thus, it was felt necessary to investigate the spectral and temporal characteristics of auditory neurons in the central nervous system in a combined way.

By that we mean that instead of measuring separately the spectral characteristics (e.g. the tuning curve) and temporal characteristics (e.g. the PSTH) of the neuron, we obtain simultaneously these characteristics in one experiment. This procedure can be based on stimulation with Gaussian wide-band noise [9] or on tonal stimuli with random frequency and amplitude [10] and has been applied to neurons in the auditory midbrain of the lightly anaesthetised grass frog. One-third of the neural units therefrom responded in a sustained way to pseudo-random wide-band noise. Noise is an example of an unstructured complex sound and as such seems ideally suited to investigate unknown systems [15,20]. In the present paper we will investigate how well a neuron can be characterised by determination of its spectro-temporal response characteristics to noise. This will be done by calculating the probability of response on basis of this spectro-temporal characterisation and comparing it to the experimentally determined PSTH.

As in previous studies [9,10] we will use the concepts of stimulus ensemble (SE) and pre-event stimulus ensemble (PESE) as defined in [1]. The series of action potentials generated by the neuron is, for example, by using a Schmitt-trigger, converted into a series of events. It is assumed that only the times of occurrence of action potentials carry information. While the complete set of stimuli presented to the neuron constitutes the SE, the subset that consists of stimulus elements that precede an event forms the PESE. In this particular approach to the study of the response characteristics of neurons, functionals of the stimulus are averaged over the PESE. By this we mean that for the time segments preceding neural events an average functional is calculated; a functional is defined as a function of a function. The noise signal itself is a function of time $x(t)$; the noise intensity, for instance, is in turn a function of this $x(t)$ and therefore a functional. In the present paper we consider the dynamic spectrum or sonogram as the functional of interest. The differences between the average dynamic spectrum of the PESE and the average dynamic spectrum of the SE are related to those spectro-temporal characteristics of the stimulus that influence the firing probability of the neuron. When the stimulus ensemble is Gaussian wide-band noise this difference is called the spectro-temporal receptive field (STRF). In this paper is outlined a procedure to calculate the share in the response of the neural unit that can be derived from its STRF. It is therefore an investigation into how far the STRF characterises the neural unit.

The procedure to calculate that part of the PSTH of a neural unit that is accounted for by its STRF basically consists of a convolution between the STRF and the sonogram of the sound followed by a summation over the various frequency bands. This can also be seen as moving a template representing the neuron's spectro-temporal properties over the sonogram of the sound and taking the degree of correspondence as a measure for the firing probability of the neuron. The similarity between the calculated STRF-based response probability and the measured PSTH reflects the amount to which the unit is characterised by its STRF. This quantitative procedure will only result in a complete characterisation if there exists a linear relation between the sonogram of the sound stimulus and the response of the neural unit. Threshold mechanisms and other non-linearities in the transduction between sound and neural response as well as spontaneous activity of the neural unit will

cause the characterisation to be incomplete. In this study it will be made clear that a model-based approach incorporating such nonlinearities improves the characterisation procedure.

Methods

Experimental data

Adult grass frogs from Ireland were lightly anaesthetised (MS 222) and single-unit activity was recorded from auditory neurons in the midbrain. All details about preparation, stimulation, recording of single-unit activity and data acquisition were described in a previous paper [9]. In that paper a population of units was described that responded in a sustained way to pseudorandom wide-band noise. The quantitative characterisation procedure described hereafter was applied to twelve of these neurons from which sufficient data were available.

The stimulus

The spectro-temporal characteristics of a neuron were determined under stimulation with pseudorandom Gaussian wide-band noise. This noise was generated by low-pass filtering (6 dB/octave) of a binary sequence with a length of 1 048 575 steps [8]. The cut-off frequency (-3 dB) of the filter was either 1500 or 5000 Hz. A complete noise stimulus consisted of several sequences presented immediately after each other.

The dynamic spectrum

The dynamic spectrum or sonogram represents the spectral content of a signal as function of time. The spectral and temporal characteristics are presented in a combined way. A possible way to construct the dynamic spectrum is to pass the signal through a bank of bandpass filters and to measure the temporal intensities, i.e. the square of the envelopes, of the outputs of these filters. In this study the dynamic spectrum was measured by means of a real-time dynamic spectrum analyser (DSA) [2]. Eighteen third-octave bandpass filters were used, having central frequencies equidistant on a logarithmic scale between 100 and 5000 Hz. The intensity of the output of each filter was sampled with intervals of 1.92 ms.

Theory

The theoretical basis for the characterisation of the neural response to an auditory stimulus consists of two assumptions. First, the effective stimulus for auditory neurons in the midbrain of the grass frog is given by the dynamic spectrum $\{P_k(t)\}$ of the sound. Second, the response of the neuron is described by the PSTH, $p(t)$.

We will derive a relation between $\{P_k(t)\}$ and $p(t)$ for a given stimulus ensemble and use this to predict responses to parts of that stimulus ensemble or different stimulus ensembles. Use will be made of results from linear and nonlinear system theory [15,20].

The dynamic spectrum or spectro-temporal intensity of the acoustic stimulus is given by $\{P_k(t)\}$ where the index k indicates central frequency f_k . $P_k(t)$ is therefore the power of the stimulus at time t around frequency f_k . The spectro-temporal intensity can be written as:

$$P_k(t) = P_{k_0} + I_k(t) \quad (1)$$

where P_{k_0} is the average spectro-temporal intensity of the SE, and I_k is the deviation of the spectro-temporal intensity from the average level.

The PSTH can be written in the form:

$$p(t) = p_0 + p_1(t) + \epsilon(t) \quad (2)$$

where p_0 is the average firing rate of the neuron, $p_1(t)$ is the component of $p(t)$ that is linearly related to the STRF, and $\epsilon(t)$ is the remainder caused by nonlinear components of the neural sensitivity.

Linear characterisation of the neural sensitivity

The linear component $p_1(t)$ of the neural response $p(t)$ is that part which depends in a linear way on the variation of the spectro-temporal intensity of the stimulus. As a consequence it can be written as:

$$p_1(t) = \sum_{l=1}^K \int_{-\infty}^{\infty} h_l(\sigma) I_l(t - \sigma) d\sigma \quad (3)$$

which states that the linear part of the response is found by a linear weighting of the difference in spectro-temporal intensity of the PESE and that of the SE. In other words, the intensity of the signal around a given frequency f_l is convolved with the sensitivity of the neuron in that frequency region, $h_l(\sigma)$, and the contributions of the different frequency regions are then summed to give the linear contribution of the PSTH. Eqn. 3 may be considered as the definition of $p_1(t)$ on the condition that the STRF $\{h_l(\sigma)\}$ is known.

$h_k(\sigma)$ can be determined on the basis of stimulus response correlation. We define the autocorrelation function of the spectral intensity at frequency f_k by

$$R_{kk}(\tau) = \int_{-\infty}^{\infty} I_k(t - \tau) I_k(t) dt, \quad (4)$$

the cross-correlation of spectral intensities at differing frequencies f_k and f_l by

$$R_{kl}(\tau) = \int_{-\infty}^{\infty} I_k(t - \tau) I_l(t) dt, \quad (5)$$

the cross-correlation of $I_k(t)$ and $p_1(t)$ by

$$R_{kp_1}(\tau) = \int_{-\infty}^{\infty} I_k(t - \tau) p_1(t) dt, \quad (6)$$

and the cross-correlation of $I_k(t)$ and the response $p(t)$ by

$$R_{kp}(\tau) = \int_{-\infty}^{\infty} I_k(t - \tau) p(t) dt \quad (7)$$

These correlation functions form the basis for the determination of the STRF

$\{h_k(\sigma)\}$. Multiplication of both sides of Eqn. 3 with $I_k(t - \tau)$, integration over t and interchange of summation and integration leads to

$$R_{k p_1}(\tau) = \sum_{l=1}^K \int_{-\infty}^{\infty} h_l(\sigma) R_{kl}(\tau - \sigma) d\sigma \quad (8)$$

This equation relates the two correlation functions $R_{k p_1}(\tau)$ and $R_{kl}(\tau)$. A considerable simplification of Eqn. 8 is possible if the temporal intensities in different frequency bands are uncorrelated: $R_{kl}(\tau) = 0$ for $k \neq l$.

In Appendix 1 it is shown that for Gaussian white noise as acoustic stimulus the intensities in non-overlapping different 1/3-octave bands as used in the dynamic spectrum analyser are uncorrelated.

Although two neighbouring bandpass filters showed overlapping amplitude characteristics from their -3 dB points on, actual calculation of the cross covariance function of the temporal intensities of two neighbouring frequency bands showed that they were uncorrelated within the error resulting from the finite lengths of the eight records used and the systematic error due to the use of pseudorandom noise. Therefore, Eqn. 8 reduces to

$$R_{k p_1}(\tau) = \int_{-\infty}^{\infty} h_k(\sigma) R_{kk}(\tau - \sigma) d\sigma \quad (9)$$

which means that $R_{k p_1}(\tau)$ is obtained by a convolution of the unknown $h_k(\sigma)$ with $R_{kk}(\tau)$. It follows therefore that the STRF $\{h_k(\sigma)\}$ can be computed by a deconvolution of $R_{k p_1}(\tau)$ with $R_{kk}(\tau)$. This, however, implies knowledge of the linear part $p_1(t)$ of the PSTH.

This type of problem can be solved in a simple way if we impose another condition on the decomposition in Eqn. 2: the three components p_0 , $p_1(t)$ and $\epsilon(t)$ should be uncorrelated with respect to the stimulus ensemble used (compare the Wiener-type approach [15]). This effectively implies that a truncated version of Eqn. 2 can be used to calculate the system's response to the stimulus ensemble, and if possible it can be extended to include higher order terms without affecting the contributions already described by the truncated model.

