

H.R. DE JONGH

MODELLING THE PERIPHERAL
AUDITORY SYSTEM

FITTING
MODEL PULSE PATTERNS
TO
NEURONAL IMPULSE TRAINS

R E L

KNO-VERENIGING
MODELLING THE PERIPHERAL
AUDITORY SYSTEM

1978

MODELING OF SPATIALLY SEPARATED SOUND SOURCES

ACADEMISCH PROEFSCHRIFT

van de kandidaat voor de
graad van doctor in de
wetenschappen, te weten
de natuurkunde, door
de heer
Herman Bollen

Herman Bollen

Het onderzoek dat aan dit proefschrift ten grondslag ligt kon mede dankzij financiële steun van Z.W.O. worden verricht.
Dank aan allen die hebben bijgedragen tot het tot stand komen van dit proefschrift; opmerkelijk kort was de afstand tussen Amsterdam en Nijmegen (A1-A7).

MODELLING THE PERIPHERAL
AUDITORY SYSTEM
FITTING
MODEL PULSE PATTERNS TO NEURONAL IMPULSE TRAINS

ACADEMISCH PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR IN
DE WISKUNDE EN NATUURWETENSCHAPPEN AAN
DE UNIVERSITEIT VAN AMSTERDAM OP GEZAG VAN
DE RECTOR MAGNIFICUS DR. G. DEN BOEF, HOOG-
LERAAR IN DE FACULTEIT DER WISKUNDE EN
NATUURWETENSCHAPPEN, IN HET OPENBAAR TE
VERDEDIGEN IN DE AULA DER UNIVERSITEIT
(TIJDELIJK IN DE LUTHERSE KERK, INGANG SINGEL
411, HOEK SPUI) OP WOENSDAG 28 JUNI 1978 TE
16.00 UUR

DOOR

HERMAN ROBERT DE JONGH

GEBOREN TE ENSCHEDE

DRUKKERIJ ELINKWIJK BV - UTRECHT

- 1 -

CONTENTS

	page
Chapter I INTRODUCTION	2
Chapter II Boer, E. de and Jongh, H.R. de: On cochlear encoding: Potentialities and limitations of the reverse-correlation technique. J. Acoust. Soc. Am. 63(1), Jan. 1978	7
Chapter III MODELS OF THE AUDITORY PERIPHERY	28
III-1. Introduction	28
III-2. Duifhuis' model	28
III-3. Johannesma's model	31
III-4. Weiss' model	33
III-5. Discussion	35
Chapter IV MATHEMATICAL ANALYSIS OF THE MODELS	43
IV-1. Introduction	43
IV-2. The analysis of the models of Johannesma, Duifhuis and Siebert	43
IV-3. The analysis of Weiss' model	47
Chapter V EXPERIMENTAL DATA AND CHARACTERISTICS OF THE MODELS	57
V-1. Introduction	57
V-2. An experimental result	57
V-3. The experimental result and a fixed threshold	58
V-4. The experimental result and a fluctuating threshold	59
V-5. The experimental result and the non-Weiss models	63
V-6. An influence of $y'(t)$ on the firing probability	64
V-7. The revcor function for a generalised firing model	66
V-8. The experimental data and the generalised firing model	69
Chapter VI FINAL CONSIDERATIONS	74
VI-1. Introduction	74
VI-2. Possible causes of skewness	75
VI-3. The skewness effect and the models	83
VI-4. The Peripheral Auditory System model	86
VI-5. Concluding remark about the Volterra-Wiener approach	91
SUMMARY	92
REFERENCES	93

PROMOTOR: PROF. DR. E. DE BOER
COREFERENT: PROF. DR. F.A. MULLER

Academisch Ziekenhuis bij de Universiteit van Amsterdam
Wilhelmina Gasthuis
Keel-, Neus- en Oorheelkundige Kliniek (Prof. Dr. L.B.W. Jongkees)
Fysisch Laboratorium (Prof. Dr. E. de Boer)
Met steun van het Laboratorium voor Neurofysiologie (Prof. Dr. P.E. Voorhoeve)

ter I. INTRODUCTION

Sound, as we perceive it, is physically nothing but vibration of the air, of a certain kind, strong enough to be detected by the ear. This detection starts with relatively simple mechanical movements of the tympanic membrane. Following the path of the hearing centralward, via the middle ear, the inner ear, the several auditory nuclei to ultimately the auditory cortex, the complexity of neural processing increases explosively. In the central nervous system the underlying processes are electrochemical in nature; in the cochlea (or inner ear) the acoustical energy of the sound stimulus is transformed into electrochemical activity in neurones. In Fig. 1-1 we see the field of our interest: the peripheral auditory system.

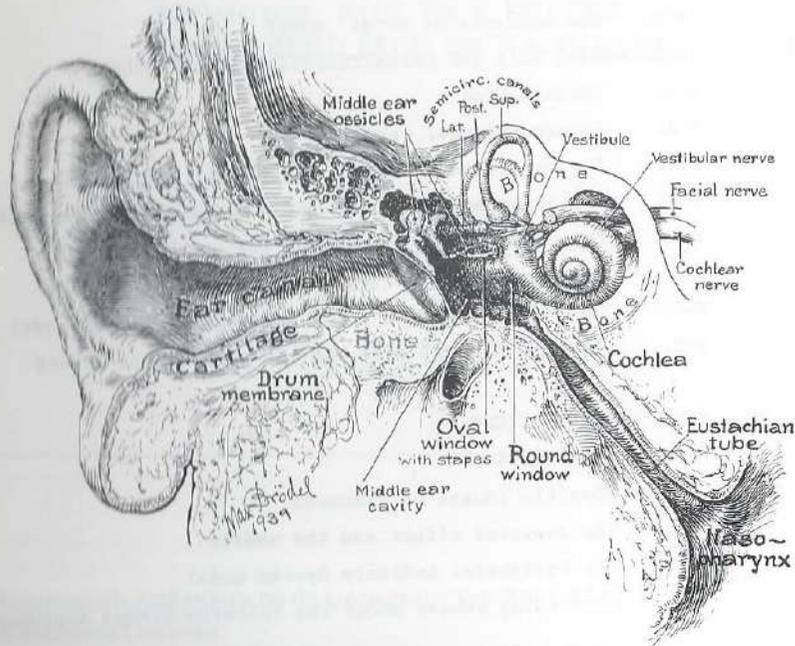


Fig. 1-1 Cross section of the human ear (reproduced from Dallos (1973))

It performs an encoding of sound into impulse patterns in the auditory nerve.

A short description of the peripheral auditory system follows:

Sound waves cause vibrations of the tympanic membrane which are transmitted to the cochlea via the middle ear ossicles. The middle ear takes care of the impedance match which is needed because vibrations in the air have to be transferred to vibrations in fluids, the peri- and endolymph within the cochlea. From our modelling point of view the middle ear is of minor importance. Its filtering action is readily absorbed in the much more drastic filtering which takes place in the inner ear. The transformation of mechanical movement to electrical activity, the so-called mechano-electrical transduction, is accomplished in the inner ear (see Fig. 1-2).

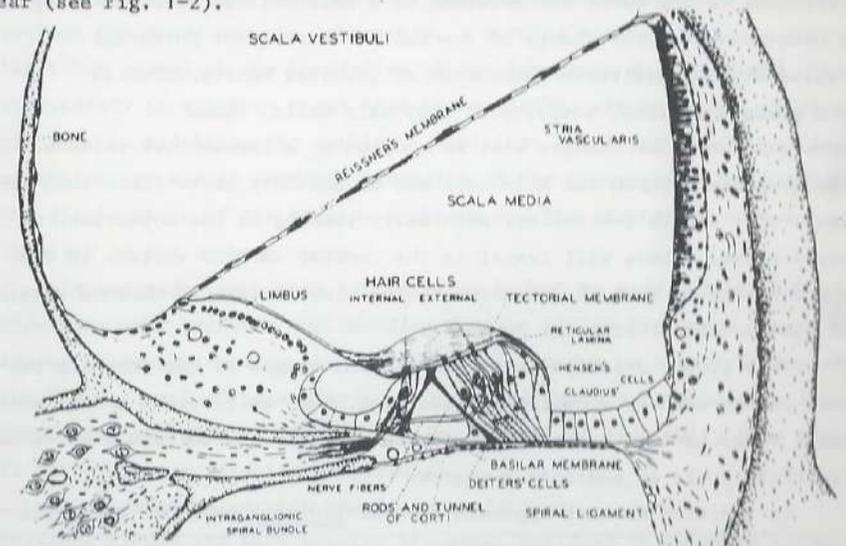


Fig. 1-2 Cross section of one turn of the cochlea. Prepared from a midmodiolar section of a cat's cochlea (reproduced from Dallos (1973)).

Movements of the fluids in the cochlea cause vibrations of the basilar membrane. The vibrations disclose a frequency selective action of the membrane such that different frequency components of the acoustic stimulus are more or less separated on it. It is on the basilar membrane that

the mechano-electrical transduction is accomplished by the auditory receptor cells, the hair cells. Mechanical excitation of a hair cell leads in an unknown way to excitation of the primary auditory neuron which innervates the hair cell. The primary auditory neuron conducts all or nothing events, impulses which can be picked up with the aid of a micro-electrode.

The ultimate mechanical movements in the inner ear or cochlea are extremely small in amplitude; the cochlear processes are so vulnerable and its structures so delicate that direct observation of cochlear functioning as a whole is still beyond all possibilities.

Notwithstanding the scarcity of direct evidence, many theories exist about the action of a hair cell (see e.g. Evans (1975)). Roughly the following is the case: The movement of a hair cell is accompanied by a membrane potential change of the cell: the receptor potential. Presumably the wellknown phenomenon of cochlear microphonics is the group electrical activity of many hair cells. These membrane potential changes will be coupled to a transmitter release. Via synapses between the hair cell and an auditory nerve fibre this release will excite the latter, eventually leading to the appearance of nerve spikes. These will travel to the central nervous system. In cat the activity in some 30.000 nerve fibres constitutes the encoded version of the acoustical stimulus to one ear.

This global description of the various stages of the encoding process can be used as a basis for modelling. Such quite crude models which exist today have interesting properties and they can certainly serve as fruitful tools in auditory research.

In addition to this approach several indirect physical and physiological methods for measurement of cochlear functioning exist which, in the last decades, yielded a wealth of experimental data. These are also of great use in studying the cochlear encoding process.

For several reasons it is important to understand the encoding process: apart from pure scientific curiosity and academic interest, insight is necessary whenever we attempt to help people suffering from loss of function of their peripheral hearing apparatus. If we want these patients to reestablish contact with the for them so dead world (Davis (1947)

page 395) , a profound understanding of the encoding process might assist the construction of an encoding apparatus that could deliver meaningful messages via an electrode array on their intact VIIIth nerve.

One way of gaining insight is to structure the many experimental data by means of models. Modelling can be done in several ways. In this context, a model is a more or less abstract, mathematical, description of the signal transformations which are involved. Such a model can be realised in software, as a computer program, or in hardware, as physical components lumped together. In the ideal case, a model should not only show the correct input-output relations but its steps should also, in details, correspond to processes in the actual course of the signal transformations. In this thesis, our modelling work will be concerned with both these aspects. It is hoped that in this way, via an interaction between computation and verification, between hypothesising and rejection, between the study and the laboratory, a working method is chosen that at least has gnawed at the boundaries of our ignorance in this field. The uncertainty is still so great that Evans (1975), after summing up several problems in modelling the cochlea, complains: "At least all these findings underline our almost complete ignorance of the processes lying between the displacement of the basilar membrane and the generation of nerve impulses in the cochlear nerve."

In our model work some restrictions had to be made. First, the modelling only concerns the encoding process of low frequency tuned fibres (frequencies less than 5 kHz.). This is due to a restriction of the "revcor method" (see chapter II). Second, mainly responses to stationary, wideband stimuli are treated for reasons of simplicity (see chapter II for different stimulus situations).

The structure of this thesis is as follows:

Chapters II and III both contain the above outlined interaction between theory and experimental outcome. Chapter II describes an analytical method. This method enables us to test a most simplified and condensed model via a "simulation method". The non-linearity of the encoding process gets much attention. Central in chapter II stands the "revcor function". (Looked upon in a more general way, the "revcor function" is proportional to the first-order kernel of the Volterra-Wiener expansion, see Hung and Stark (1977) for a review of the

Volterra-Wiener approach. This aspect is being dealt with in chapter VI). In chapter III three more specific and refined versions of the basic model from chapter II are treated and their properties, as far as relevant for our study, described. These properties are brought into connection with physiological phenomena known from literature, notably attention is given to analogue reasoning with the better known process of spike initiation in motoneurons and synaptical processes at the nerve-muscle transition. It turns out that the most fundamental difference between the three models resides in their tails: the firing models.

In chapter IV a mathematical analysis is given of the models of chapter III. Relations between the model's outcomes and the predictions via the simulation method are calculated. The results of this analysis enable us (chapter V) to interpret an experimental result obtained from the cat's peripheral auditory system. It is concluded that, without adjustment, the outcomes of all three models do not agree with the experimental result. The reason for this disagreement is, in the second part of chapter V, sought in the firing model. Then a generalised firing model is introduced, which is submitted to the same mathematical analysis as the other models. This procedure leads to new insight for the interpretation of the experimental result in terms of the firing model. In chapter VI the models of chapter III are reconsidered in the light of the outcomes of chapters IV and V.