In Appendix 2 it is shown for a stimulus ensemble with uncorrelated intensities in different frequency bands (cf. Eqn. A1.1) that this condition leads to the relation

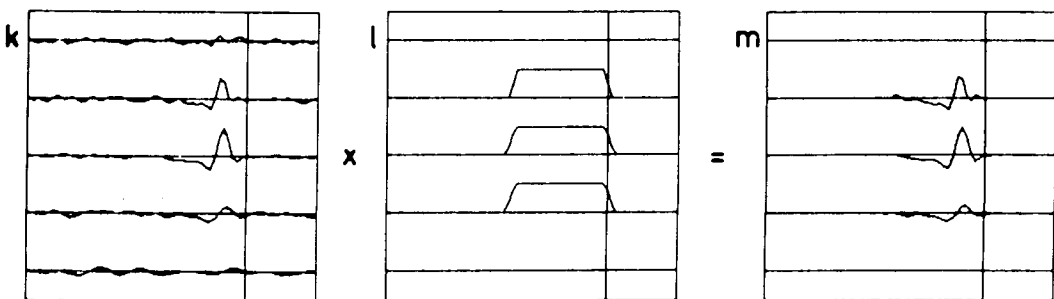
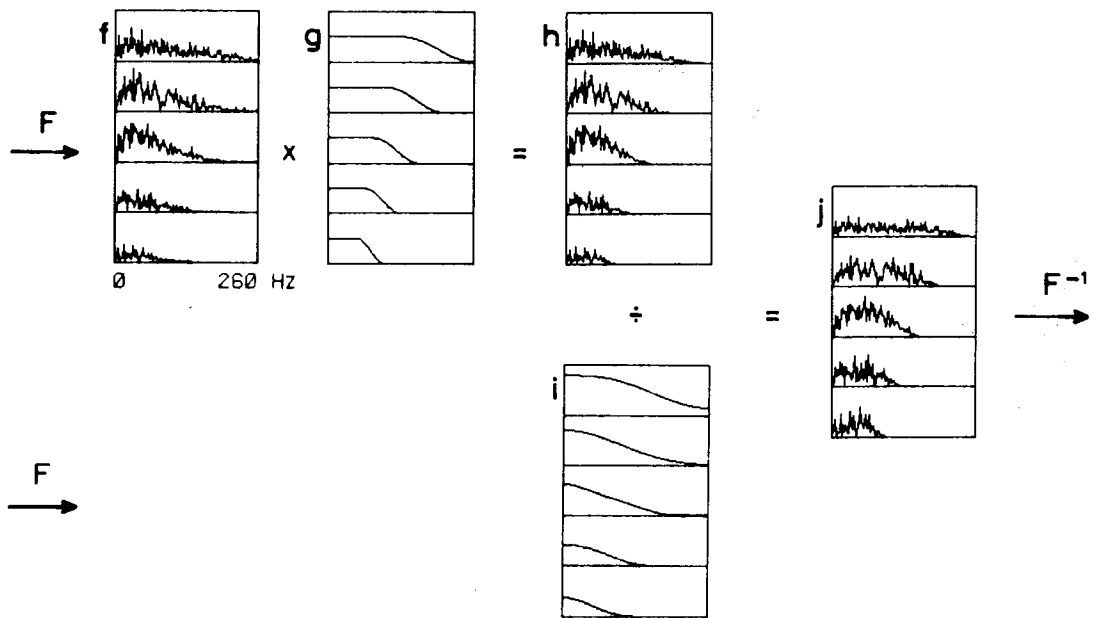
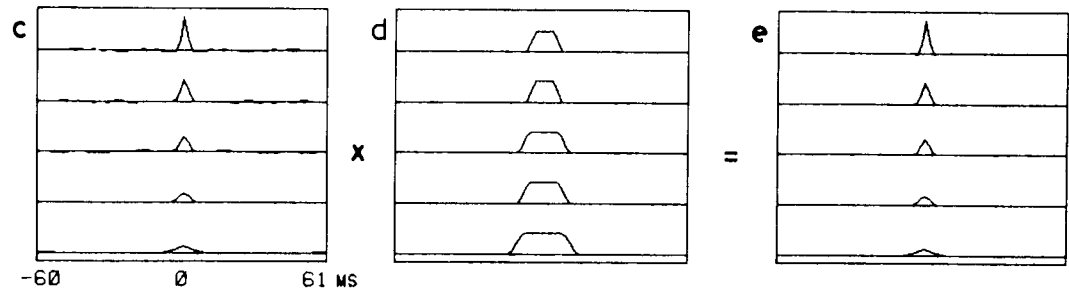
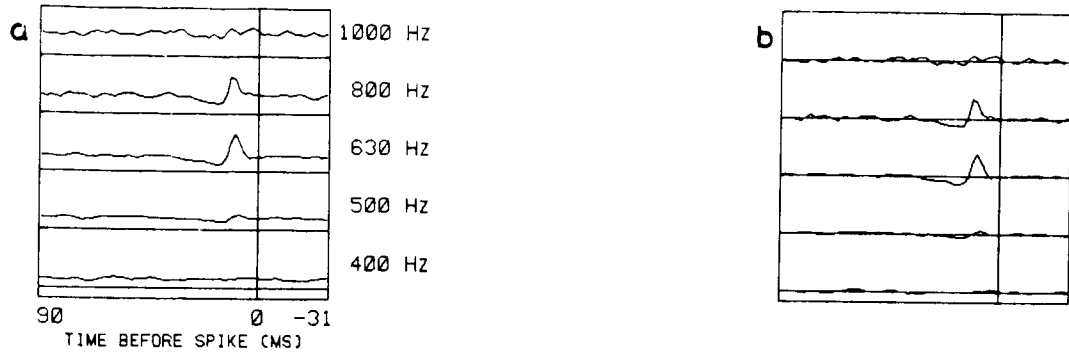
$$R_{kp}(\tau) = R_{k p_1}(\tau) \quad (10)$$

Combining Eqns. 9 and 10 therefore leads to

$$R_{kp}(\tau) = \int_{-\infty}^{\infty} h_k(\sigma) R_{kk}(\tau - \sigma) d\sigma \quad (11)$$

and $h_k(\sigma)$ can be computed from the experimentally measurable $R_{kp}(\tau)$ and $R_{kk}(\tau)$. This procedure is carried out in the frequency domain after Fourier transformation (see Fig. 1).

It should be realised that even when the auditory system does not adapt its properties to the characteristics of the stimulus ensemble the $\{h_k(\sigma)\}$ depend on the choice of the stimulus ensemble. The STRF represents the best linear relation



between spectro-temporal intensity of the noise stimulus and the neural response. In reality the relation between the spectro-temporal sensitivity (STS) and the neural response will be nonlinear and the STS may differ from stimulus ensemble to stimulus ensemble. This paper and a companion paper will consider two important points regarding the STRF. The first concerns the quality of the STRF, and will be investigated in this paper: how well does the STRF characterise the PSTH? The second point concerns the validity of the STRF determined with a noise stimulus to predict the response to a different stimulus ensemble, i.e. species-specific vocalisations, and is investigated in the companion paper.

Calculation of the STRF-based response

In order to verify the quality of the STRF, $\{h_k(\sigma)\}$ derived from Eqn. 11 we will use this STRF for the calculation of the response of the neuron to selected parts of the SE. The assumptions made in the Theory section limit the applicability to a stimulus ensemble for which the intensities in different spectral regions are uncorrelated. The test stimulus both for determining the STRF and for the response calculation is Gaussian white noise.

By calculating $p_1(t)$ with Eqn. 3 and substituting into Eqn. 2 the response of the neuron as far as it is determined by p_0 and the STRF, i.e. the *STRF-based response* is obtained:

$$\hat{p}(t) = p_0 + p_1(t) \quad (12)$$

A linear characterisation calculated in this way is shown in Fig. 5b. As can be seen there is some similarity between the experimentally measured PSTH and the theoretical characterisation. At least one systematic deviation is noted: while the PSTH is always non-negative, as it should be by definition, the characterisation sometimes assumes negative values. This is caused by the neglect of the spike generating mechanism (cf. Fig. 2) and the nature of the neural activity. A heuristic approach is followed to take this into account: a rectification stage is added in cascade to the linear prediction (Fig. 3).

Two choices have been made for the rectifier:

1. A one-sided linear rectifier giving zero output for negative input values, which

Fig. 1. Illustration of the calculation of the spectro-temporal receptive field (STRF) for 5 frequency bands. The average dynamic spectrum of the pre-event stimulus ensemble is shown in a. In b the average dynamic spectrum of the stimulus ensemble has been subtracted to give $R_{kp}(\tau)$. The amplitude of the Fourier transform of b is shown in f. This Fourier transform is multiplied by the window shown in g. This window is 1 for $|\omega| \leq \Omega$, a squared cosine $\cos^2\{\pi(\omega - \Omega)/2\Omega\}$ for $\Omega < |\omega| \leq 2\Omega$ and zero otherwise, in which $\Omega = 0.1302 \omega_0$ Hz where ω_0 is the central frequency of the bandpass filter used. The result of this windowing is shown in h. $R_{kk}(\tau)$ is shown in c, which is multiplied with a window shown in d. This window is 1 for $|t| \leq T$, $\cos^2\{\pi(t - T)/2T\}$ for $T < |t| \leq 2T$ and zero otherwise, in which $T = 3.072/\omega_0$ s. The result is shown in e. This is Fourier transformed, the amplitude of which is shown in i. Then the Fourier transform, shown in h, is divided by i, the result of which is shown in j. The inverse Fourier transform is then shown in k: the STRF. This is in its turn windowed l. The window was 1 for regions in which the STRF differed from zero, while the tails of the window consisted of the same squared cosines as in d. The final result, shown in m, is used to calculate the STRF-based response.

gives a modified characterisation

$$\begin{aligned} \hat{p}_l(t) &= a_l \hat{p}(t) && \text{for } \hat{p}(t) > 0 \\ &= 0 && \text{for } \hat{p}(t) \leq 0 \end{aligned} \tag{13}$$

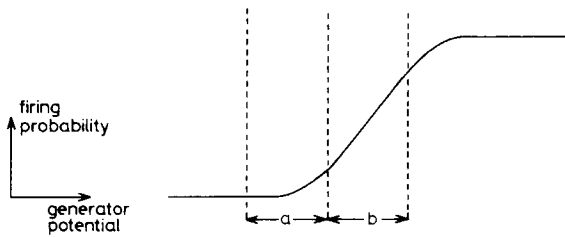


Fig. 2. Example of a static nonlinearity of a spike-generating mechanism. In the interval a a strong rectifying effect is present, whereas in b the relation between the probability of an action potential and the generator potential is about linear.

where $\hat{p}(t)$ is the linear prediction given by Eqn. 12. The constant a_l is chosen such that the average level of the prediction $\hat{p}_l(t)$ is again equal to p_o .

2. A one-sided quadratic rectifier giving zero output for negative input and quadratic output for positive input, which gives the modified characterisation

$$\begin{aligned} \hat{p}_q(t) &= a_q \hat{p}^2(t) && \text{for } \hat{p}(t) > 0 \\ &= 0 && \text{for } \hat{p}(t) \leq 0 \end{aligned} \tag{14}$$

the constant a_q is again chosen to assure that the average of $\hat{p}_q(t)$ is p_o .

These procedures lead to characterisations which on one hand are not completely consistent with the general theoretical approach presented, but on the other hand do eliminate some a priori impossible aspects from the linear characterisation. Its practical value should be estimated from the experimental data; its theoretical

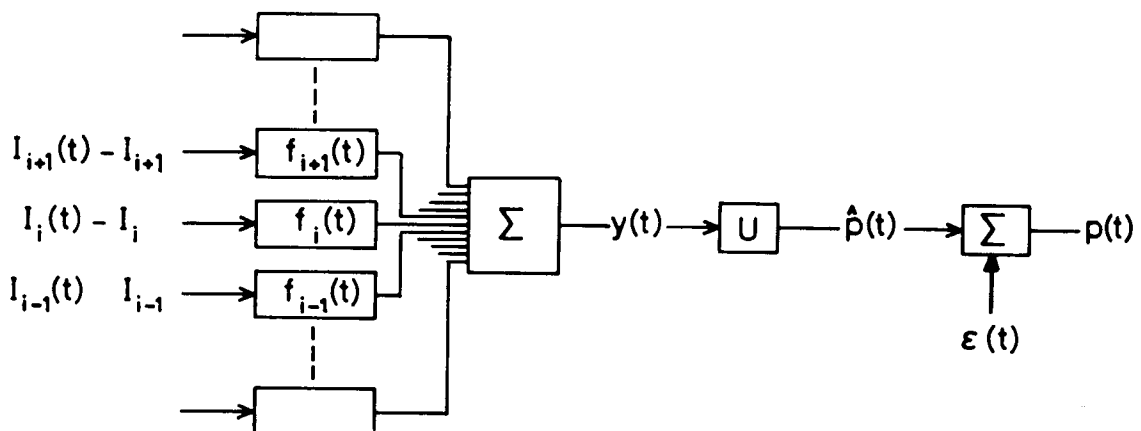


Fig. 3. Functional model of an auditory neuron consisting of a parallel set of linear filters with impulse responses $f_i(t)$, the outputs of which are summated and passed through a static nonlinearity. It must be noted that $\epsilon(t)$ does not represent a 'noisy' component but a functional of input $\{P_k(t) - P_{k^*}\}$.

evaluation is outside the limits of this paper. Results of these procedures are shown in Fig. 5c and 5d for the linear and quadratic one-sided rectifier respectively.

Quality of the response characterisation

In order to evaluate the quality of the calculated response as compared to the measured response of the neuron two measures can be chosen. One may take the deviation of the calculated data from the experimental ones or one may take their similarity. The deviation is given by the relative mean square error defined as

$$\delta_{p\epsilon}^2 = \frac{\sigma_\epsilon^2}{\sigma_p^2} \quad (15)$$

where σ_ϵ^2 is the mean square error

$$\sigma_\epsilon^2 = \overline{\{p(t) - \hat{p}(t)\}^2} \quad (16)$$

where the bar denotes 'taking the time average', and σ_p^2 is the variance of the measured PSTH:

$$\sigma_p^2 = \overline{\{p(t) - p_o\}^2} \quad (17)$$

The similarity is given by the correlation coefficient [18]

$$\rho_{p\hat{p}} = \frac{\overline{\{p(t) - p_o\}\{\hat{p}(t) - p_o\}}}{\sigma_p \sigma_{\hat{p}}} \quad (18)$$

where σ_p^2 is the variance in the PSTH and $\sigma_{\hat{p}}^2$ is the variance in the calculated response curve.

These two measures are not independent; for the linear characterisation procedure leading to Eqn. 12 there exists a simple relation [18]:

$$\rho_{p\hat{p}}^2 = 1 - \delta_{p\epsilon}^2 \quad (19)$$

If the neural unit behaves as a linear system and no stochastic fluctuations in the measured PSTH are present then the characterisation will be perfect. In that situation $\delta_{p\epsilon} = 0$ and $\rho_{p\hat{p}} = 1$. If the neural unit behaves as a nonlinear system and/or stochastic fluctuations are present, then $\delta_{p\epsilon} > 0$ and $\rho_{p\hat{p}} < 1$.

The linear prediction procedure can be extended to nonlinear prediction in two ways. The general approach would be the measurement and use of higher order correlation functions as in the Volterra–Wiener characterisation of nonlinear systems [15]. If applied correctly this leads to a decrease in the error $\delta_{p\epsilon}$ and an increase in the similarity $\rho_{p\hat{p}}$. A more specific and model-oriented approach can be based on the incorporation of information available for the system under consideration. The use of a modified characterisation procedure as given by Eqns. 13 and 14 is an example of such approach. We expect in these situations a decrease of $\delta_{p\epsilon}$ and an increase of $\rho_{p\hat{p}}$.

To evaluate the quality of the characterisation we will use $\rho_{p\hat{p}}$. The results for 12 units are given in Table I. The similarity for the linear prediction is given by $\rho_{p\hat{p}}$ in the second column, the modified result with the one sided linear rectifier by $\rho_{p\hat{p}_1}$ in

TABLE I
SIMILARITIES BETWEEN MEASURED PSTH AND STRF-BASED RESPONSE

N , number of presentations of the noise sequence; $n_{\hat{p}}$, number of smoothings to the STRF-based response; n_p , number of smoothings to the PSTH; uns., unsmoothed; smo., smoothed.

Unit	$\rho_{p\hat{p}}$		$\rho_{p\hat{p}_1}$		$\rho_{p\hat{p}_q}$		$\rho_{p'p''}$		N	$n_{\hat{p}}$	n_p
	uns.	smo.	uns.	smo.	uns.	smo.	uns.	smo.			
133-4	0.70	0.77	0.75	0.81	0.68	0.71	0.54	0.71	32	1	2
161-4	0.44	0.53	0.54	0.64	0.58	0.66	0.85	0.94	64	1	2
161-5	0.44	0.52	0.53	0.62	0.60	0.66	0.42	0.77	16	0	4
166-9	0.57	0.70	0.67	0.81	0.79	0.88	0.90	0.93	32	1	2
167-4	0.20	0.22	0.26	0.28	0.24	0.26	0.50	0.66	16	0	2
167-6	0.36	0.42	0.45	0.50	0.57	0.64	0.92	0.95	64	1	2
168-1	0.64	0.58	0.69	0.62	0.70	0.60	0.08	0.21	32	1	2
169-3	0.49	0.56	0.57	0.62	0.61	0.64	0.20	0.46	32	0	4
171-2	0.36	0.37	0.45	0.45	0.55	0.52	0.56	0.83	16	0	2
174-1	0.30	0.30	0.37	0.36	0.51	0.49	0.51	0.75	16	0	2
175-2	0.28	0.35	0.35	0.42	0.50	0.65	0.78	0.93	32	1	2
177-2	0.39	0.45	0.48	0.55	0.66	0.72	0.87	0.90	32	1	2

the third column and for the one-sided quadratic rectifier by $\rho_{p\hat{p}_q}$ in the fourth column.

The system characterisation as applied here is valid for deterministic systems. The nervous system, however, is inherently stochastic. Part of this stochastic component is eliminated by the averaging procedure to obtain the PSTH, $p(t)$. Part of the variability, however, remains and leads to a decrease of similarity between the STRF-based response and the experimental data. In order to estimate the variability in the PSTH we measured it twice, leading to the estimates $p'(t)$ and $p''(t)$, and correlated the first PSTH $p'(t)$ with the repeat determination $p''(t)$.

The quantitative results for $\rho_{p'p''}$ are given in the fifth column of Table I. It appears from the table that $\rho_{p\hat{p}}$ is uncorrelated with $\rho_{p'p''}$.

Smoothing

When comparing the autocovariance function of $p(t)$ with the autocovariance function of $\hat{p}(t)$ we observe a much sharper peak at $\tau = 0$ for the experimental PSTH, $p(t)$ as shown in Fig. 4a, b. This indicates that the action potentials are more accurately timed than can be derived from the STRF-based response $\hat{p}(t)$. In other words, the various $p(t)$ contain comparatively more high frequencies than the corresponding $\hat{p}(t)$. The averaging procedures to obtain the STRF as well as the bandwidth of the filters in the dynamic spectrum analyser obviously destroy the fine timing relations between stimulus and response. In order to get the spectral contents of $p(t)$ and $\hat{p}(t)$ more in line, $p(t)$ was smoothed two or four times by applying a Hanning window in the frequency domain [15]. When this procedure removed too much of the high frequencies in $p(t)$, $\hat{p}(t)$ was Hanning filtered once (cf. Fig. 4c, d).

The effect on the real data can be seen in Figs. 6 and 7. The effect on the various correlation coefficients is shown in Table I on the right-hand side of each column.