Attempts to adapt the models in order to remove the disagreement found in chapter V lead, in a quite natural way to a reappraisal of Weiss' firing model. The chapter ends with some questions that remain.

On cochlear encoding: Potentialities and limitations of the reverse-correlation technique

E. de Boer and H. R. de Jongh

Physics Laboratory ENT Department (KNO) Wilhelmina Hospital Amsterdam, The Netherlands
(Received 29 November 1976; revised 27 June 1977)

This paper presents a description of the interrelation between two major properties of the responses recordable from auditory nerve fibers: *frequency selectivity* and *partial synchrony* between stimulus and response. In the course of this work the influence of nonlinearity on the cochlear encoding process can be assessed. The theory of the *reverse-correlation technique* is derived in a most general way. It is based on a model in which a filter—assumed to be linear—is followed by a stochastic pulse generator—the probability of producing an output pulse being an instantaneous but nonlinear function of its input signal. Insofar as such a model represents stimulus transformations in a primary auditory neuron, the technique can be applied to the responses recorded from an auditory nerve fiber. Several illustrative examples of experimental reverse-correlation functions—abbreviated: *revcor functions*—are presented and discussed. These functions have the general character of impulse responses of sharp bandpass filters. They show very little phase modulation. For noise stimuli of up to 70 dB per third octave the revcor functions are almost invariant. Above that level some (but not all) of the revcor functions show a loss of frequency selectivity. If a nerve fiber can be contacted for a sufficiently long time, it is possible to compare the response with that of a model filter, in which the revcor function of that fiber is substituted as its impulse response. The output signal of the model filter is shown to be a very good predictor of the firing probability of the fiber under study. This property is demonstrated for noise as well as for tone stimuli. There is surprisingly little evidence of nonlinear filtering in these results. This so-called simulation method can also be applied when the stimulus is switched on and off. The results show, apart from effects due to filtering, clear manifestations of fast adaptation. Again, the filtering appears to be independent of the latter effect. It is concluded that for wide-band noise and single-tone signals the firing probability is predominantly controlled by a linearly filtered version of the acoustical stimulus; this constitutes the principle of *specific coding*. The conspicuous absence of nonlinear effects in the results can partly be explained in terms of the response properties of a class of networks in which sharp filtering occurs after the generation of nonlinear distortion products. It can then be predicted that this property will hold only for wide-band and tonal stimuli. That our results show so little evidence of cochlear distortion appears to be a property of signal transformations and is not due to linearization tendencies of the experimental method.

PACS numbers: 43.63.Pd, 43.63.Bq, 43.63.Nc, 43.63.Rf

INTRODUCTION

In the past decades many studies have been dedicated to the response properties of single auditory-nerve fibers (primary auditory neurons¹). In this paper we will describe the relation between two of the main properties, namely frequency selectivity and time locking (synchrony), as revealed by the application of modern on-line computing techniques. The results described in the paper extend the concept of partial synchrony between stimulus and response to include stimulation with stochastic and with nonstationary signals: noise and tonebursts. It is shown that for each neuron a linear transform of the stimulus can be constructed in such a way that the nerve fiber's firings show a great deal of synchrony with it. For the type of stimuli used, this confirms the validity of a model of a primary auditory neuron which contains a linear filter as the only frequency-selective part. The model can be used as a functional model: i.e., it can predict input-output relations in detail (input being the acoustical input to the ear and output the train of action potentials in a nerve fiber). A large number of such models, each representing a group of primary auditory neurons, can serve as a simplified representation of the cochlear encoding process.

The frequency selectivity that responses of auditory-

nerve fibers exhibit is usually described in terms of tuning curves, determined for pure-tone stimuli. (See e.g., Kiang *et al.*, 1965; Evans, 1972.) A tuning curve shows the boundary along which a certain increase in spike rate due to the presence of the stimuli is observed. As a result of several nonlinear phenomena the curves determined for different increases in spike rate above the spontaneous rate are not parallel (e.g., Rose *et al.*, 1967).

A second important property of auditory-nerve fibers is the partial synchrony of the firings with the waveform of the stimulus (Hind *et al.*, 1967; Anderson *et al.*, 1971; de Jongh, 1972). There is a limit to the synchrony between stimulus and response: For stimuli above 6 kHz the effect can no longer be demonstrated. This limiting behavior does not seem to be due to errors or inaccuracies in the measurement of the epochs of nerve-fiber action potentials, but to an intrinsic timing limitation in the nervous excitation process.

The third property of cochlear transduction which is reflected by the response pattern of auditory-nerve fibers is nonlinearity (Sachs and Kiang, 1968; Sachs, 1969; Goldstein and Kiang, 1968; Gobleck and Pfeiffer, 1969). See for a review paper Pfeiffer and Kim, 1973.

To these nonlinear phenomena we may add properties

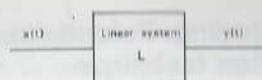


FIG. 1. Linear system driven by a stochastic signal.

pure tones and it cannot be expected to hold true for signals with steep slopes in their spectrum (single-sideband signals and tone complexes).

For stimuli with an abrupt onset the picture is complicated by adaptation. The influence of this effect can easily be separated from the filtering process and again there appears very little evidence of nonlinear filtering. In other words, nonlinear filtering or time-varying effects do not substantially upset detailed temporal relations between the stimulus signal and the nerve fiber's firing probability. This holds true for stationary as well as transient wide-band stimulus signals and for tonebursts.

In Sec. IV several specific topics are discussed regarding the use of the revercor function and the consequences of the results. It is concluded that, as far as the prediction of firing patterns is concerned, the model with a linear filter is sufficient for describing cochlear encoding for most if not all signals occurring in everyday life. A tentative explanation is offered for the observed near-invariance of revercor functions for variations in stimulus intensity. Furthermore, the conspicuous absence of nonlinear filtering effects is discussed in terms of the BPNL network (de Boer, 1976a). As far as cochlear filtering is concerned, nonlinear effects appear to be of a type that leaves hardly a discernible trace in the processing of wide-band signals. This is consistent with the idea that nonlinear effects occur on a "wide-band" basis and that the distortion products generated are affected by sharp bandpass filtering. However, this is not the only possible explanation.

1. REVCOR FUNCTIONS; THEORY AND EXPERIMENTS

We first derive the basic theory behind the reverse correlation technique. Note that the present derivation differs from the one used when the method was originally derived (de Boer and Kuypers, 1968). Consider a linear system L (Fig. 1) with a stochastic signal $x(t)$ at its input terminals, and let $y(t)$ be the output signal. The cross-correlation function $\varphi_{xy}(\tau)$ is defined as the time (or ensemble) average of the product $x(t+\tau) \times y(t)$:

$$\varphi_{xy}(\tau) = \overline{x(t+\tau) \times y(t)}, \quad (1)$$

where the bar indicates the averaging operation. For the system of Fig. 1 the following relation holds:

$$\varphi_{xy}(\tau) = \int_0^{\infty} \varphi_{xx}(\tau + \sigma) \times h(\sigma) d\sigma, \quad (2a)$$

where $\varphi_{xx}(\tau)$ is the auto-correlation function of $x(t)$ and $h(t)$ is the impulse response of the linear system L . If white noise is chosen for $x(t)$, the relation becomes particularly simple

$$\varphi_{xy}(\tau) = h(-\tau). \quad (2b)$$

The system of Fig. 2 is more complex; it can be used as a model for a primary auditory neuron¹ to describe the transformation of the acoustic stimulus signal $x(t)$ into a train $s(t)$ of impulses recordable from the pertinent nerve fiber (cf. Siebert, 1968; Dufhuis, 1972). The probability of firing $p(t)$ of the pulse generator PG is assumed to be a no-memory nonlinear transform of the signal $y(t)$ produced by L . For this system the cross-correlation function $\varphi_{xs}(\tau)$ can conveniently be defined as a time average

$$\varphi_{xs}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t+\tau) s(t) dt, \quad (3)$$

If $s(t)$ consists of a number of δ functions

$$s(t) = \sum_{i=1}^{N_T} \delta(t - t_i) \quad (4)$$

this can be reduced to

$$\varphi_{xs}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=1}^{N_T} x(t_i + \tau), \quad (5)$$

where t_i ($i = 1, \dots, N_T$) denote the N_T instants at which nerve impulses are detected in the time interval T .

For the purpose of analysis, the system of Fig. 2 can be redrawn as in Fig. 3. Here the pulse generator is linear: The probability of firing $p(t)$ is proportional to its input signal $x(t)$. All nonlinearity is lumped into the no-memory nonlinear element NL which, in particular, ensures that the signal $x(t)$ never becomes negative. It is easy to understand that the cross-correlation function $\varphi_{xs}(\tau)$ is proportional to $\varphi_{xy}(\tau)$. According to Price's theorem (Price, 1958) the latter function is proportional to $\varphi_{xx}(\tau)$, at least when $x(t)$ is Gaussian— $y(t)$ is then Gaussian too. Hence

$$\varphi_{xs}(\tau) = C \times \varphi_{xy}(\tau). \quad (6)$$

In terms of Eqs. (2) this means that for a composite system as depicted by Fig. 2, the waveshape of the linear system's impulse response $h(t)$ can be recovered from a correlation measurement on stochastic input and output signals. If $x(t)$ is a white-noise signal, the relation becomes

$$h(\tau) = C^{-1} \times \varphi_{xs}(-\tau). \quad (7)$$

The function $\varphi_{xs}(-\tau)$, obtained with white-noise input, is the *reverse-correlation function*, abbreviated: *revcor function*.

Let us assume that the model of Fig. 2 represents a primary auditory neuron. Then the same analysis can be applied to measure the properties of the linear filtering stage. The appropriate experiment involves simultaneous observation of the analog input signal (preferably

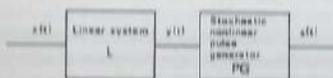
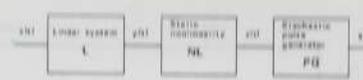
FIG. 2. Nonlinear frequency-selective pulse generating system, can be used as a model of a primary auditory neuron.¹

FIG. 3. The network of Fig. 2 shown as a cascade of simpler devices.

white noise) and the output signal, i.e., the series of action potentials ("events") recorded from a single auditory nerve fiber. The computation should be triggered by the fast initial edges of the action potentials, and proceed according to Eq. (5). Since this equation contains only a sum of parts of the $x(t)$ -signal—taken at and around (or, prior to) the instants t_i at which $s(t)$ contains an impulse—the actual computation of the cross-correlation function is very simple indeed.

The derivation presented above provides the general theory behind the method of reverse correlation as applied to the study of the cochlear encoding process. It should be noted that the method does not "exist" without a model: i.e., its result can only be expressed in terms of that model. In other words, the result of the experimental procedure, which comes in the form of an (unbiased) estimate of the revercor function, is nothing more than a parameter of the underlying model (Fig. 2). If the appropriate parameter is substituted, the model should provide an optimal match to the experimental observations. The validity of the model in a more general sense must be justified by an independent test. Such a test is described in Sec. II.

We now present some typical results of the application of the reverse-correlation procedure to responses recorded from auditory-nerve fibers. The experimental setup is schematized in Fig. 4. Anesthetized cats are used as experimental animals. The stimulus signal—white noise— $x(t)$ is presented by an earphone via a closed acoustic system. The (electrical) signal $x(t)$ is fed also to the computer (DEC PDP-9). From this moment on, the signal $x(t)$ assumes the role of the stimulus signal. Thus, the electroacoustical transfer function is absorbed in the linear system L of the model. Action potentials recorded from an auditory-nerve fiber are converted into pulses and sent to the computer via a digital channel. The procedure outlined in the previous paragraphs calls for processing of the stimulus signal $x(t)$ as guided by the occurrence of action potentials recorded. Since only negative values of τ are involved [cf. Eq. (2b)] the processing involves addition of stimulus signal fragments that occur prior to each action potential. To achieve this, the computer continuously reads in the x signal via its analog-to-digital system, so that the immediate past of $x(t)$ (over, e.g., 50 ms) is available every time an action potential is detected. It is to be noted that in an actual experiment only an estimate of the revercor function can be obtained, due to the limited time of observation. Appendix A gives more details about experimental and computational procedures, calibration, polarity conventions, etc.

Figures 5-8 show experimental results for four primary auditory neurons measured in data. Figures 5(a), 6(a), 7(a), and 8(a) show the waveshapes of the revercor func-

that seem to be mainly of a typical neural character, such as adaptation (e.g., Smith, 1973), saturation (Kiang *et al.*, 1965), and the depletion effect described by Gray (1966). This is certainly a bewildering situation, and we may well ask what type of simple model will describe the encoding for everyday signals such as animal cries and human speech.

In this paper it will be shown to what extent the properties frequency selectivity and synchrony can be tied together by a *linear* type of model. For linear systems it is possible to extract the parameters by operations on stochastic input and output signals. Such a technique can also be used to extract the dynamic response characteristics from a stochastic, frequency-selective system that accepts an analog signal as input signal and produces a series of impulses at the output—the technique of triggered correlation (de Boer and Kuypers, 1968). The application of this technique to cochlear physiology is known as reverse correlation (de Boer, 1967, 1968, 1969a, 1969b; de Boer and Jongkees, 1968). The method involves stimulation of the ear of an experimental animal (cat) with white noise, recording of the train of action potentials from a single fiber of the auditory nerve, and computation of a specific type of cross-correlation function. For each fiber the result comes in the form of a time function, which can be associated with the nerve fiber studied. That function is known as the "revercor function" (reverse-correlation function) of the fiber. The Fourier transform of the revercor function has the character of a bandpass-filter frequency response; the corresponding frequency response function has extremely steep slopes and, in fact, is very similar to the pure-tone tuning curve (de Boer, 1973). Section I of the present paper presents the derivation of the theory in a most general way and reports the waveshapes of typical revercor functions for a number of primary auditory neurons.

In view of the many and strong nonlinearities in the cochlea it seems at first sight not fruitful to pursue a linear analysis too far. It is shown in Secs. II and III, however, how well the revercor function of a nerve fiber can be used to predict detailed temporal properties of the fiber's response to wide-band signals and tonebursts. For this purpose the response of the actual physiological preparation is compared with that of a model in which a linear filter is the only frequency-selective element. The revercor function is substituted as the impulse response of this filter. It is found that the firing probability of the neuron is nearly proportional to the rectified output of the filter, this holds true for neurons with resonance frequencies (CF's) of 2 kHz and lower. For neurons with higher resonance frequencies the firing probability behaves as a linear (low-pass) transform of the rectified output of the model filter. The success of this so-called "stimulation procedure" can be interpreted to mean that the neural transducer mainly reacts to a linearly filtered version of the stimulus, the so-called principle of specific coding (de Boer, 1973), and that in the cochlear filtering process, nonlinear effects are not very important. This conclusion is subject to the restriction of wide-band stimulus signals and single

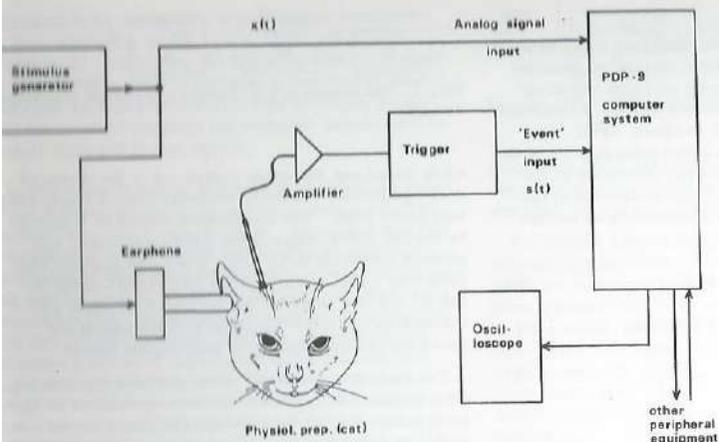


FIG. 4. Setup to measure a reverberation function.

tions as they are obtained after the processing of more than 4000 action potentials. The functions are measured with white-noise stimuli presented at a level of 20–40 dB above the fiber's threshold, see the legends to the figures. Threshold is defined here as the acoustical level at which the presentation of a 50-ms burst of noise produces a just discernible increase in firing rate. The reverberation functions show several oscillations with a fairly rapid build up and a tail of slowly decaying amplitude. The frequency of these oscillations corresponds to the characteristic frequency of the fiber. Hence we associate a resonance frequency with each reverberation function, and we may identify this with the characteristic frequency (CF) of the fiber. The waveshapes reported in these figures are representative of the hundreds of reverberation functions we obtained from cat experiments, with resonance frequencies ranging from 0.34 to 5 kHz. In general the relative bandwidth decreases with increasing resonance frequency. The variations are quite large, however; the neuron shown by Fig. 7, for instance, has a larger-than-average bandwidth.

All reverberation functions show an initial delay; this delay not only reflects propagation time in the cochlea but also includes the buildup time of the action potential from its locus of initiation to the recording electrode. In the model of Fig. 3 the neural pulse generator is assumed to be an instantaneous device. Hence all delays show up as an initial delay in the recovered impulse response $h(t)$ for L . As mentioned before, the transmission characteristics of the electroacoustic system are absorbed in L too. As long as the same system is used, no correction is necessary for the comparison of responses to different stimuli.

Figures 5(b), 5(c), 6(b), 6(c), 7(b), 7(c), 8(b), and 8(c) show the "reverberation spectra," i.e., the Fourier transforms of the reverberation functions. The amplitude spectra [5(b), 6(b), 7(b), and 8(b)] are plotted on logarithmic scales and upside down so as to stress the resemblance to pure-tone tuning curves. The spectra are cut off at

a distance of up to 40 dB from the top since components further away from the top are due to measurement inaccuracy. When plotted on logarithmic coordinates, these amplitude spectra are somewhat asymmetrical around their resonance frequencies. On a linear frequency scale, however, they appear as nearly symmetrical. The amplitude spectra of reverberation functions have been reported to be quite similar to pure-tone tuning curves (de Boer, 1969a, 1973) in the region around the resonance frequency, albeit that the tuning curves may be somewhat steeper.

Figures 5(c), 6(c), 7(c), and 8(c) show the phase responses associated with the spectra. Two points are of interest here. First, the reverberation function shows an initial delay and, as a consequence, the phase changes fairly rapidly as a function of frequency. To improve clarity, each phase value at a particular frequency is plotted as the difference of the actual phase and the phase due to a pure delay at the same frequency. The delay used for this correction is indicated in the figure. For reference purposes, the pure-delay phase is also shown (dotted). The second point is that the spectrum cannot be estimated accurately down to zero frequency and hence it is not possible to find out how many times a phase of 2π is accumulated. For this reason the phase at the resonance frequency is plotted as its true value between $-\pi$ and $+\pi$, and all phase values are plotted with respect to this reference point. The correction phase is taken as zero at this point. We observe that the phase responses are fairly regular throughout the resonance region.

The main source of error in the reverse-correlation procedure is finite sample size: Several thousands of action potentials are processed for each reverberation function shown, yet random fluctuations are clearly visible. Under the assumption that errors are independent and additive, the experimental functions can be smoothed by removing from the waveform or the Fourier spectrum those components that appear to be entirely random.

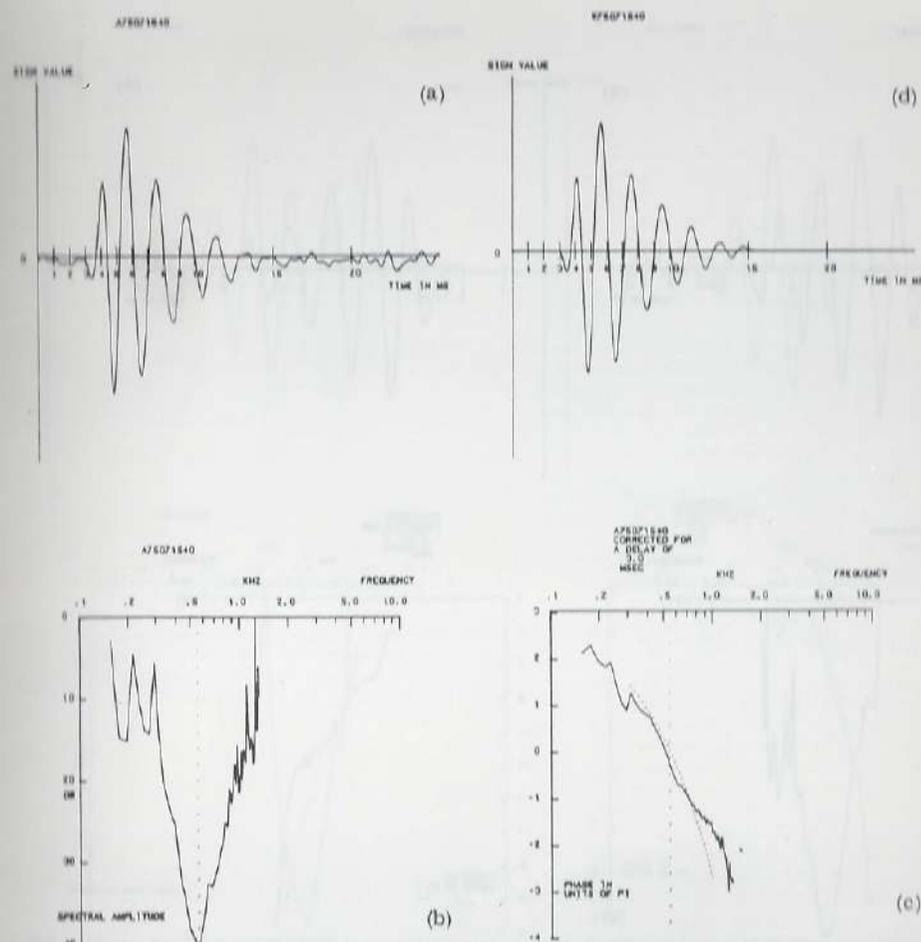


FIG. 5. Reverberation function of unit 75-07-15, resonance frequency: 530 Hz, threshold for noise: 26 dB (SPL)/third oct., stimulation level: 40 dB/third oct. (a) Waveform of reverberation function (vertical scale arbitrary). (b) Spectral magnitude (inverted). (c) Phase characteristic (see text). (d) Smoothed waveform.

The smoothing procedure is alternately carried out in the frequency and the time domain. Figures 5(a), 5(b), 7(a), and 8(a) of the figures show the raw data, Figs. 5(d), 6(d), 7(d), and 8(d) the smoothed reverberation functions. Most of the characteristic features of the functions are preserved. The smoothed versions of the reverberation functions are convenient for further analysis of stimulus-response relationships, see Sec. II.

In view of the linearizing tendencies of correlation methods, separate tests are necessary to assess the part played by cochlear nonlinearity. One impression of nonlinearity can be obtained by measuring via the reverse-correlation technique whether frequency selec-

tivity of a neuron is dependent upon stimulus intensity. For low-to-moderate levels of stimulation, the reverberation functions are found to be almost invariant. The largest variations occur in the phase response, they appear as a uniform phase shift limited to $\pi/4$ rad. At levels of over 70 dB SPL per $\frac{1}{2}$ octave (83 dB SPL overall level for a bandwidth of 5 kHz) some reverberation functions change their appearance. In terms of the waveform an increased damping is evident. In terms of the spectrum, the relative bandwidth becomes broader with increasing level and the top shifts to a lower frequency. This behavior is reminiscent of the nonlinearity in mechanical tuning that was reported by Rhode (1971). We did not find this behavior, however, in all fibers tested at high

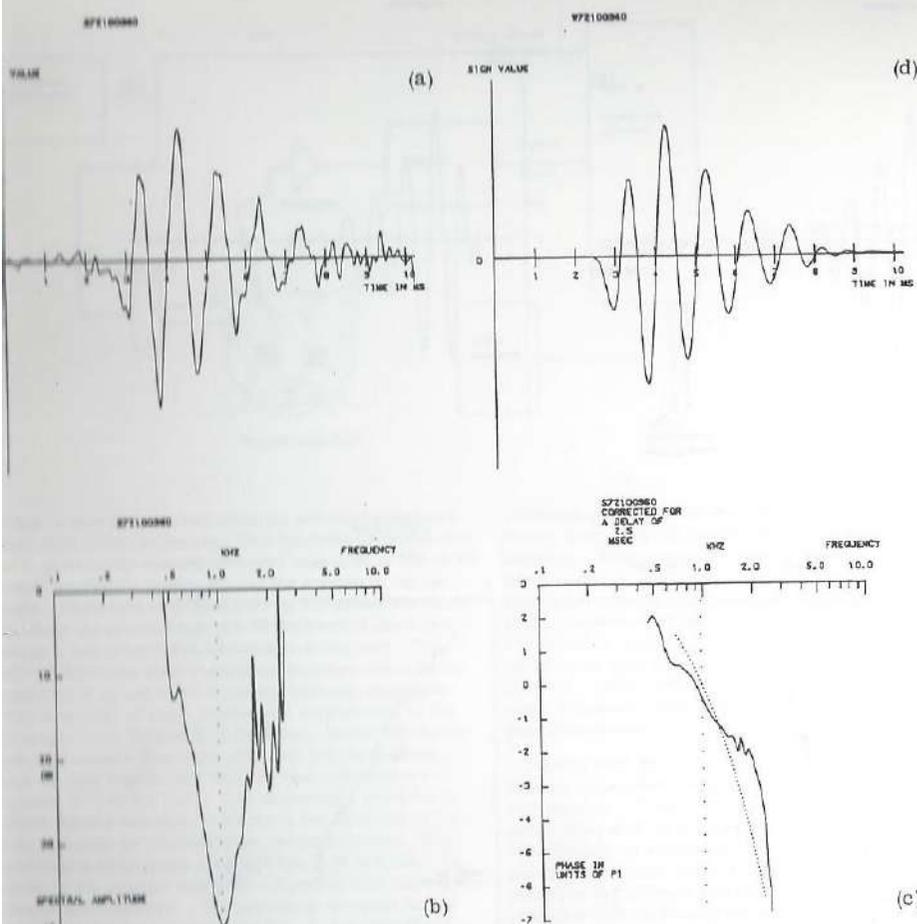


FIG. 6. Reverb function of unit 72-10-09, resonance frequency: 1 kHz, threshold for noise: 30 dB/third oct., stimulation level: 60 dB/third oct.

intensities. Comparable results on properties of reverb functions were recently reported by Evans (1976, 1977). A possible explanation for these findings will be described in the discussion. Further experiments that give an indication as to how important cochlear nonlinearity is in signal processing are presented in the next section.