Furthermore, since the $p(t)$ were obtained from at most 64 presentations (see

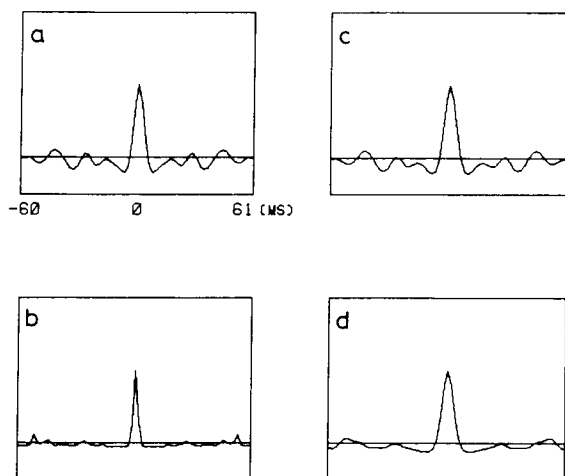


Fig. 4. Autocovariance functions of the unsmoothed and smoothed versions of the STRF-based response and the PSTHs of unit 161-4. a and b show the autocovariance functions of the unsmoothed STRF-based response and the PSTH, respectively. c shows the autocovariance function of the STRF-based response after it was Hanning-filtered once. d shows the autocovariance function of the twice Hanning-filtered PSTH.

sixth column in Table I) and generally less (32 or 16 presentations) the smoothed $p(t)$ may appear to be a better estimate of the neuron's firing probability especially since the firing rate is rather low. The number of smoothings applied to $\hat{p}(t)$ and $p(t)$ is given in the last column of Table I.

Results

Reproducibility of experimental results

The reproducibility of an STRF was tested for six neural units by presenting the same noise twice and comparing the STRF-based responses for both presentations. Even when the calculation of the STRF was based on only some hundreds of pre-event stimuli, the STRF-based responses were almost identical with correlation coefficients between the two being 0.986 or more.

The reproducibility of the actual PSTHs could, however, be very low. The correlation coefficient between the PSTH obtained in the stimulus presentation resulting in the STRF and a second PSTH to an identical stimulus could be as low as 0.08 and as high as 0.92 (no smoothing was applied) and from 0.21 to 0.95 when smoothing was applied as shown in Table I fifth column.

The quantitative characterisation

The results of the quantitative characterisations are shown in Table I. The second column shows the similarity between actual response and STRF-based response, with and without smoothing, for the linear procedure. For nine units the similarity was less than 0.5 (unsmoothed data), for three units better than 0.5. The best

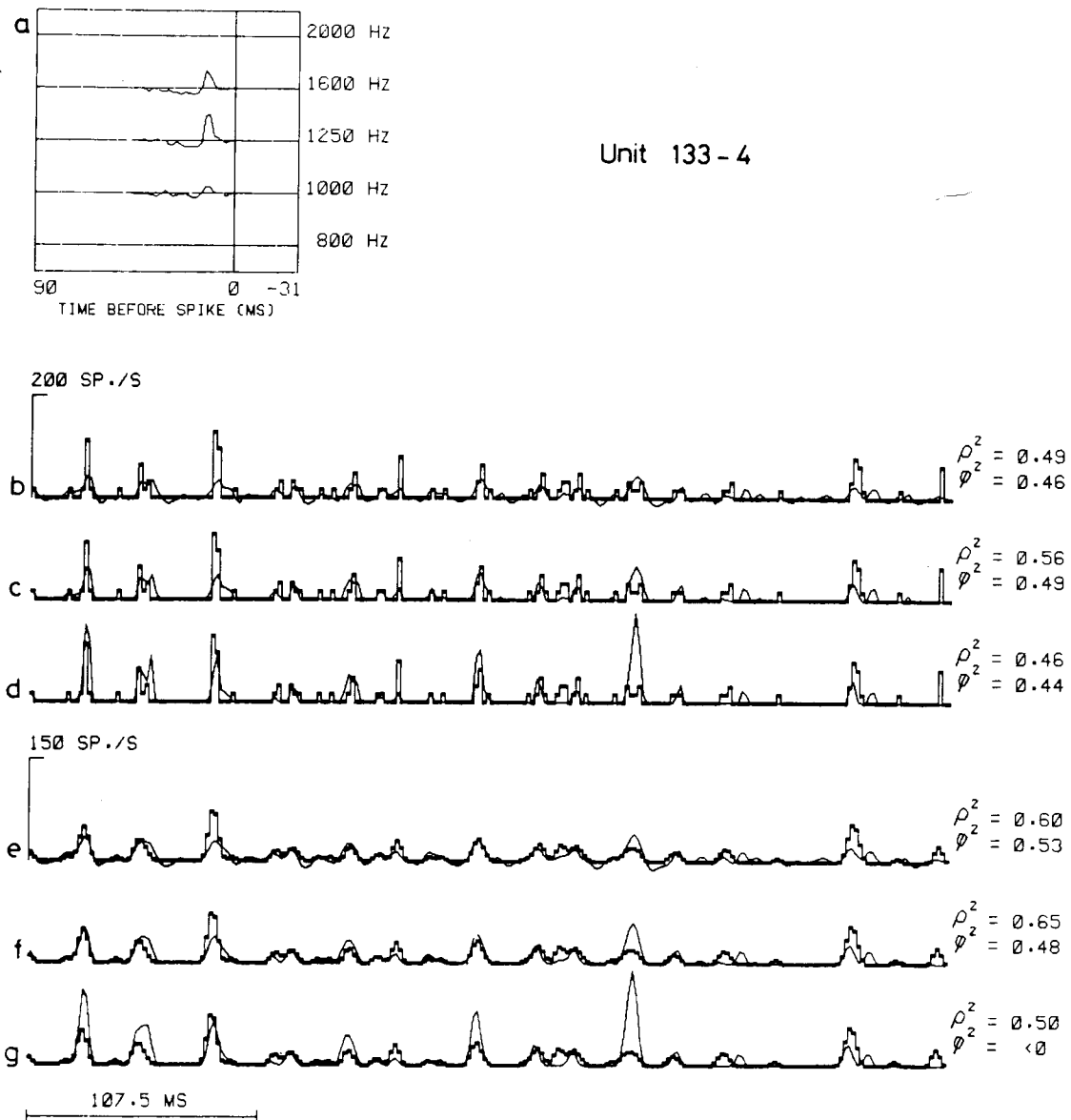


Fig. 5. Quantitative characterisation of unit 133-4. The STRF, shown in a, was calculated from 1288 pre-event stimuli while during the presentation of the stimulus 2087 events occurred. So, because of the dead time of the dynamic spectrum analyser 799 (38%) pre-event stimuli were omitted from the analysis. In b, c and d segments of the STRF-based responses (thin plot) and the PSTH (thick staircase plot) are shown without applying smoothing under no assumption of a static nonlinearity (b), under the assumption of a linear half-wave rectifier (c), and under the assumption of a quadratic halfwave rectifier (d). In e, f and g the same STRF-based responses and PSTHs are shown after applying smoothing. (Note the different scalings.)

characterisation was obtained for unit 133-4, the results of which are shown in Fig. 5b. This neuron was most sensitive for frequencies in the 1250 Hz band, and showed a clear postactivation suppression expressed as a negative region in the STRF that precedes the positive one [9]. This unit was somewhat better characterised by its STRF under the assumption of a linear half-wave rectifier (Fig. 5c) as can be seen from the third column in Table I.

All units, except unit 133-4 (Fig. 5d) and 167-4, were best characterised under the assumption of a quadratic halfwave rectifier (see Table I, fourth column). For example unit 166-9 could then be characterised with a similarity of 0.79. The STRF,

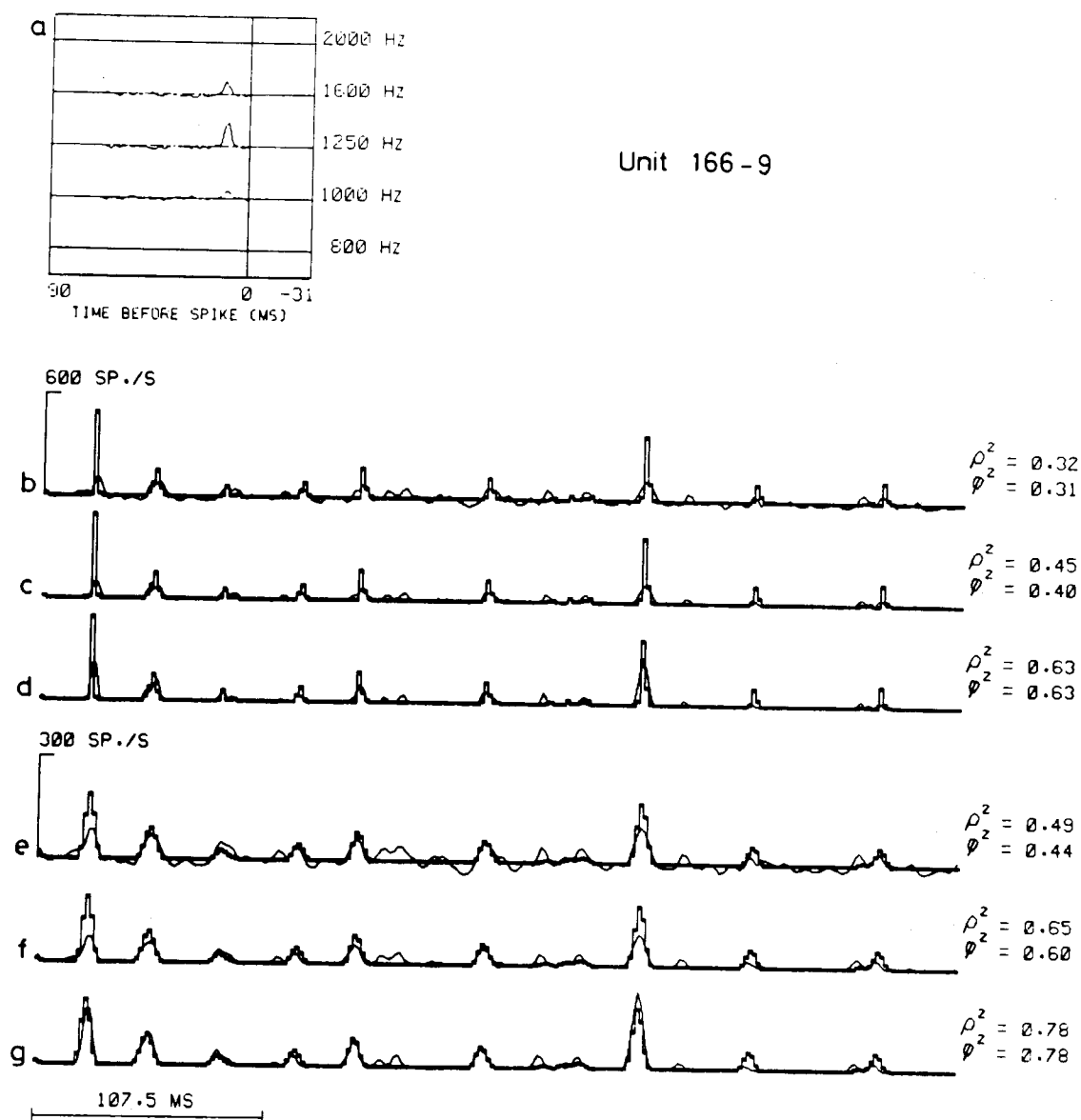


Fig. 6. Quantitative characterisation of unit 166-9. The STRF (a) was calculated from 1645 pre-event stimuli out of 2823 available events. Therefore 42% of the spikes did not contribute to the analysis.

the STRF-based responses and the PSTHs of this unit are shown in Fig. 6. The unit was most sensitive for frequencies in the 1250 Hz band and hardly showed any postactivation suppression.