H. VERIFICATION (SIMULATION PROCEDURE)

The experiments described in the present section are designed to test the predictive power of the model underlying the reverse-correlation procedure. The principal idea is to measure the reverb function for a primary auditory neuron, to substitute this function as the impulse response in the model of Fig. 2 and to compare the timing pattern of actual nerve firings with predic-

tions from the model. If the model holds true, the firing probability $p(t)$ of the neuron should be a monotonic function of the output $y(t)$ of the linear filter L . We will presently see to what extent $y(t)$ is a good predictor of $p(t)$. The procedure is called the *simulation procedure* (cf. de Boer and de Jongh, 1971; de Jongh, 1972, 1973; Grashuis, 1974). In the present report the emphasis will be on the accuracy of timing prediction and the evidence of specific nonlinear effects.

Figure 9 shows a flow diagram of the simulation procedure. First, the reverb function of the nerve fiber under study is determined, using white noise as the stimulus [Fig. 9(a)]. In smoothed form this reverb function—to be called $h^*(t)$ —is substituted as the impulse response of the model filter L^* (an asterisk denoting

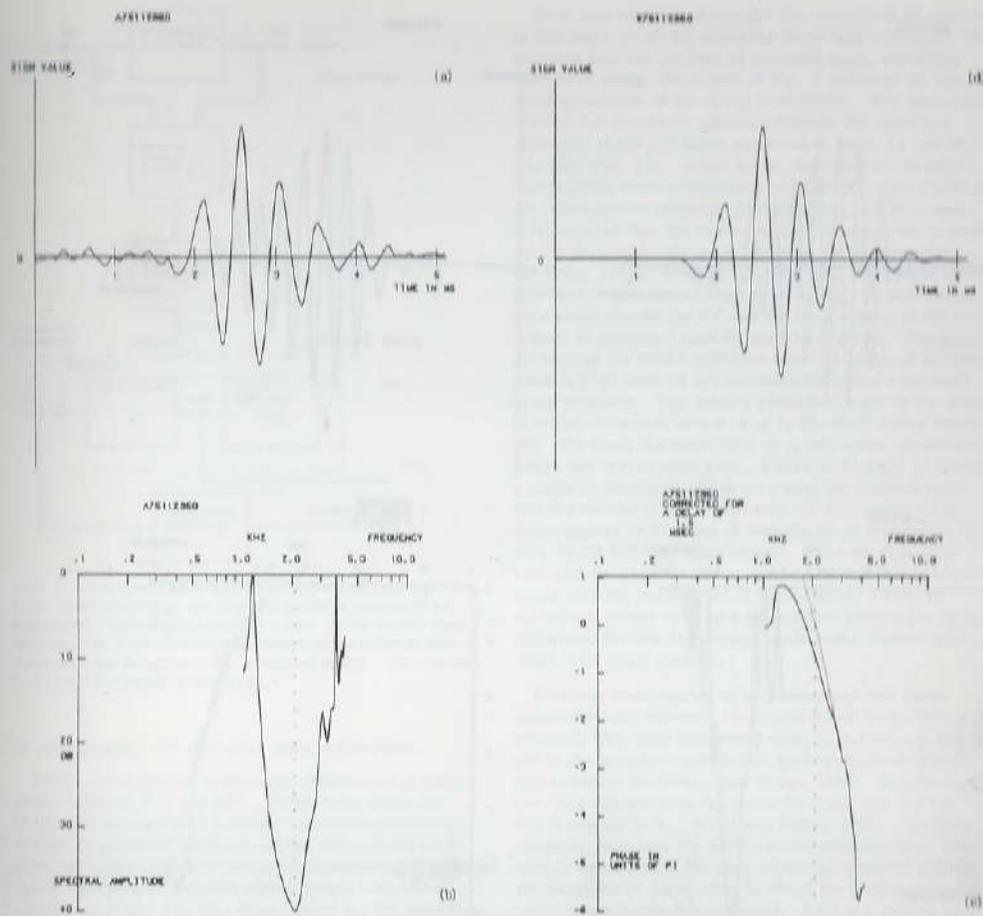


FIG. 7. Reverb function of unit 75-11-29, resonance frequency: 2.1 kHz, threshold for noise: 30 dB/third oct., stimulation level: 50 dB/third oct.

The fact that the quantities and concepts involved are based on a statistical estimate of the reverb function). In the second stage [Fig. 9(b)] a new stimulus signal $x(t)$ is generated which is periodic with the period T ; this signal $x(t)$ is led to the animal's ear as well as to the input of the filter L^* . The firing probability $p(t)$ of the (same) nerve fiber is determined from the physiological preparation in the form of a PST histogram of the nerve firings over the interval T . The response $y^*(t)$ of the filter L^* is computed according to the well-known convolution theorem. The signal $y(t)$ —of which we can only measure its estimate $y^*(t)$ —is called the excitatory signal. It should be noted that the excitatory signal is used in the physical (and not in the physiological) sense: it is just a filtered form of the stimulus $x(t)$ and nothing else.

The first case to be described is that of stimulation with repetitive noise (pseudorandom noise). After the reverb function of a primary neuron was determined with truly random noise, the same neuron was stimulated continuously by presenting pseudorandom periodic Gaussian noise generated by the Hewlett Packard noise generator type 3722 A. The bandwidth of the noise was 5 kHz and the period of repetition 0.082 s ($N=13$). The uppermost panel of Fig. 10(a) shows the noise waveform. The PDP-9 computer was programmed to accumulate a PST histogram of the nerve firings with the onset of the noise repetitions as the zero point for time. The lowest panels of Fig. 10(a) show parts of the measured PST histogram, i.e., the function $p(t)$, together with the computed waveform of the filter output $y^*(t)$. For the computation of $y^*(t)$ a smoothed version $h^*(t)$ of

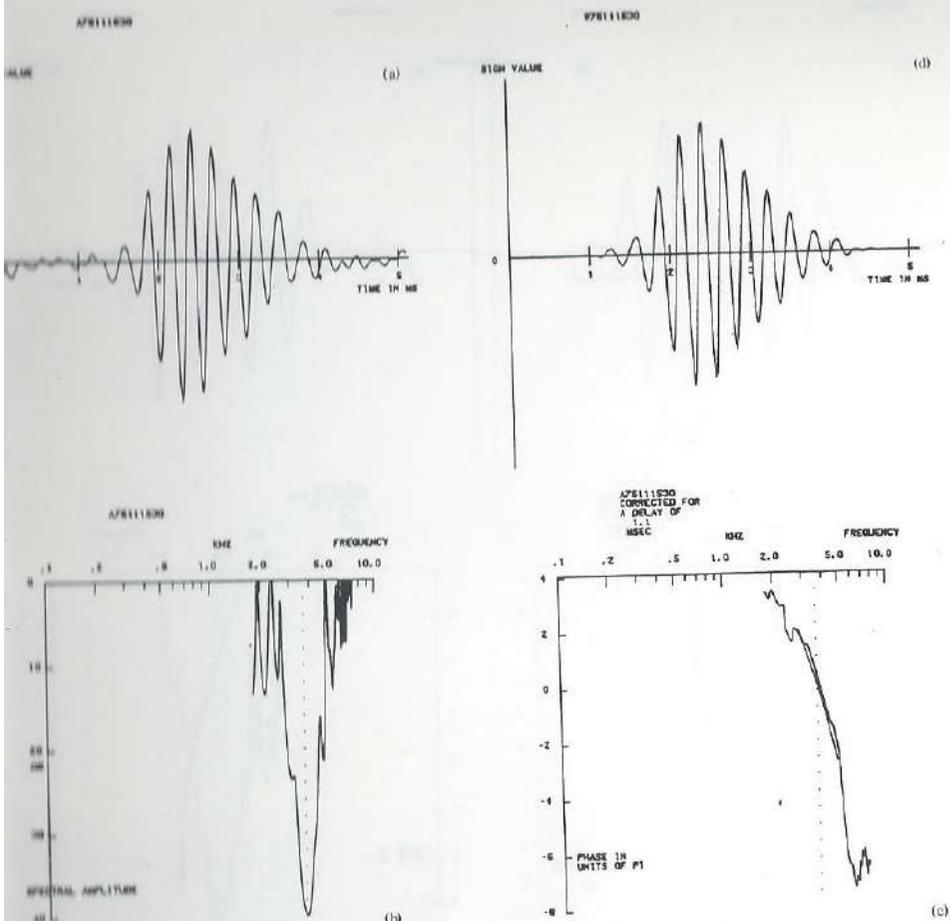


Fig. 8. Reverse function of unit 75-11-15, resonance frequency: 4.66 kHz, threshold for noise: 10 dB/third oct, stimulation (a) 30 dB/third oct.

reverse function was used, see the inset of this figure. The amplitude scales for $y^*(t)$ and $p(t)$ are arbitrary.

The neuron under study (75-08-73) had a resonance frequency (CF) of 800 Hz. It is evident that $y^*(t)$ predicts the time course of the probability of firing $p(t)$ to large degree. For positive values of $y^*(t)$, the average relation between $p(t)$ and $y^*(t)$ appears to be nearly linear. For negative values of $y^*(t)$, the probability of firing is practically zero. Figures 10(b) and 10(c) show other examples of measured $p(t)$ and computed $y^*(t)$ signals for two other neurons with low resonance frequencies, 1.2 and 2.1 kHz, respectively. The stimulus was the same repetitive noise signal. In Fig. 10(b) we note that the largest positive values of $y^*(t)$ do not give

rise to a fully proportional $p(t)$. In Fig. 10(c) we note just the opposite: The smaller excursions of $y^*(t)$ seem to produce a disproportionately small number of spikes. It is as if the neuron has an "instantaneous threshold," i.e., when $y^*(t)$ is below a certain (positive) level, the neuron is not likely to fire. Still, the actual threshold of this neuron is lower, see Fig. 7. We may conclude that the relation between $y^*(t)$ and $p(t)$ appears to be nearly monotonic but its character varies from neuron to neuron. There is another interesting feature in Fig. 10(c): In the smaller excursions of $y^*(t)$ we observe a small shift in timing of the firings. It is as if a certain phase modulation is present in these periods. It is probable that errors in the estimation of the reverse function $h^*(t)$ produce time shifts of $y^*(t)$; it is also probable, however, that this effect is a manifestation

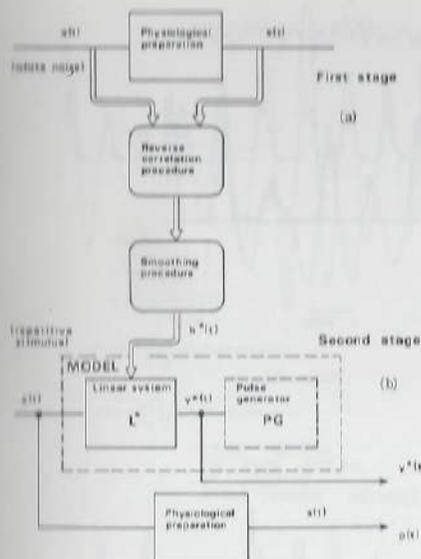


FIG. 9. Illustrating the stages of the simulation method. See text. In the first stage the reverse function is measured and processed. The stimulus is white noise. In the second stage the response of the (same) physiological preparation is compared with the response of the linearized model. The stimulus is a repetitive signal in this stage.

of nonlinearity. We will come back to this later.

There are small but systematic differences in wave-shape between $y^*(t)$ and $p(t)$. These differences are larger for neurons with a higher resonance frequency. Figure 11 presents the result of the simulation procedure for a neuron with a resonance frequency of 2.6 kHz. Again the firing probability $p(t)$ is nearly a rectified version of $y^*(t)$. For this fiber, however, the individual lobes of the $p(t)$ function are asymmetrical. Their left-hand flanks are steep, and their right-hand flanks are shallower and have a tendency for "crossing over" to the next lobe. This is not a result of processing or measurement errors but, quite likely, reflects intermediate steps in the initiation of nerve impulses that are not accounted for by the simple model of Fig. 2 and that show up more clearly when the resonance frequency is high. For primary auditory neurons with still higher resonance frequencies this tendency is more pronounced and the firing probability ceases to follow every lobe of $y^*(t)$. In a certain sense, the firing probability tends to be controlled by the envelope of the $y^*(t)$ signal. Figure 12 shows the result for a neuron with a resonance frequency of 4.7 kHz, an extreme case that serves well to illustrate this point. Note the enlarged time scale in this figure. We observe here a manifestation of the same cause that prevents us from measuring reverse functions in neurons with resonance frequencies above 6 kHz.

Up to now we have processed the records of 31 neurons in this way. In all we found the same type of result. We conclude that for the type of stimulus used, repetitive wide-band noise, the model of Fig. 2 produces an adequate prediction of the firing probability. The main correction that should be applied concerns the observed skewness of the $p(t)$ lobes as shown in Figs. 11 and 12 (cf. also Fig. 14). In any event, the reverse correlation function (revcor function) provides the main parameter, the impulse response function $h^*(t)$, of the model. It is recalled that the revcor amplitude spectrum around its top approximately coincides with the tuning curve (de Boer, 1973). It appears, then, that the revcor function describes several aspects of tuning: the tuning for pure tones around the CF and the time course of the response to (pseudo-) random wide-band noise. The predictions of the model hold true over the range of fluctuations in $y^*(t)$ such as are encountered with a stationary noise stimulus. The results presented apply to the case of stimulation with levels of up to 30-40 dB above threshold. We found the same type of result under conditions where the neuron saturates. Within each lobe of $y^*(t)$ only a slight flattening of $p(t)$ occurs near its maxima [note that the neuron of Fig. 10(b) was not saturated]. The same applies to the case of stimulation at intensities of over 70 dB SPL per third octave. Here we obtain the best results when we use a reverse function that is measured with the same level of stimulation. These remarks are meant to be of a descriptive nature; we have processed too few recordings made under these conditions to be more specific.