Better characterisations were for most units obtained when the smoothing procedure was applied. The results are found on the right-hand sides of column 2, 3 and 4 in Table I. In the linear procedure unit 133-4 was again the best characterised unit (Fig. 5e). For most of the units the one-sided quadratic rectifier assumption resulted in the best characterisation. Unit 166-9 could then be characterised with a similarity of 0.89.

The results for unit 161-4 are shown in Fig. 7. This neuron was maximally sensitive in the 630 Hz band. Clear postactivation suppression can be observed. The unit was best characterised under the assumption of a quadratic halfwave rectifier. Part of the traces show high activity in the STRF-based responses while the PSTH is zero (see arrows). All significant major peaks in the PSTH are, however, also present in the STRF-based response. The data shown are the smoothed versions.

The results of unit 175-2 are shown in Fig. 8. This unit was maximally sensitive in the 630 Hz band. Hardly any postactivation suppression was seen. The best characterisation was obtained under the assumption of a quadratic halfwave rectifier. This unit fired only at three instances in the segments shown. Inspection of the

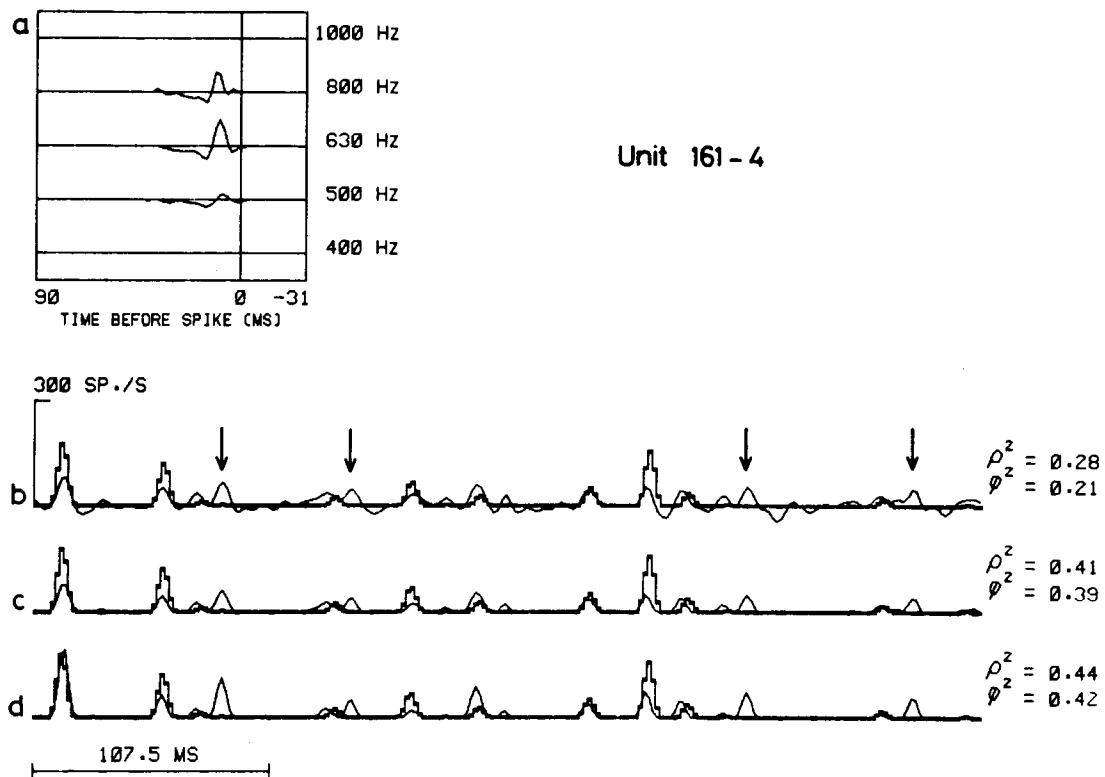


Fig. 7. Quantitative characterisation of unit 161-4. The STRF (a) was calculated from 3582 pre-event stimuli, 2963 (45%) were omitted from the analysis. Only the results after applying smoothing are shown. The arrows indicate where the STRF-based response is high, whereas the PSTH is zero.

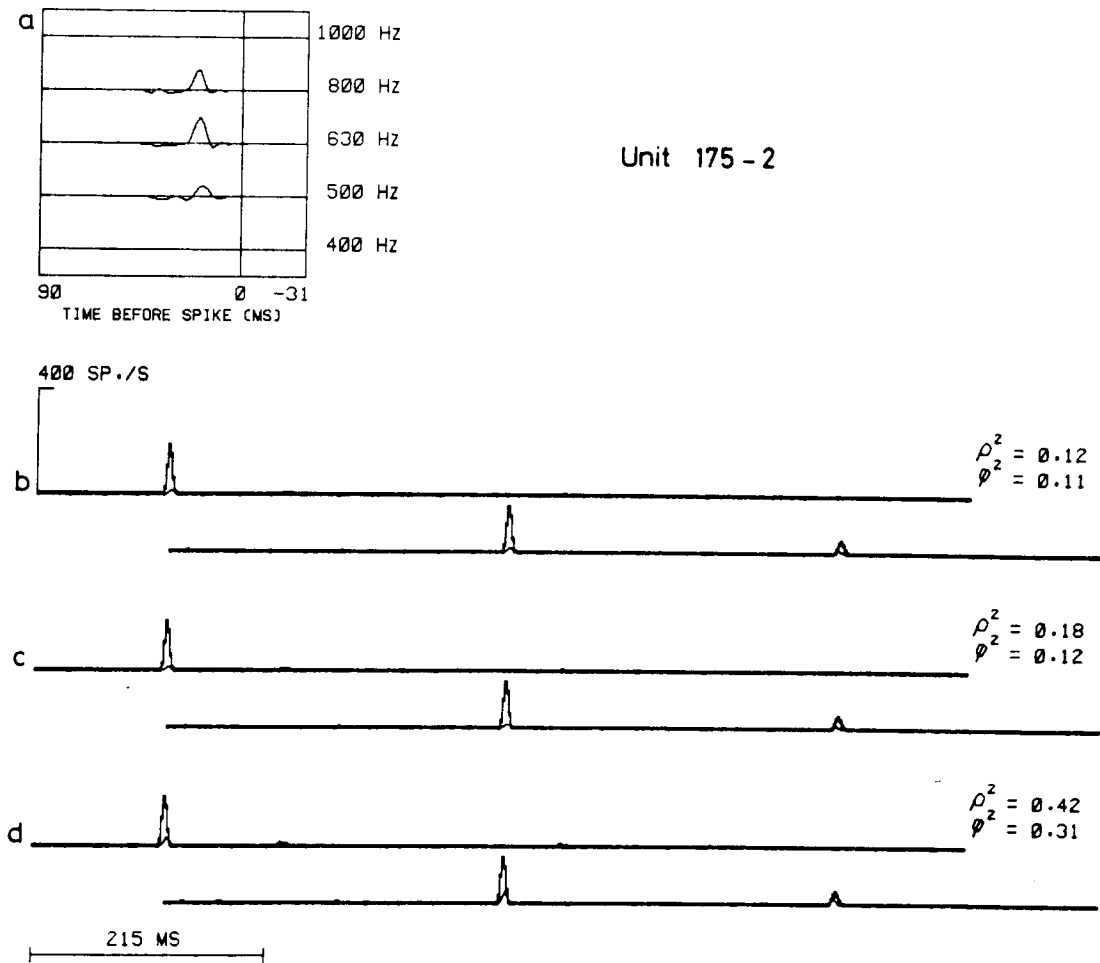


Fig. 8. Quantitative characterisation of unit 175-2. The STRF (a) was calculated from 860 pre-event stimuli, 54 (6%) were omitted from the analysis. Only the results after applying smoothing are shown.

STRF-based responses showed that these are the instances where the calculated response was highest. These instances were very reproducible in the PSTH as shown by a $\rho_{p'p''} = 0.93$.

The results for unit 169-3 are shown in Fig. 9. The best frequency of this unit fell in the 250 Hz band. A region of activation is preceded by a region of suppression that also extended upwards in the higher frequency bands and in time equal to the presence of the activation region. This lateral suppression [9] may appear unclear from the figure but becomes evident when the inherent delay of the bandpass filters is taken into account [2]. Furthermore, this unit had a first-order cross-correlogram (reverse correlation function [5]) that differed significantly from zero, indicating that this unit responded in phase-lock to the noise signal. The correlogram was shown in Fig. 8 of [9]. To compensate for this phaselock the STRF-based responses as well as the PSTHs for inverted and non-inverted noise were averaged. For the smoothed PSTH the similarity with the STRF-based response was 0.57 under the assumption of the linear halfwave rectifier.