Cochlear nonlinearity is so pronounced that under specific circumstances, responses occur to distortion products (the cubic difference tone, in particular) and *not* to the primary components generating these distortion products (Goldstein and Kiang, 1968). Similarly, two-tone suppression can occur by tones that are below threshold (e.g., Sachs and Kiang, 1968). In these respects, nonlinearity seems to override cochlear frequency selectivity. We may expect components outside the passband of the system to affect the transmission of components inside the passband. This is a reason to pay attention to effects that betray the influence of nonlinear filtering. Those effects occur in the phase and amplitude modulations that are inherent in narrow-band noise signals (de Boer, 1976b). With this in mind, we turn again to Figs. 10-12 and observe that there are no obvious phase deviations between $y^*(t)$ and $p(t)$ [except for the slight effect in Fig. 10(c)]. In general terms, nonlinear filtering in the cochlea, however strong it may be, seems to leave no trace in the response to wide-band stimuli such as pseudorandom noise. See the next section for a further analysis.

All this applies to stationary stimuli. The next question is: What happens in the case of nonstationary stimuli? For the results shown by Fig. 13 a toneburst of 50-ms duration was presented periodically with a period of 0.1 s. The tone was switched on and off with zero rise time (the uppermost trace shows the stimulus waveform). It is seen from the central trace that the preferred cycles of firing as well as the cycles of suppression are cor-

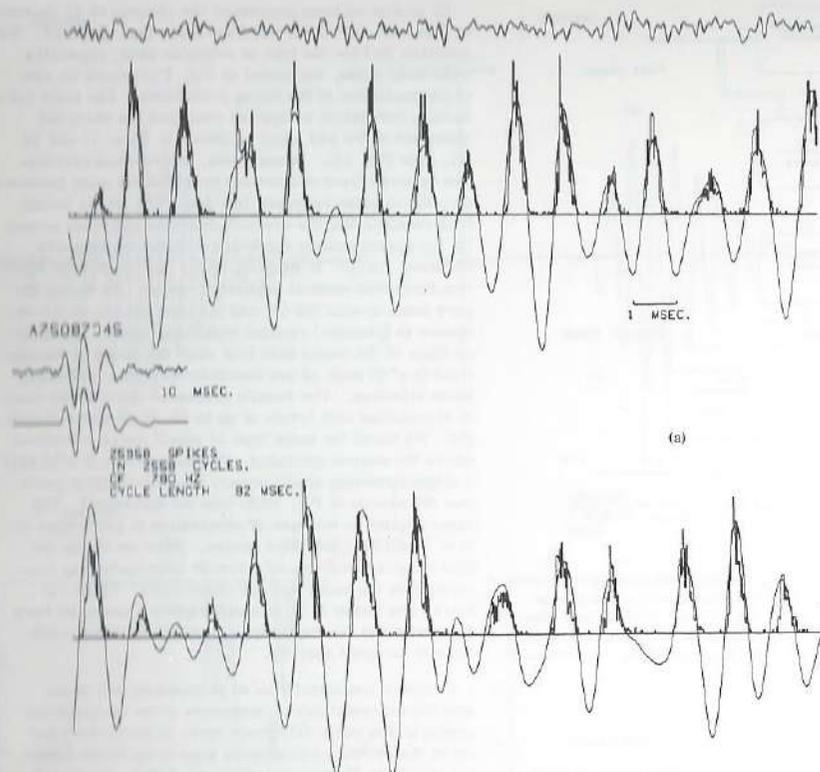


FIG. 10(a). Results of the simulation method for unit nr.: 75-08-73, resonance frequency: 0.780 kHz, stimulation level: 45 dB/4 oct. The upper trace shows the stimulus waveform $x(t)$; the other traces show the computed $y^*(t)$ signal (smooth line) and the measured $p(t)$ function (as a histogram). The lower panel shows a different part of the recording. The first shows the revcor function, in both raw and smoothed forms. The latter version is used as $h^*(t)$ to compute $y^*(t)$. The unit, as are those in Figs. 10(b) and 10(c), is stimulated at 30 dB above its threshold for noise.

rectly predicted by the (linear) model. Even during the buildup phase of the $y^*(t)$ response, there is no trace of a decrease or a change of synchronization. However, in the first few cycles of $y^*(t)$ the firing probability $p(t)$ is far larger than in later cycles; this is the manifestation of adaptation as it occurs during the onset phase. The lowest panel of this figure shows the response to the sudden cessation of the stimulus. Again it is seen that the linear response $y^*(t)$ correctly predicts the lobes of the histogram but in this case the amplitudes of the $p(t)$ lobes decay more slowly than those of $y^*(t)$ —this is probably another manifestation of the adaptation process. Finally, it should be noted that the “after oscillations” in the firing probability, after the cessation of the stimulus, occur with the resonance frequency, just as is the case of $y^*(t)$.

We did not explore stimulation with pure tones in a

systematic way. We satisfied ourselves by noting that the average phase of the response to tones presented at up to 20 dB above threshold corresponds (within, say, 30°) to the phase of the revcor function at the corresponding frequencies. This comparison was done while measuring tuning curves and was found to hold over the frequency range over which the definition of the revcor phase spectrum is meaningful.

Figure 14 presents the result of the simulation procedure when a suddenly starting noise burst is used as the stimulus. In each panel the upper trace shows the stimulus waveform, the other traces show the signals $y^*(t)$ and $p(t)$. In this case the burst was chosen long enough (82 ms) to permit observation of the onset transient [Fig. 14(a)] as well as the steady-state behavior [Fig. 14(b)]. In Fig. 14(a) the gradual building up of the $y^*(t)$ waveform, reflecting the resonance implicit in the revcor

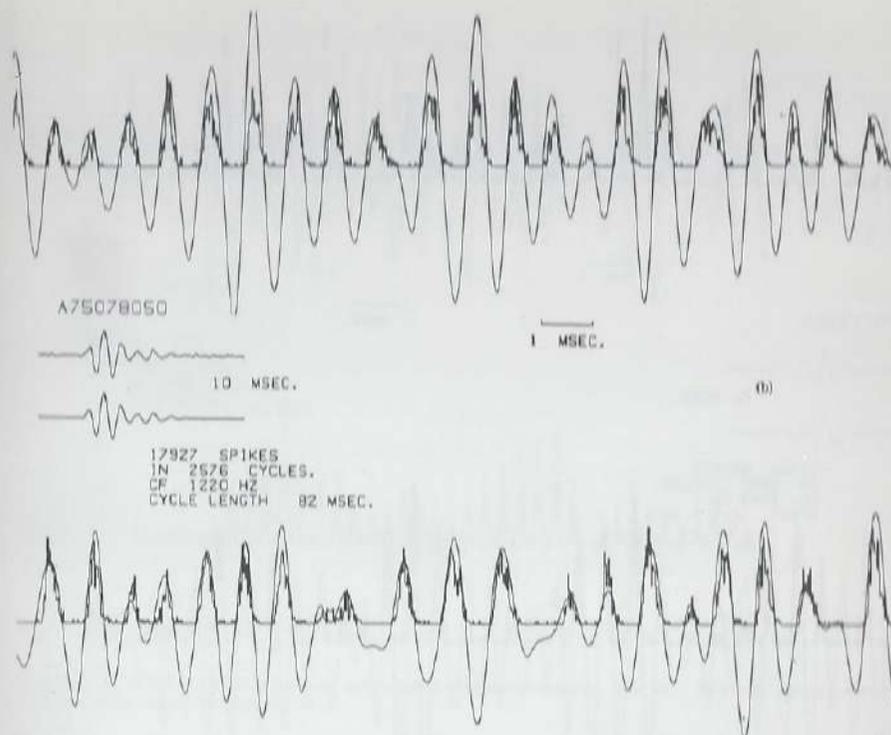


FIG. 10(b). Same as Fig. 10(a), but for unit nr.: 75-07-08, resonance frequency: 1.22 kHz, stimulation level: 50 dB/4 oct. The stimulus waveform is omitted.

function, is recognized. The firing probability appears to follow $y^*(t)$ closely but its initial amplitude is far larger than in later sections of the signal. There is no apparent distortion of the amplitude and phase modulations of the signal. From the instant marked by a dashed vertical line on, the waveform of the stimulus in Fig. 14(a) is identical to that in Fig. 14(b). Apart from the aforementioned difference in amplitude, the waveforms of $p(t)$ in these two parts are found to be practically identical. We observe again the tendency for “overlapping” between the individual lobes of $p(t)$ that was mentioned in connection with Fig. 11.

This figure, then, confirms that for a wide-band noise stimulus a linear transform of the stimulus is the prime controller of firing probability. Of the nonlinear effects only short-term adaptation is manifest, it appears as a multiplicative effect, modulating the firing probability when sudden changes in stimulus occur, and reducing firing probability as the stimulus continues. Note that the first phase of adaptation is very fast, it is almost completed within 2 ms. It is remarkable that fast adaptation shows up only in the form of onset and offset effects. In the course of the response to a stochastic signal there occur many fluctuations in the excitatory signal

$y^*(t)$, but short periods in which the oscillations have a nearly constant amplitude occur as well. We do not observe systematic adaptation effects during these periods. Apparently, fast adaptation shows up only when the amplitude changes rapidly over a wide range.

III. AUTOCORRELATION ANALYSIS

Of the many nonlinear effects in the cochlea, we want to pay particular attention to nonlinear filtering. The recordings presented so far do not betray much evidence of a derangement of filtering due to nonlinear effects. On the contrary, the results confirm that a linear transform of the stimulus signal is the prime controller of firing probability. This holds true within the following constraints: the application of wide-band stochastic stimuli or single pure tones in the main passband, and the use of stimuli with constant (average) intensity. In this section we subject the results to further analysis with the aim to detect possible effects of nonlinear filtering of a more subtle nature. In particular, we want to determine the fine structure of the spectrum of the $p(t)$ function for noise excitation. As said earlier, nonlinear filtering is most likely to affect the relation between intrinsic amplitude and phase modulation of signals

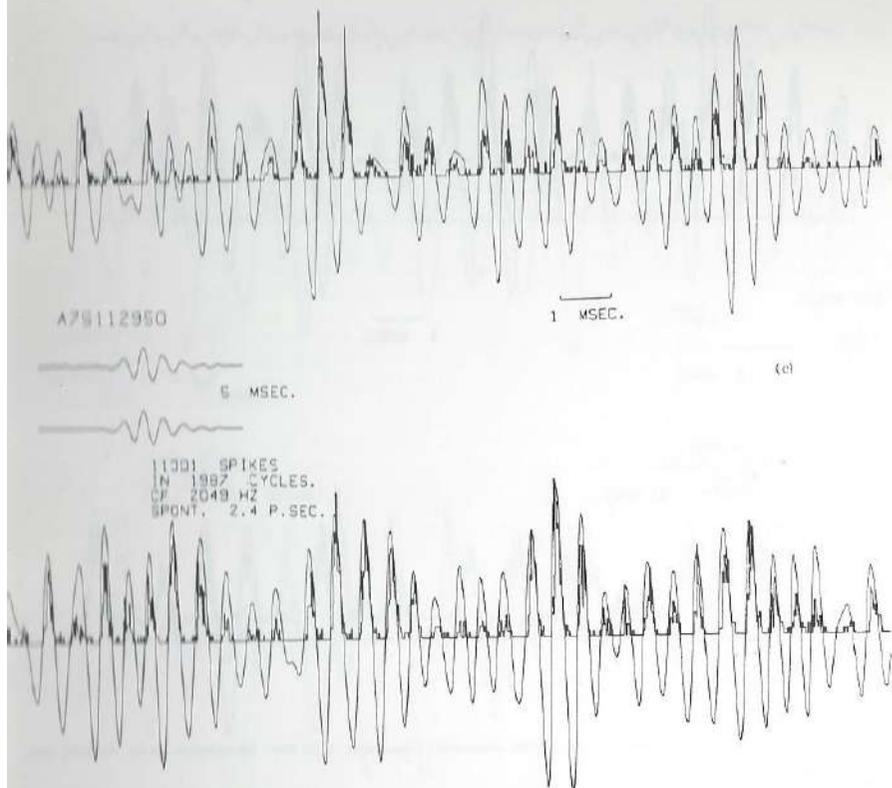


FIG. 10(c). Same as Fig. 10(a), but for unit nr.: 75-11-29, resonance frequency: 2.05 kHz, stimulation level: 50 dB/3 oct. The stimulus waveform is omitted.

(de Boer, 1976b) and this should be evident in the form of a modification of the spectrum. To this aim we first compute the autocorrelation function (acf) $\psi_p(\tau)$ of $p(t)$ and compare this with the acf $\psi_y(\tau)$ of $y(t)$. (For reasons of clarity we leave out the asterisks here.) The comparison will be executed in the frequency domain on the basis of the power spectra $\Psi_p(\omega)$ and $\Psi_y(\omega)$ corresponding to these functions.