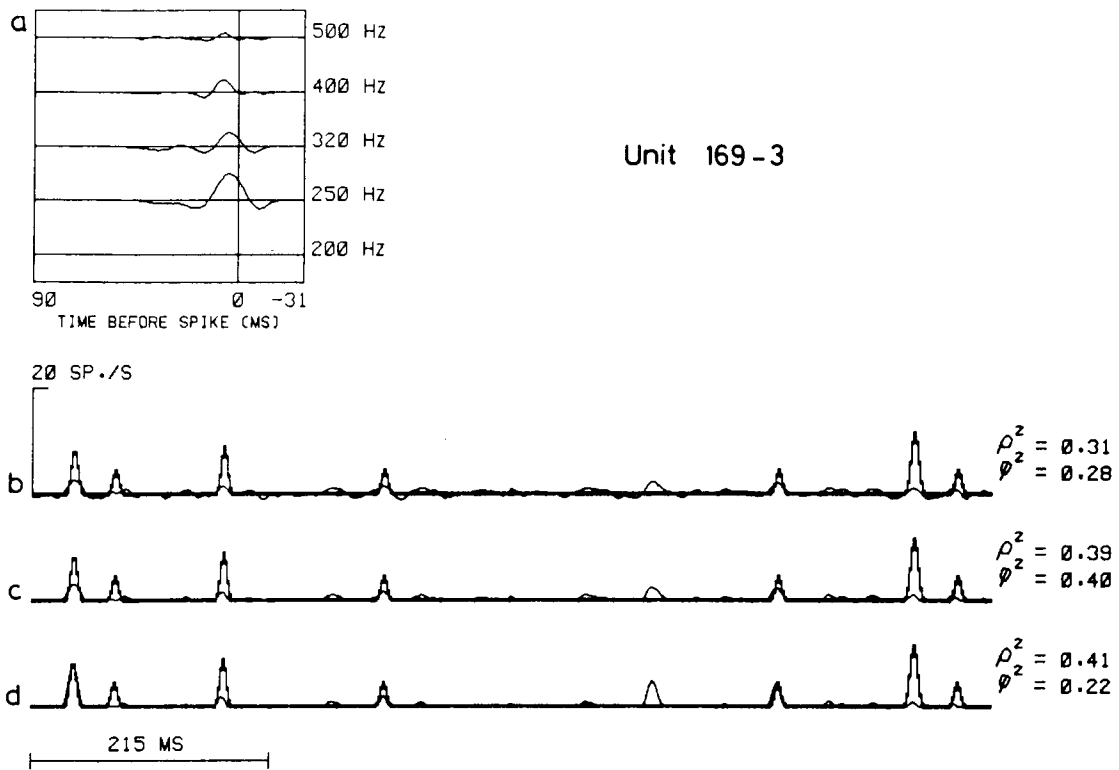


Fig. 9. Quantitative characterisation of unit 169-3. The STRF (a) was calculated from 270 pre-event stimuli, 2 (1%) were omitted from the analysis. Only the results after applying smoothing are shown.

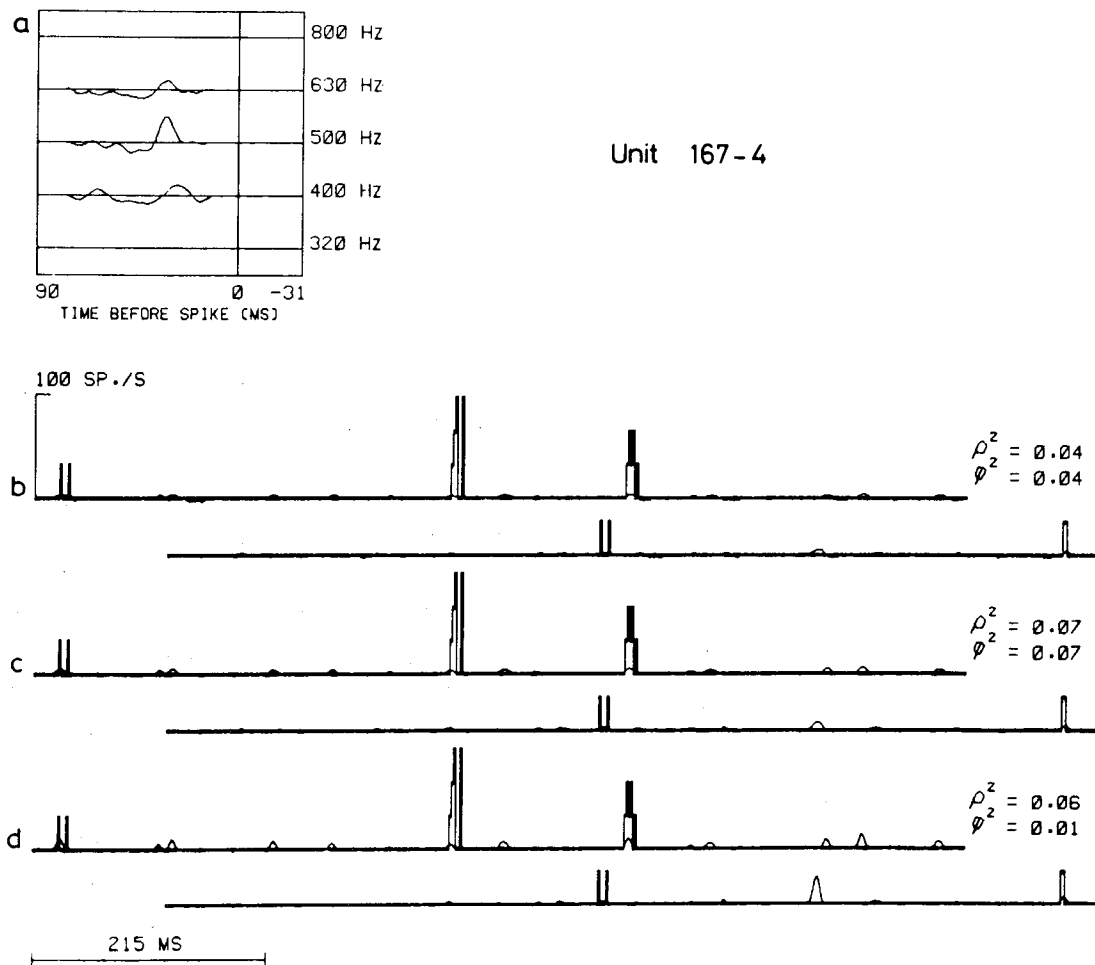


Fig. 10. Quantitative characterisation of unit 167-4. The STRF (a) was calculated from 245 pre-event stimuli, 95 (28%) were omitted from the analysis. Only the results without applying smoothing are shown.

The worst characterisation was obtained for unit 167-4 (Fig. 10). Its similarity between STRF-based response and PSTH did not exceed 0.24. Its best frequency fell in the 500 Hz band; some postactivation suppression can be observed. Close inspection of the PSTHs revealed that this unit unlike the other ones often fired twice. To show this most clearly the unsmoothed results are shown. The first spike always came at an instance that the STRF-based response was high. Many parts show high calculated responses also at moments that the PSTH is zero. It must be remarked that, due to the dead time of the dynamic spectrum analyser, the second firings were not taken into account in the calculation of the STRF (see Discussion).

Discussion

The spectro-temporal receptive field (STRF)

In the present study we have investigated auditory neurons in the midbrain of the lightly anaesthetised grass frog using Gaussian wide-band noise. This stimulus is a good approximation of Gaussian white noise as long as the bandwidth of the noise is considerably larger than the frequency band to which the auditory neurons are sensitive. For the characterisation procedure we used the dynamic spectrum of the noise as input to the neurons. This has several advantages in this case since the dynamic spectrum represents spectro-temporal characteristics of the sound signal that are believed to play an important role in central auditory information processing [1,19,21,23,26]. When Gaussian wide-band noise is used as stimulus, the average pre-event dynamic spectrum, or the STRF derived from it, presents a lucid picture of the spectro-temporal differences between the PESE and the SE. The interpretation of the STRF is much more straightforward than the interpretation of the second-order Wiener kernel to which it is related [3]. Properties like spectral sensitivity, latency, post-activation suppression and lateral suppression are clearly exhibited in the STRF [9].

The determination of the STRF on basis of the dynamic spectrum allows an implementation in hardware, the dynamic spectrum analyser [2], thereby reducing the necessary computer power and at the same time allowing a real-time availability of the neuron's spectro-temporal sensitivity during the experiment. A disadvantage of considering the dynamic spectrum as input to the neuron will emerge when the neurons respond to properties of the stimulus that are lost in the dynamic spectrum. This is inherently connected to the present approach which essentially is a 'model-based' one, in contrast to the more general Wiener-type methods which also take the results from first-order cross-correlation [5,17] into account. On the other hand, model assumptions generally lead to more specific and more efficient identification techniques.

The actual procedure to determine the STRF (cf. [9]) is similar to the equation presented by Schetzen [20] in which the estimation of the first-order Wiener kernel is corrected for a non-white input that has, however, still to be Gaussian. The estimation of the STRF does neither require the input to be white nor Gaussian. Moreover, when the input lacks certain frequencies these will also be absent in the

input–output correlogram. This shows that the deconvolution required to estimate the STRF can theoretically be carried out for any stationary stimulus with uncorrelated frequency bands, whatever the amplitude distribution or spectral content of its dynamic spectrum. For all these stimuli the STRF produces the optimal characterisation in mean-square-error sense when a time-invariant system with finite memory is described according to Eqns. 2 and 3. It must be emphasized, however, that the characterisation is only optimal for the input with which the STRF is determined.

So, the characterisation procedure does not require the input to be white or Gaussian. However, the interpretation of an STRF in terms of activation and suppression regions [9] does require that the presence of these regions reflects neuronal characteristics and not possible correlations within the dynamic spectrum of the stimulus. For this reason Gaussian wide-band noise was used as stimulus.

Sources of error in the characterisation procedure

The calculation of the STRF is subject to several errors. First, because of the dead-time of the DSA a sometimes considerable number of elements of the PESE was omitted in the construction of the STRF. These stimulus elements were followed by spikes that were less than 107.52 ms preceded by other spikes. These omitted elements of the PESE may constitute a subset of the PESE with spectro-temporal properties that are different from those of the PESE [25]. Comparison of the average pre-event dynamic spectrum obtained with the DSA and the average pre-event complex spectro-temporal intensity density [9], where the complete PESE was taken into account, however, showed no indication of major errors. In both cases the distribution of suppression and activation regions in the STRF were similar. When a unit structurally fires twice at some instances, as unit 167-4, the stimuli preceding these second firings will not be represented in the STRF. This may explain why in the STRF-based response the second firings (cf. Fig. 10) are not present.

Second, the estimations of $\{P_{k\circ}\}$ and $\{R_{kk}(\tau)\}$ are subject to some errors. The estimation of $\{P_{k\circ}\}$ was accurate compared with that of $\{R_{kk}(\tau)\}$. The biggest component in this latter error originates in the finite record length of the segments with which $\{R_{kk}(\tau)\}$ was estimated.

A third error was introduced in the deconvolution of $\{R_{kp}(\tau)\}$ by $\{R_{kk}(\tau)\}$. This procedure was performed in the frequency domain (cf. Fig. 1). $\{R_{kk}(\tau)\}$ was first windowed in the time domain in order to reduce the variance of its Fourier transform. Then, the Fourier transform of $\{R_{kp}(\tau)\}$ was divided by the Fourier transform of this windowed $\{R_{kk}(\tau)\}$. Despite this windowing procedure high frequencies in $\{R_{kp}(\tau)\}$ had subsequently to be eliminated because otherwise several numerical errors produced severe oscillations in the STRF. Nevertheless, some oscillatory phenomena remained. This explains, for example, the small negative parts in the STRF at about 8 ms before the spike in Fig. 7a.

A fourth error was introduced by the use of pseudo-random binary-noise sequences. $\{R_{kk}(\tau)\}$ is a fourth-order characteristic of noise (cf. Eqn. 4) and it is not known if the pseudorandom sequences used in this study were satisfactory approximations of perfectly random signals with ideal fourth-order properties (cf. [6,22]). It appeared that fluctuations far outside the proper STRF were reproducible

and did not diminish when more than about 1000 pre-event stimuli were analysed. These fluctuations are most likely due to statistical relations between the various frequency bands of the noise stimulus. Their maximum amplitudes could amount to 20% of the highest peak in the STRF. When they were outside the STRF they could be removed by use of windowing. Inside this window they will lead to a systematic error.

The peristimulus time histogram (PSTH)

The PSTH was considered to represent the output of the neuron because it is an experimental estimate of the instantaneous firing rate or event density of the neuron. By the stochastic nature of the event generation process, the PSTHs remain subject to sometimes large statistical fluctuations. The correlation coefficient of the two PSTHs to an identical stimulus could be as low as 0.08 and was never higher than 0.92. This means that, in terms of a signal plus noise model, between 92% and 8% of the variance of the PSTHs were due to noise on them. These errors can theoretically be reduced by presenting the stimulus many more times. When under a certain condition the correlation coefficient is 0.5, the stimulus period has to be presented 81 times more often to obtain a correlation coefficient of 0.9, if at least the noise is uncorrelated with the signal. This, however, would require a stimulus duration of hours, which is generally much longer than the experimental lifetime of a neuron. Another solution can be found in using much shorter noise sequences. But because these short sequences have worse statistical properties (e.g. of $R_{kk}(\tau)$) they do not allow a satisfactory interpretation in terms of activation and suppression of the STRF which is estimated on basis thereof. It must be emphasized that the amount of reproducibility of the PSTH did not seem to influence the similarity between the STRF-based response and the measured PSTH (cf. Fig. 4).

The reproducibility of a PSTH is often improved by additional smoothing. Marmarelis and McCann [14] used a smoothing window with a bandwidth that was derived from the cross-correlogram of the two PSTHs to the same stimulus. French and Holden [7] used a window that eliminated the frequency content of the PSTH as far as it was higher than the bandwidth of the stimulus. In the smoothing procedure that was employed in the present study every action potential that contributes to the PSTH is treated equally. So only the timing properties of the neuron are influenced to some extent. The amount to which this influences the various similarity measures can be obtained from Table I.

The static nonlinearity

When one inspects the PSTHs of the neurons in response to a sequence of noise, one observes in many cases lengthy intervals where the PSTH is zero. This means that the firing probability of the neuron will be very low in these parts. In some units (see e.g. unit 175-2 in Fig. 8) it even appeared that the firing probability only obtained realistic values during some very small parts of the stimulus presentation. This suggests a kind of threshold mechanism, a strong static nonlinearity. For this reason we have tried to improve the characterisation of the neurons by two a priori assumed relatively simple nonlinearities. The choice of these nonlinearities was

based on the assumption that they approximated a threshold mechanism in the spike generator of the neuron. The threshold was put at $\hat{p}_l(t) = \hat{p}_q(t) = 0$ (Eqns. 13 and 14). This choice is in fact arbitrary. In some units better results could be obtained by assuming a higher threshold. An example is presented in Fig. 8, where comparison of $\hat{p}(t)$ and $p(t)$ revealed that the unit simply fired at the three instances where $\hat{p}(t)$ was above a certain level. By raising the threshold up to just below these three peaks the correlation coefficient increased from 0.35 to 0.80 under the assumption of a linear halfwave rectifier. This shows that this method may reveal the character of the spike-generating mechanism of the neuron (or from more peripheral neurons afferent to the neuron recorded from). When the system consists of a linear part followed by a non-even static nonlinearity, the first-order Wiener kernel equals the impulse response of the linear part except for a constant factor. This means that the output of the linear part can be calculated except for a constant factor. The character of the nonlinearity can then simply be found by plotting this calculated output of the linear part against the actual output of the system. This requires, however, that the input of the system is Gaussian noise. This was obviously not fulfilled in this study because the dynamic spectrum of Gaussian noise was taken as the input to the system. The incorporation of static nonlinearities in this study is justified by the practical argument of the often considerable improvements that were obtained for the characterisation. It provides a further step towards a model-based identification technique to characterise auditory neuron response functions. This type of approach seems unavoidable in order to gain a basic understanding of the transduction process; at the same time it may circumvent some of the difficulties associated with the more general approach [12].

Dynamic nonlinearities

Despite all these manipulations the neurons could only be characterised on the basis of their STRFs to a limited extent. A significant fraction of the unexplained part of the response of the unit will be due to the errors described above. Another part will probably be due to dynamic nonlinearities in the neuron. A realistic example of such nonlinearities may be found in interactions between different parts of the STRF. When the STRF, for example, exhibits postactivation suppression in a certain frequency band, the average pre-event dynamic spectrum shows that the unit on the average fired when the temporal intensity of this band is first low and then high. This interpretation applies only to the average. One can then imagine two possibilities. First that a low intensity raises the firing probability of the neuron for some time independent of what comes later, while a high intensity raises the firing probability independent of what happened before. On the other hand, one can imagine that the unit fires only after a high-intensity part of the stimulus in strict combination with a preceding low-intensity part. The averaging procedure does not distinguish between these two possibilities and higher order analyses are necessary to reveal them. It is rather likely that mechanisms like these are present. In most units high parts in the PSTH corresponded to high parts in the STRF-based response. High activity in the STRF-based response was, however, regularly observed where the PSTH was low or zero. When postactivation suppression is present, a possible

explanation might be that this high activity in the STRF-based responses is due to either a low or a high intensity in the frequency band to which the unit is sensitive without the necessary succession of a low and a high intensity being realised. Similar arguments might be put forward for different parts of an activation region, for example. This indicates that higher order analyses may considerably improve the characterisations of the units. This is also suggested by the results of Bibikov and Gorodetskaya [4] who observed nonlinear phenomena for neurons in the torus semicircularis of the frog in response to amplitude modulation of tones, whereas Møller [16] describes a basically linear response in the cochlear nucleus of the rat. Finally, the rare occurrence of a high-activity part in the PSTH when the STRF-based response was low indicates that responses to harmonics of the stimulus only play a role as far as they affect a rectification of the response.

The quantitative characterisations

The results show that the response of the neurons to stimulation with Gaussian wide-band noise can to some extent be derived from its STRF. All information used to obtain these characterisations was derived from a comparison of the average dynamic spectrum of the PESE and the statistical properties of the dynamic spectrum of the stimulus. This means that to the extent to which the unit is characterised, the response of the neuron is based on properties of the stimulus that are present in its dynamic spectrum. It was argued that better characterisations might possibly be obtained when more suitable static nonlinearities are tested and when higher-order contributions are taken into account. This would indicate that the response of these units to noise then could to a considerable part be derived from the dynamic spectrum of the stimulus.

From the point of view of nonlinear-system identification the characterisation can be much more straightforward if the conventional first- and second-order Wiener kernels are calculated. It would, however, have taken much more computation time to get the present characterisations. In addition, the interpretation of these kernels is difficult, and therefore it would be hard to relate the obtained results to properties of the investigated neurons. The STRF, in contrast, presents an intuitively lucid picture of various response properties that are assumed relevant for auditory information processing. In this study it was deliberately attempted by the model approach to relate the results to physiological phenomena as generator potentials, thresholds and postactivation suppression (short time adaptation). Only when this is done can the characterisation be more than a series of figures and may describe part of the transducer mechanism.

Appendix 1: Statistical characteristics of the stimulus ensemble as measured by the dynamic spectrum analyser

For the derivation of the spectro-temporal receptive field we need the following statistical characteristics of the dynamic spectrum of the stimulus: expected value $\{P_{k\circ}\}$ (cf. Eqn. 1) and autocorrelation function $\{R_{kk}(\tau)\}$ (cf. Eqn. 4). Furthermore,

we require that the temporal intensities of different frequency bands are uncorrelated (cf. Eqn. 5).

$\{P_{k\circ}\}$ was obtained by averaging the dynamic spectrum of noise signals preceding a few thousand random events. This was calculated for noise with a cut-off frequency of 1500 Hz as well as for noise with a cut-off frequency of 5000 Hz. As the overall intensities of these two different kinds of noise were equal, their intensities in a third-octave band differ by 5.2 dB, the 1500 Hz noise having the higher intensity. In the estimations of $\{P_{k\circ}\}$ this was found indeed.