The processing of the computed $y(t)$ signal presents no difficulties. In the processing of experimental $p(t)$ functions, however, we have to correct for two (trivial) types of nonlinearity. The most important one of these is rectification: The firing probability $p(t)$ is a non-negative function of time. In Appendix B it is described how it is possible to correct for rectification. In essence, the acf computed from the measured $p(t)$ function is pre-distorted so that an acf $\psi_p(\tau)$ is obtained that corresponds to a Gaussian variable $q(t)$ which is identical to $p(t)$ whenever the latter is positive. This transformation of the acf is called "derectification." Consequently, when we refer to the power spectrum $\Psi_p(\omega)$ we actually

mean $\Psi_q(\omega)$. The only proviso is that $p(t)$ is assumed to have the probability distribution of a rectified Gaussian variable. As it turns out, the effect of "derectification" on the main part of $\Psi_p(\omega)$ is very small indeed. (In effect we measured and processed some functions $p(t)$ in the form of a "compound histogram" as originated by Pfeiffer—see, e.g., Goblick and Pfeiffer (1969)—but this was found to be an unnecessary complication for the present purpose.) Thus the condition is not a very stringent one.

The second trivial type of nonlinearity concerns the form of the (instantaneous) relation between $y(t)$ and $p(t)$. See the remarks made in connection with Figs. 10(a)–10(c). Since derectification has so little effect on the shape of $\Psi_p(\omega)$ in the main frequency band, further corrections of instantaneous nonlinearities are not necessary (cf. Davenport and Root, 1958). Further details of data processing are given in Appendix B.

Figure 15(a) shows the result of this procedure applied to the $p(t)$ function of the neuron of Fig. 10(a). The figure

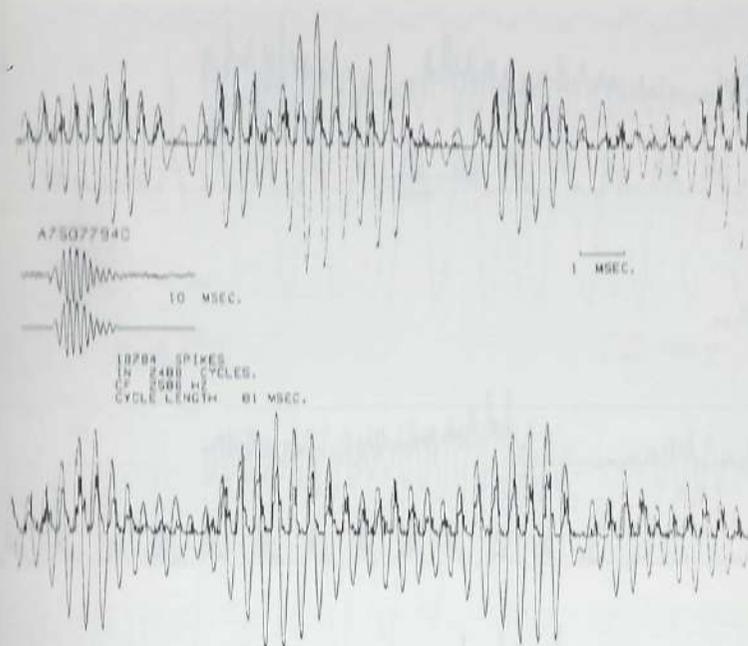


FIG. 11. Same as Fig. 10 but for a neuron with a higher resonance frequency. Unit nr.: 75-07-79, resonance frequency: 2.6 kHz, stimulation level: 40 dB/third oct.

shows the power spectra in much the same way as Figs. 5(b)–8(b) but the frequency scale is widened considerably. Three spectra are shown; below, the one corresponding to the measured $p(t)$ function [actually $q(t)$], center, the one corresponding to the $y(t)$ function, the third one (top) being the power spectrum belonging to the acf computed from the reverberation function $k(t)$. The three spectra are, for reasons of clarity, displaced over 5 dB with respect to one another. The latter two spectra should be identical when the input signal is (true) white noise; any remaining differences are to be attributed to processing effects, and, more important, to the fact that $\psi_p(\tau)$ is computed from a finite sample of a nonstochastic signal. As seen from the figure, the differences between these two spectra are less than 1 dB and can, therefore, be neglected. The third spectrum, that of $\psi_p(\tau)$ (actually, that of the corresponding q function) is very similar to the other two. The deviations are less than 1 dB over a range of 15 dB and may well be attributed to processing effects and sampling errors. We conclude that in this recording nonlinear filtering has not left any observable effect in the distribution of spectral components as they are represented in the firing pattern of the neuron. In particular, there is no widening of the resonance peak. This conclusion is corroborated by the results shown in Fig. 15(b) for another neuron with a fairly low resonance frequency, 1.5 kHz. It is seen that the p spectrum begins to be somewhat wider at a distance of more than 12

dB from the top. This may be caused by nonlinear effects (but this is by no means certain). We found the same type of results in ten other neurons with low resonance frequencies, with one exception. The deviations amount to 2 dB in this case and cannot be ascribed to processing errors.

Neurons with resonance frequencies well above 2.0 kHz show $p(t)$ functions that are definitely asymmetrical and tend to be more related to the envelope of $y(t)$ than to $y(t)$ itself, see Figs. 11 and 12. The autocorrelation functions $\psi_p(\tau)$ for such neurons reflect this aspect, of course. In fact they show an important component related to the acf of this envelope. But apart from this, $\psi_p(\tau)$ still reflects individual lobes of the $p(t)$ function and it is well possible to compare the spectra $\Psi_p(\omega)$ and $\Psi_y(\omega)$ in the region of resonance. Figure 16(a) shows the result of the processing of the data of the neuron of Fig. 14 of which the resonance frequency is 3.7 kHz. For this figure a still wider frequency scale was chosen in view of the small bandwidth involved. It appears that processing errors are somewhat larger but the result is consistent with the conclusion drawn above. Figure 16(b), finally, shows the result for the neuron of Fig. 13 with a resonance frequency of 4.7 kHz. For this neuron $\psi_p(\tau)$ shows only a small contribution related to the individual lobes of $y(t)$. With exactly the same processing it appears possible to study details of the representation

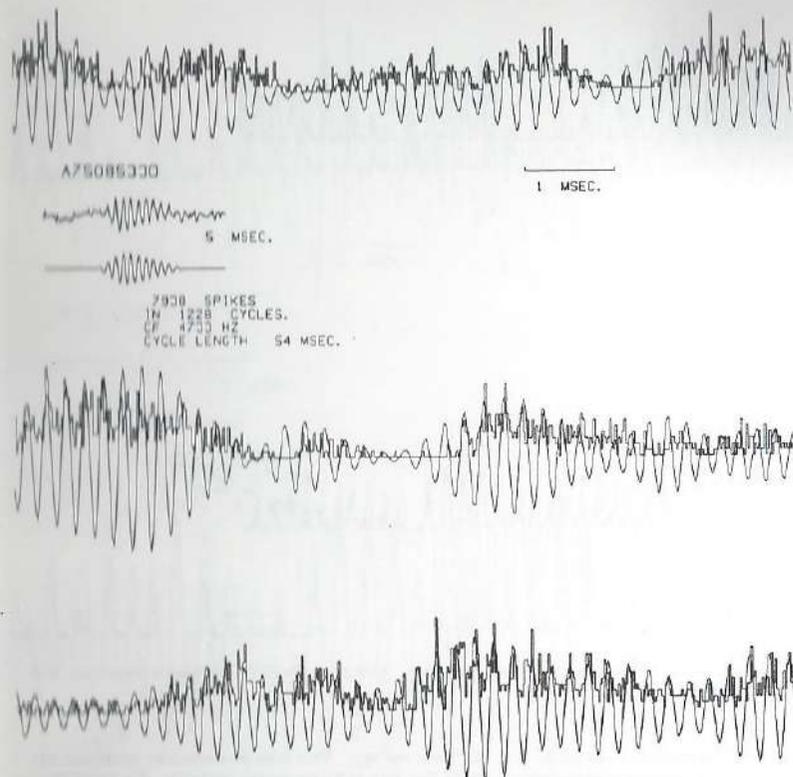


FIG. 18. Same as Fig. 10, for a neuron with a high resonance frequency. Unit nr.: 75-08-53, resonance frequency: 4.7 kHz, stimulus level: 30 dB/third oct. (20 dB above threshold), noise bandwidth: 15 kHz.

of frequency components in this neuron. The figure shows, then, that the spectrum $\Psi_p(\omega)$ in the passband is slightly wider than the spectrum $\Psi_s(\omega)$ but the slopes of the responses down from the resonance region are the same. Again we note that the effects of nonlinear filtering on the distribution of frequencies are marginal, if at all present. In quite general terms, cochlear non-linearity (nonlinear filtering) is of such a nature that it does not noticeably affect the timing pattern insofar as this represents components in the frequency region of resonance. Nonlinear effects do show up in the form of fast adaptation and in the intensity dependence of reverber functions but neither of these two effects involves impairment of cochlear filtering. It should be remembered that this holds true for wide-band stimuli and for single pure tones.

IV. DISCUSSION

The reverse-correlation method is observed to yield a characteristic function representing fundamental aspects of the stimulus-response relation of a primary

auditory neuron. This function turns out to be useful also for the prediction of firing patterns. In this respect the revercor function has advantages over conventional descriptors of the response pattern observed in auditory-nerve fibers, like the tuning curve or the click PST histogram. The method of reverse correlation has the additional advantage that it reveals details of frequency selectivity around the resonance frequency which are difficult to get otherwise. However, the method is not suited to reveal details about the skirts of the frequency response curve far from the top. In particular, the resolving power is not sufficient to assess the presence or absence of the low-frequency "tails" of the tuning curve in the revercor spectrum (cf. Kiang and Moxon, 1974).

All revercor functions have the general characteristics of the impulse response of a bandpass filter. We recall that the shapes of the revercor function's spectrum and the pure-tone tuning curve for the same fiber are most similar (de Boer, 1969a, 1973). General determinants of experimental revercor functions like initial delay, effective bandwidth, and slope of the spectrum are in good

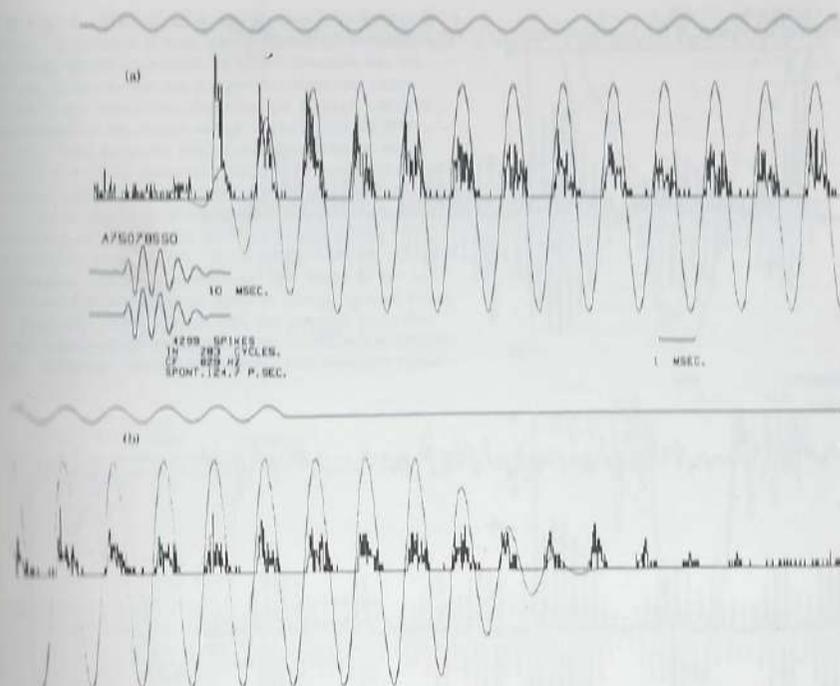


FIG. 19. Simulation result for a tone burst (a) shows the starting phase, and (b) the final phase. The layout is similar to that of Fig. 10(a). Unit nr.: 75-07-85, resonance frequency: 829 Hz, threshold at CF: 30 dB, stimulation frequency: 730 Hz, stimulus level: 80 dB.

agreement with corresponding data obtained by conventional methods (click PST histograms and pure-tone tuning curves). On a logarithmic frequency scale the amplitude spectra appear as asymmetrical. Mechanical response curves (e.g., Rhode, 1971) are highly asymmetrical. This difference suggests that a (second) filtering action is interspersed between mechanical events and excitation of the primary neuron (cf. Evans, 1978; Wilson and Evans, 1971). The phase characteristic of a revercor function shows around the resonance frequency a behavior that is reminiscent of a single resonating system. If we correct for the infinite delay (latency), the phase appears as an anti-symmetrical function around the resonance frequency. The only rapid and asymmetrical phase changes occur at frequencies for which the spectral amplitude is more than 30 dB below maximum. The influence of these phase changes on the response pattern to wide-band stimuli is negligible.