$\{R_{kk}(\tau)\}$ was estimated by averaging the autocorrelation functions of the dynamic spectra $\{I_k(t)\}$ of eight different segments of the noise stimulus. These spectrograms were sampled with 224 points separated by 1.92 ms. It can be shown that for Gaussian noise the autocorrelation function of a 1/3-octave band of noise with a cut-off frequency of 1500 Hz will be 11.1 times (10.5 dB) the autocorrelation function of the same 1/3-octave band of noise with a cut-off frequency of 5000 Hz and the same overall intensity. For the 1/3-octave band with central frequencies, f_k , from 160 to 1250 Hz this appeared to be realised within 9%.

Although two neighbouring bandpass filters in the DSA showed overlapping amplitude characteristics from their -3 dB points on, actual calculation of the cross-correlation function $R_{kl}(\tau)$ of the temporal intensities of two neighbouring frequency bands of Gaussian noise showed that these were uncorrelated within the error resulting from the finite lengths of the eight records used and the systematic error due to the use of pseudo-random noise:

$$R_{kl}(\tau) = 0, \quad k \neq l \quad (\text{A 1.1})$$

Appendix 2: Orthogonal decomposition of the PSTH

To facilitate the actual determination of the STRF, $\{h_k(t)\}$, from experimentally measurable correlation functions we demand that the expansion in Eqn. 2 be such that the three components p_\circ , $p_1(t)$ and $\epsilon(t)$ are mutually orthogonal with respect to the stimulus ensemble involved, i.e. they should be uncorrelated:

$$R_{p_\circ p_1}(\tau) = 0 \quad \text{for all } \tau \quad (\text{A 2.1})$$

$$R_{p_\circ \epsilon}(\tau) = 0 \quad \text{for all } \tau \quad (\text{A 2.2})$$

$$R_{p_1 \epsilon}(\tau) = 0 \quad \text{for all } \tau \quad (\text{A 2.3})$$

This approach is similar to the Gram-Schmidt orthogonalisation procedure used in the Wiener-type methods (see e.g. [15,20]).

We start by noting that the conditions A 2.1 and A 2.2 are not really new, they follow in a straightforward manner from the definition of p_\circ , the average firing rate, which leads to expected values for $p_1(t)$ and $\epsilon(t)$ that are zero given the statistics of the stimulus ensemble used. To ensure the validity of Eqn. A 2.3 we demand that

$$R_{k\epsilon}(\tau) = 0 \quad \text{for all } k \text{ and } \epsilon \quad (\text{A 2.4})$$

i.e. the remainder $\epsilon(t)$ should be uncorrelated to the stimulus intensity in the various frequency bands. By substitution we will find out if this condition leads to a consistent set of relations, especially with regard to the expression which the STRF, $\{h_k(\tau)\}$, will have to fulfil. Substitution of $p_1(t)$ (see Eqn. 3) into

$$R_{p_1\epsilon}(\tau) = \int_{-\infty}^{\infty} p_1(t-\tau)\epsilon(t)dt \quad (\text{A 2.5})$$

and combination with A 2.4 leads to the desired orthogonality of $p_1(t)$ and $\epsilon(t)$ as expressed in (A 2.3).

The input-output correlation $R_{kp}(\tau)$ can be written (cf. Eqn. 2) as

$$R_{kp}(\tau) = R_{kp_0}(\tau) + R_{kp_1}(\tau) + R_{k\epsilon}(\tau) \quad (\text{A 2.6})$$

Because of the definitions of p_0 and I_k it is obvious that R_{kp_0} equals zero. Substitution of (A 2.4) then leads to

$$R_{kp}(\tau) = R_{kp_1}(\tau) \quad (\text{A 2.7})$$

By assuming a stimulus ensemble with uncorrelated intensities in different frequency bands we can use Eqn. 9 to yield

$$R_{kp}(\tau) = \int_{-\infty}^{\infty} h_k(\sigma)R_{kk}(\tau-\sigma)d\sigma \quad (\text{A 2.8})$$

which expresses the STRF, $\{h_k(\sigma)\}$, in terms of measurable quantities: the STRF equals the stimulus response correlation 'corrected' for the spectral composition of the stimulus ensemble.

By analogous reasoning it can be shown that the converse procedure, i.e. imposing Eqn. A 2.8 as a condition, leads to the decomposition of $p(t)$ as in Eqn. 2 that is orthogonal to the stimulus ensemble.

Acknowledgements

This investigation was supported by the Netherlands Organisation for the Advancement of Pure Research (ZWO). Dik Hermes performed all the calculations and elaborated on a slightly different theoretical procedure in his Ph.D. thesis: 'Spectro-temporal characterisation of auditory neurons in the torus semicircularis of the grass frog, *Rana temporaria* L.' (University of Nijmegen). A former version was critically read by Ton Vendrik. Brian Johnstone provided helpful comments for this final version. Wim van Deelen and Jan Bruijns were of great help in the various aspects of data analysis. Marianne de Leng prepared the manuscript.

References

- 1 Aertsen, A.M.H.J., Smolders, J.W.T. and Johannesma, P.I.M. (1979): Neural representation of the acoustic biotope: on the existence of stimulus-event relations for sensory neurons. *Biol. Cybernetics* 32, 175-185.

- 2 Aertsen, A.M.H.J. and Johannesma, P.I.M. (1980): Spectro-temporal receptive fields in auditory neurons in the grassfrog. I. Characterization of tonal and natural stimuli. *Biol. Cybernetics* 38, 223–234.
- 3 Aertsen, A.M.H.J. and Johannesma, P.I.M. (1981): The spectro-temporal receptive field, a functional characteristic of auditory neurons. *Biol. Cybernetics* 42, 133–143.
- 4 Bibikov, N. and Gorodetskaya, O. (1981): Coding of amplitude-modulated tones in the midbrain auditory region of the frog. In: *Neuronal Mechanisms in Hearing*, pp. 347–352. Editors: J. Syka and L. Aitkin. Plenum Press, New York.
- 5 De Boer, E. and de Jongh, H.R. (1978): On cochlear encoding: potentialities and limitations of the reverse-correlation technique. *J. Acoust. Soc. Am.* 63, 115–135.
- 6 Eckhorn, R. and Pöpel, B. (1979): Generation of gaussian noise with improved quasi-white properties. *Biol. Cybernetics* 32, 243–248.
- 7 French, A.S. and Holden, A.V. (1971): Alias-free sampling of neuronal spike trains. *Kybernetik* 8, 165–171.
- 8 Grashuis, J.L. (1974): The pre-event stimulus ensemble, an analysis of the stimulus-response relation for complex stimuli applied to auditory neurons. Ph.D. Dissertation, Catholic University, Nijmegen, The Netherlands.
- 9 Hermes, D.J., Aertsen, A.M.H.J., Johannesma, P.I.M. and Eggermont, J.J. (1981): Spectro-temporal characteristics of single units in the auditory midbrain of the lightly anaesthetised grass frog (*Rana temporaria* L.) investigated with noise stimuli. *Hearing Res.* 5, 147–178.
- 10 Hermes, D.J., Eggermont, J.J., Aertsen, A.M.H.J. and Johannesma, P.I.M. (1981): Spectro-temporal characteristics of single units in the auditory midbrain of the lightly anaesthetised grass frog (*Rana temporaria* L.) investigated with tonal stimuli. *Hearing Res.* 6, 103–126.
- 11 Johannesma, P.I.M. (1972): The pre-response stimulus ensemble of neurons in the cochlear nucleus. In: *Proceedings of the IPO Symposium on Hearing Theory*, pp. 58–69. Editor: B.L. Cardozo. Eindhoven, The Netherlands.
- 12 Johnson, D.H. (1980): Applicability of white-noise nonlinear system analysis to the peripheral auditory system. *J. Acoust. Soc. Am.* 68, 876–884.
- 13 Lee, Y.W. and Schetzen, M. (1965): Measurement of the Wiener kernels of a non-linear system by cross-correlation. *Int. J. Control* 2, 237–254.
- 14 Marmarelis, P.Z. and McCann, G.D. (1973): Development and application of white-noise modeling techniques for studies of insect visual nervous system. *Kybernetik* 12, 74–89.
- 15 Marmarelis, P.Z. and Marmarelis, V.Z. (1978): *Analysis of physiological systems: the white-noise approach*. Plenum Press, New York.
- 16 Møller, A.R. (1973): Statistical evaluation of the dynamic properties of cochlear nucleus units using stimuli modulated with pseudorandom noise. *Brain Res.* 57, 443–456.
- 17 Møller, A.R. (1977): Frequency selectivity of single auditory-nerve fibers in response to broadband noise stimuli. *J. Acoust. Soc. Am.* 62, 135–142.
- 18 Papoulis, A. (1965): *Probability, random variables and stochastic processes*. McGraw-Hill Kogakusha, Tokyo.
- 19 Scheich, H. (1977): Central processing of complex sounds and feature analysis. In: *Recognition of Complex Acoustic Signals*, pp. 161–182. Editor: T.H. Bullock. *Life Sciences Res. Rep. Vol. 5*, Dahlem Konferenzen, Berlin.
- 20 Schetzen, M. (1980): *The Volterra and Wiener Theories of Nonlinear Systems*. Wiley, New York.
- 21 Suga, N. (1978): Specialization of the auditory system for reception and processing of species-specific sounds. *Fed. Proc.* 37, 2342–2354.
- 22 Swerup, C. (1978): On the choice of noise for the analysis of the peripheral auditory system. *Biol. Cybernetics* 29, 97–104.
- 23 Symmes, D. (1981): On the use of natural stimuli in neurophysiological studies of audition. *Hearing Res.* 4, 203–214.
- 24 Wiener, N. (1958): *Nonlinear problems in random theory*. MIT Press, Cambridge, MA.
- 25 Wilson, J.P. and Evans, E.F. (1975): Systematic error in some methods of reverse correlation. *J. Acoust. Soc. Am.* 57, 215–216.
- 26 Worden, F.G. and Galambos, R., editors (1972): *Auditory processing of biologically significant sounds*. Neuroscience Research Program Bulletin 10. NRP, Brookline, Mass.