Revercor functions have little or no phase modulation in the waveform. Evidence on the presence or absence of phase modulation in the mechanical response of the cochlea is conflicting. On the one hand, a clear break in the phase function near the point of resonance (Rhode,

1971) and a clear frequency modulation in the impulse response (Rhode and Robles, 1974) are reported from measurements with the Mössbauer effect. On the other hand, no evidence on phase modulation was obtained with the capacitive-probe method (Wilson and Johnstone, 1972). Again, if we consider cochlear frequency selectivity as a two-stage process, we are led to conclude that the second filter is much more selective than the first so that in the narrow range of resonance the break in the phase curve can no longer be observed. On the other hand, if sharpening is achieved by a simple transformation of the output of the mechanical filter (Hall, 1977; Allen, 1977), the latter conclusion holds true as well.

The results presented in Sec. II illustrate the predictive power of the revercor function and confirm the validity of the associated simple linear model. The acoustical signal $x(t)$, when it is filtered by an appropriate filter, produces a signal $y(t)$ that clearly delineates the periods of preferred firing and predicts the course of the firing probability in several details. This property holds true for wide-band stimuli, the predictions of the linear model are valid over the range of intrinsic variations in a stationary random stimulus. For slow variations

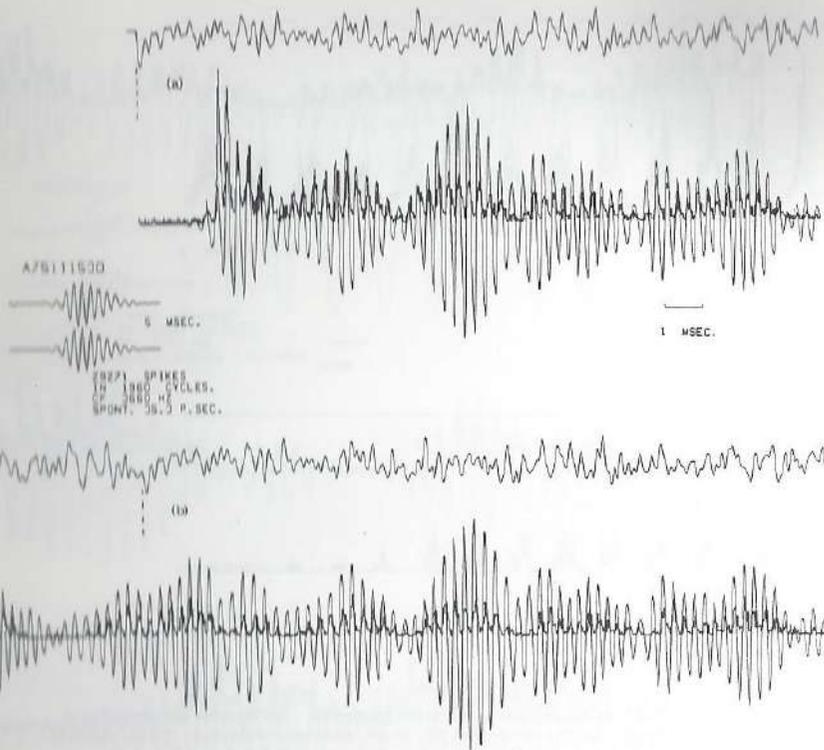


FIG. 14. Simulation result for a noise burst (a) shows the starting phase, and (b) shows a part of the recording where the waveform of the stimulus is exactly the same as in the starting phase. The dashed lines mark the beginnings of these sections. Unit nr.: 75-11-15, resonance frequency: 3.66 kHz, stimulus level: 30 dB (30 dB above threshold).

of average intensity we must take into account the dependence of the reverse function on intensity. For rapid variations we observe evidence of fast adaptation but only if the stimulus level changes rapidly over a large foveal range.

These two effects are the only manifestations of cochlear nonlinearity that are clearly present in our results. In view of known data on combination-tone generation and two-tone suppression we would have expected more evidence of nonlinearity, particularly of nonlinear filtering. Visual inspection of the simulation results in Sec. II does not yield clear manifestations of nonlinear filtering. The autocorrelation analysis of Sec. III reveals that the representation of the stimulus spectrum as it is processed by a primary auditory neuron is not noticeably different from a linear one. In the low-to-moderate intensity range, the nonlinearity that is responsible for the generation of combination tones and two-tone suppression seems to be of a type that does not affect frequency selectivity. At least, not for stimuli with a wide frequency spectrum. It seems safe to expect that this will hold true as long as the stimulus spec-

trum has slopes that are much lower than those of reverse amplitude spectra. (That is, it will hold for most, if not all, everyday sounds.) For these stimuli a linear transform of the stimulus is the principal controller of firing probability: the principle of specific coding (de Boer, 1973). Expressed in another way, this principle states that the components of a stimulus partake in the excitation of a neuron in relative proportions that are given by the neuron's reverse amplitude spectrum. This idea is consistent with findings on the near linearity of cochlear filtering obtained with quite different methods (Evans and Wilson, 1973).

It is most remarkable that the reverse correlation and the simulation method give so little evidence of nonlinear filtering. As far as the reverse function is concerned, this property can be understood. At present the model proposed by Pfeiffer (1970) comes nearest to the goal of describing all manifestations of cochlear nonlinearities. This model is composed of three independent elements in cascade: a first linear filter L_1 , a no-memory nonlinear circuit NL, and a second linear filter L_2 . This network should replace the linear network L in the mod-

el of Fig. 2. The filters L_1 and L_2 should be bandpass filters. A network of this type is known as a bandpass nonlinear (BPNL) network. A BPNL network has the curious property that for a Gaussian input the input-output cross-correlation function has a shape which is independent of the nature of the nonlinearity (de Boer, 1976a). This property can be considered as an extension of Price's theorem. It holds for any type of no-memory nonlinearity. A proof for the special case that the nonlinearity is an ideal rectifier, is implicit in the discussion of an extended primary neuron model by Johannesma *et al.* (1971). If the BPNL network really represents cochlear nonlinearity, the shape of the reverse-correlation function must be independent of stimulus intensity. This agrees with the general character of the experimental results for low-to-moderate intensities. However, the slight changes with intensity found

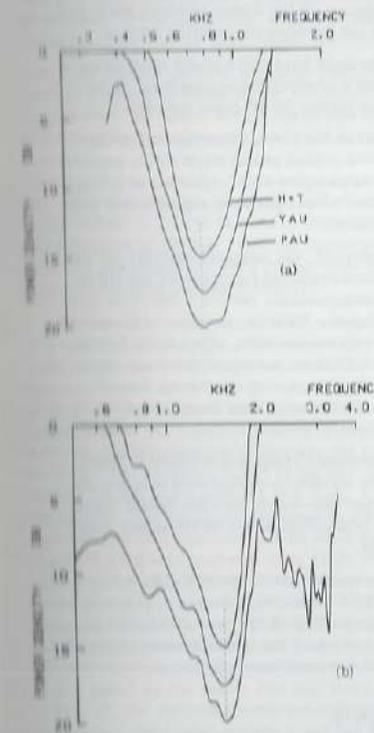


FIG. 15. Autocorrelation analysis applied to simulation result. Layout similar to the spectra of Figs. 5-8 but with a widened frequency scale, adapted to the resonance frequency and a different ordinate scale. See Appendix B. Symbols: PAU: acf of deresified $p(t)$ function (lowest curve), YAU: acf of $y^2(t)$ function (central curve), H^*F : power spectrum corresponding to $s^2(t)$ -function (upper curve). (a) Unit nr.: 75-08-73, resonance frequency: 0.750 kHz, stimulation level: 35 dB/1 out. See Fig. 10(a). (b) Unit nr.: 75-08-14, resonance frequency: 1.5 kHz, stimulation level: 50 dB/1 out.

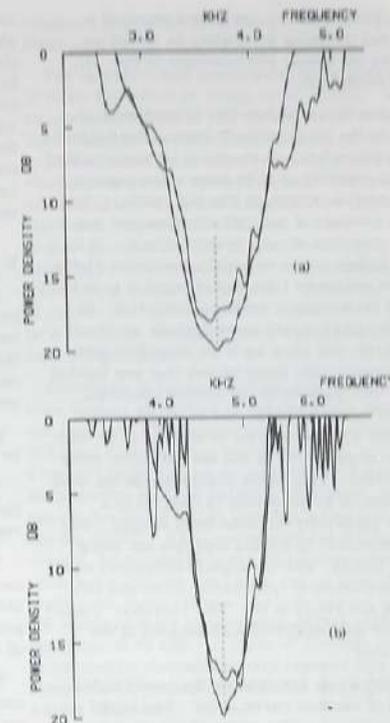


FIG. 16. As Fig. 15 but for two neurons with high resonance frequencies. Frequency scale still wider. (a) Unit nr.: 75-11-15, resonance frequency: 3.7 kHz, stimulation level: 30 dB/1 out. See Fig. 14. (b) Unit nr.: 75-08-53, resonance frequency: 4.7 kHz, stimulation level: 30 dB/1 out. See Fig. 12. Upper curve: PAU, lower curve: YAU, see legend to Fig. 15.

in the reverse functions suggest that the model does not account completely for all manifestations of nonlinearity. That is, part of the nonlinearity is of a type that affects resonance directly, particularly at high stimulus intensities.

That the simulation results show so little influence of nonlinearity is more difficult to understand. A first indication is contained in the fact that the nonlinearity produces a compression which should be (almost) the same for all stimulus intensities. Then there is no reason why the timing of the output signal of the BPNL network would be different for identical stimulus signals differing only in intensity (cf. Davenport and Root, 1958). However, it is more difficult to understand why the waveform of the output signal of the nonlinear network would be so similar to that of the corresponding linear network. It can be ascertained that the BPNL network approximately has this property, provided narrow-band filtering occurs after the generation of distortion products and the stimulus is a signal with a large

er bandwidth. Furthermore, the output signal of a BPNL network has the same properties in its acf as are demonstrated for neural $p(t)$ functions in Sec. III, under the same conditions.

In any event, the observed scarcity of nonlinear effects is not due to the characteristic linearizing tendency of a correlation procedure. It seems to be more related to a fundamental property of wide-band noise signals processed by a nonlinear system like the cochlea. We remark that the analysis of Sec. III presents just one approach to the detection of nonlinearity effects. It is probable that a higher-order correlation analysis yields more evidence of nonlinear filtering although it is to be doubted whether the accuracy would be sufficient. In this connection we predict that much clearer manifestations of nonlinearity will show up if we stimulate with signals having spectra with steep slopes that are located asymmetrically with respect to the reverber spectrum. In particular, we recommend the use of single-sideband (SSB) signals with the carrier frequency at or near the resonance frequency (CF) of the neuron under study (cf. de Boer, 1976b). This is the reason we do not wish to pursue the analysis as presented in Sec. III to a greater depth or generality for wide-band stimuli. Nor do we deem it necessary to do this analysis for more neurons. Note, finally, that the signals commonly used to demonstrate generation of combination tones and two-tone suppression are really of the "SSB" variety: Their spectra are highly asymmetrical with respect to the resonance frequency.

We conclude with a few remarks on the model underlying the method of reverse correlation. The model of Fig. 2 is a stochastic one. It is remarkable that the simulation method reveals so little evidence of the threshold crossing that is inherent in a neural triggering process. Our earlier work on the reverse-correlation method was based on a model (Weiss, 1964, 1966) in which a threshold-crossing process was explicitly built in. From our results it appears that the mechanism of triggering of a primary auditory neuron is considerably more complicated than in the simple Weiss model, even in a modified form (de Jongh, 1973). In particular, the stochasticity of the mechanism is of a different kind.

Consider Figs. 11, 12, and 14 and observe the small but systematic differences between the shapes of the functions $y^*(t)$ and $p(t)$. The lobes of the firing probability function $p(t)$ are somewhat narrower than those of $y^*(t)$, and they appear as slightly, but significantly, skewed. These properties reveal that the model of Fig. 2 is not sufficiently detailed to describe nerve fiber timing patterns on a microtime scale. Johannesma *et al.* (1971) proposed and discussed a generalized model for a neuron that takes several physiological factors into account, such as: propagation of EPSP's along a primary dendrite and temporal integration at the site of initiation of nerve firings. A similar model can be used to represent a primary auditory neuron. Quite formally, the additional effects are represented by a linear transformation, namely, a low-pass filter, inserted between the rectifying nonlinearity NL and the pulse generator PG in Fig. 3. In a qualitative way, the

observed discrepancies between $p(t)$ and $y^*(t)$ are consistent with such a model. Note that the theorem mentioned in connection with the BPNL network also applies to Johannesma's model, hence the cross-correlation function of the entire network relates directly to the response of the two linear systems in cascade. Since the secondary filter is a low-pass filter, it is readily understood that there is a high-frequency limit to the reverse-correlation method. More about the underlying model can be found in de Jongh's latest study (1977).

V. CONCLUSIONS

For the case of stimulation with wide-band signals the process of excitation of a primary auditory neuron can be split up into two parts (see Fig. 2), one involving linear filtering of the acoustical stimulus and the second one representing the generation of a stochastic series of pulses.

The impulse response of the linear filtering stage can be measured with the reverse-correlation method.

Rever functions have very little phase modulation in their waveform and a gradual phase spectrum in the region of resonance (within 20 dB from the top).

The output signal of the linear filtering stage—the excitatory signal—is a good predictor of firing probability. Positive values give rise to increased firing probability, negative values produce suppression, even of spontaneous activity.

For wide-band stimuli, the spectral content—as measured from the autocorrelation function—of the firing probability is indistinguishable (within 1 dB) from that of the excitatory signal. That is, no trace of frequency-selective nonlinearity is observed. This holds for the range within 12–15 dB from the top of the power spectrum, and for the great majority of neurons tested. (More evidence of nonlinearity may be expected for single-sideband stimuli, etc.)

Neurons with resonance frequencies above 2.5 kHz show characteristic waveform differences between the firing probability and the—rectified—excitatory signal. Nevertheless, the conclusion above on the power spectrum remains valid.

When stimulus intensity varies rapidly over a large range, adaptation effects can be observed. These do not affect the timing pattern of the neuron's action potential but appear to control the (instantaneous) rate of firing in relation to the excitatory signal.

ACKNOWLEDGMENTS

The authors wish to acknowledge their gratitude to L. C. M. van Keulen, neurophysiologist. He did not only assist in neurophysiological matters concerning the experiments but was also a most valuable partner in numerous discussions, notably on details of modeling the processing that occur in primary auditory neurons. Since the latter topic is a subordinate one in the present paper, his influence on our thinking is not too obvious in this text but it will be in future papers (cf. de Jongh

1977). The electronic setup was supervised by C. Euzenier on whom we could depend in all problems that arose in carrying out experiments.

APPENDIX A: DETAILS OF EXPERIMENTAL METHOD

The laboratory's PDP 9 computer is equipped with an ultrafast A/D conversion system and several especially designed types of interface to facilitate neurophysiological work. The A/D converter, Adage type VT 13 AB, has 14 bits of resolution and a conversion time of 5 μ s. The 16-channel interface system (DEC AF 09) allows transfer rates of over 100 000 words/s controlled by an external clock. For the work reported in this paper the standard clock frequency was 50 kHz. A modification of the A/D system allows us to use one additional bit for labeling, in this way the instants at which nerve-fiber action potentials are detected are transferred to the computer's memory with a maximum error of one clock period. For less time-critical applications, "event detectors" are used that generate a program interrupt for each firing detected.

For the measurement of a rever function the A/D system continually reads in the $x(t)$ signal. No sample is lost when the buffer is full and is starting to be refreshed from the bottom on. At each instant the immediate past of the signal is available and this can be used for the correlation computation according to Eq. (5). For this purpose the signal buffer is used as a circular buffer, i.e., the bottom is considered as the logical extension of the top. Several thousands of nerve firings are processed for each rever function, using double-precision integer arithmetic.

The stimulus noise is generated with a Hewlett Packard noise generator type 3722 A operated with a bandwidth of 5 kHz. For the measurement of the rever function the period of the noise is set to "infinite." To generate pseudorandom noise for studies with the simulation method the standard setting is "N=13." In that case the synchronization pulses are fed to the computer as well. The stimulus is fed to the ear via an earphone (Beyer DT 40) and a tube (91 mm long, 4.5-mm internal diameter). The earphone response is corrected with a filter so that it is flat within 5 dB when placed on a standard B&K artificial ear (type 4153) over the frequency range 300–8000 Hz. The tubing system introduces systematic variations of up to 7 dB and a typical resonance around 3 kHz. Until 1975 we measured the earphone response with the cat's ear *in situ* with help of a probe microphone. For the measurements reported in this paper we did not do this any more. As is stated in Sec. II, the electroacoustic transduction system is absorbed in the linear system L and for the comparison of the response to two different stimuli it is not necessary to correct for that. Such a correction would be necessary, of course, when we wish to relate properties of measured rever functions to sound pressure at the eardrum. Still the variations in the shape of the rever function or of its spectrum would be small, the main offender in this respect being the aforementioned resonance which can reach a peak of 10–15 dB (variable from animal to animal). The other wiggles in the frequency

response are virtually without effect, compared to the other sources of measurement error.

The operation and preparation techniques are essentially those described by Kiang *et al.* (1965). We have left an alternative method (cf. de Boer, 1973) in favor of the "classical" method because of its greater yield. The microelectrode is directly driven by a remote-controlled stepping motor and reduction gear. Recorded signals are amplified by standard amplifiers. All signals—stimulus, response, and timing marks—are fed directly to the computer and are also stored on magnetic tape using an Ampex FR 1500 instrumentation tape recorder. Tape speed was 30 or 60 in./s. In off-line processing of signals the tape speed was reduced by a factor of 4 or 8 to ensure that processing could keep up with the rate of input data.

Smoothing of rever function was carried out in the time domain as well as in the frequency domain. Rever functions were stored with a word count of 512 or 1024 and all truncation was carried out smoothly over a range of (at least) five successive points. Convolution, as is needed in the simulation procedure, was carried out via the fast Fourier transform routines developed for the PDP 9. When successive recordings are processed, as is needed in the simulation procedure, it is imperative that the read in of analog data occurs with exactly the same speed. For this reason we replaced the standard clock by a crystal-driven pulse generator operating at 50 kHz. A chain of dividers was used to accommodate processing with reduced tape speeds. A similar arrangement controls the tape speeds in the Ampex recorder.

In the presentation of the data we have continually used the convention that upward deflections in the rever functions correspond to positive sound pressure. In the realm of polarity conventions, it is to be noted that using a nonlinear pulse generator PG in the model of Fig. 2 with a firing probability $p(t)$ that is a nondecreasing function of the excitatory signal $y^*(t)$, automatically results in the constant of proportionality C in Eq. (6) being a positive number.

APPENDIX B: DETAILS ON THE AUTOCORRELATION PROCEDURE

Consider two variables v_1 and v_2 with a joint Gaussian distribution having a correlation coefficient ρ_w (we leave out the dependence on τ here). Two other variables w_1 and w_2 are derived from v_1 and v_2 by ideal rectification: for positive v_i , w_i and v_i are equal, and for negative v_i , w_i is zero. The correlation coefficient ρ_w of w_1 and w_2 is

$$\rho_w = 2 \int_0^{\infty} \int_0^{\infty} v_1 v_2 p(v_1, v_2) dv_1 dv_2,$$

where $p(v_1, v_2)$ is the probability density function of v_1 and v_2 . The integrals can be evaluated analytically (Bussgang, 1952) and the result is

$$\rho_w = \pi^{-1/2} [(1 - \rho_w^2)^{1/2} + \frac{1}{2} \rho_w + \rho_w \tan^{-1} \rho_w / (1 - \rho_w^2)^{1/2}]. \quad (A1)$$

Within an error of a few percent, this expression can be approximated by

$$\rho_v = (\rho_p + 1)^{1/2} / \pi \quad (A2)$$

over the full range $-1 < \rho_p < +1$. Inversely, ρ_p can be expressed in terms of ρ_v :

$$\rho_p = (\pi \rho_v)^2 - 1. \quad (A3)$$

This expression is used to convert the (normalized) autocorrelation function $\psi_p(\tau)$ of the firing probability $p(t)$ to the autocorrelation function $\psi_v(\tau)$ of a variable $q(t)$ that corresponds to v above. It is seen that this conversion—derectification—is very simple indeed.

If the signal $q(t)$ has the character of a narrow-band random signal, rectification has little effect on the spectrum in the main passband (cf. Davenport and Root, 1958), most of the distortion products arise elsewhere in the spectrum. Conversely, the derectification process has little effect in the main passband but it serves well—if not always satisfactorily in our recordings—to eliminate other distortion products due to rectification.

Records of the $y(t)$ and $p(t)$ signals are processed with a word count of 4000, the sampling period being 20 μ s as usual. Autocorrelation functions are computed for 512 values of τ , starting at 0. A triangular window over the entire τ range is applied to achieve some smoothing in the frequency domain; next, the autocorrelation function is made symmetrical to a word count of 1024 and is Fourier transformed. The plotting program was similar to the one used for plotting reverber spectra; it had the option to widen the frequency scale and to adapt the center point to the resonance frequency. It should be noted that the vertical scale has acquired a different meaning: in the reverber spectrum plots the spectral amplitude is plotted on a decibel scale according to its power but in Figs. 15 and 18 the Fourier transform of the acf has the dimension of power (density) and can be plotted directly in decibels.

We consider as a primary auditory neuron a functional unit that includes the external and the middle ear, the cochlea, the appropriate (inner) hair cell, and the process that leads to the production of an action potential in the associated nerve fiber. It transforms an analog signal—the acoustic waveform of the sound received—into a train of pulses ("events").

- Allen, J. B. (1977). "Cochlear micromechanics—a mechanism for transforming mechanical to neural tuning within the cochlea," *J. Acoust. Soc. Am.* 62, 936–939.
- Anderson, D. J., Rose, J. E., Hind, J. E., and Brugge, J. F. (1971). "Temporal position of discharges in single auditory nerve fibers within the cycle of a sine-wave stimulus: frequency and intensity effects," *J. Acoust. Soc. Am.* 49, 1131–1139.
- Buangang, J. J. (1952). "Crosscorrelation functions of amplitude-distorted Gaussian signals," MIT, RLE, Tech. Rep. No. 316.
- Davenport, W. G., Jr., and Root, W. L. (1958). *An introduction to the theory of random signals and noise* (McGraw-Hill, New York), p. 296 ff.
- de Boer, E. (1967). "Correlation studies applied to the frequency resolution of the cochlea," *J. Aud. Res.* 7, 209–217.
- de Boer, E. (1968). "Reverse correlation I," *Proc. K. Ned. Akad. Wet. C* 71, 473–486.
- de Boer, E. (1969a). "Encoding of frequency information in

- the discharge pattern of auditory nerve fibers," *Int. Audiol.* 8, 547–556.
- de Boer, E. (1969b). "Reverse correlation II," *Proc. K. Ned. Akad. Wet. C* 72, 129–131.
- de Boer, E. (1973). "On the principle of specific coding," *J. Dyn. Syst. Meas. Control* 95C, 265–273.
- de Boer, E. (1976a). "Cross-correlation function of a band-pass nonlinear network," *Proc. IEEE* 64, 1443–1446.
- de Boer, E. (1976b). "Spectral transformations by an infinite clipper," *J. Acoust. Soc. Am.* 60, 960–963.
- de Boer, E., and Jongkees, L. B. W. (1968). "On cochlear sharpening and cross-correlation methods," *Acta Otolaryngol.* 65, 95–104.
- de Boer, E., and de Jongh, H. R. (1971). "Computer simulation of cochlear filtering," in *Proceedings of the 7th International Congress on Acoustics* (Akademiai Kiado, Budapest), Vol. 3, 303–306.
- de Boer, E., and Kayper, P. (1968). "Triggered correlation," *IEEE Trans. Biomed. Eng.* BME-15, 169–179.
- de Jongh, H. R. (1972). "About coding in the VIIIth nerve," in *Proceedings of the IPO Symposium on Hearing Theory*, edited by B. L. Cardozo (Eindhoven, the Netherlands), pp. 70–77.
- de Jongh, H. R. (1973). "Analysis of a spike generator," in *Proceedings of the 3d IFAC Symposium on Identification and System Parameter Estimation*, edited by P. Eyckhof (The Hague, the Netherlands), pp. 607–612.
- de Jongh, H. R. (1977). "Models for the primary auditory neuron," (Doctoral dissertation (University of Amsterdam, the Netherlands) (to be published).
- Duijhuys, H. (1972). *Perceptual analysis of sound*, Doctoral dissertation (Eindhoven Technical University, the Netherlands).
- Evans, E. F. (1972). "The frequency response and other properties of single fibers in the guinea-pig cochlear nerve," *J. Physiol. London* 226, 263–287.
- Evans, E. F. (1976). "Place coding of frequency and intensity in the peripheral auditory nervous system: pros and cons" (Invited paper XIII Int. Congr. of Audiology), Audiology (Basel) (to be published).
- Evans, E. F. (1977). "Frequency selectivity at high signal levels of single units in cochlear nerve and nucleus," *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London).
- Evans, E. F., and Wilson, J. P. (1973). "Frequency selectivity of the cochlea," *Basic mechanisms in hearing*, edited by A. R. Møller (Academic, New York), pp. 519–551.
- Goblick, T. J., and Pfeiffer, R. R. (1969). "Time-domain measurements of cochlear nonlinearities using combination click stimuli," *J. Acoust. Soc. Am.* 46, 924–938.
- Goldstein, J. L., and Kiang, N. Y.-S. (1968). "Neural correlates of the aural combination tone $2f_1 - f_2$," *Proc. IEEE* 56, 981–992.
- Grashuis, J. L. (1974). "The pre-event stimulus ensemble. An analysis of the stimulus-response relation for complex stimuli applied to auditory neurons" (Academic thesis, Nijmegen).
- Gray, P. R. (1966). "A statistical analysis of electrophysiological data from auditory nerve fibers in cat," MIT, RLE, Tech. Rep. No. 451.
- Hall, J. L. (1977). "Spatial differentiation as an auditory 'second filter,' Assessment on a nonlinear model of the basilar membrane," *J. Acoust. Soc. Am.* 61, 520–524.
- Hind, J. E., Anderson, D. J., Brugge, J. F., and Rose, J. E. (1967). "Coding of information pertaining to paired low-frequency tones in single auditory nerve fibers of the squirrel monkey," *J. Neurophysiol.* 30, 794–816.
- Johannesma, P. I. M., van Gisbergen, J. A. M., and Grashuis, J. L. (1971). "Forward and backward analysis of temporal relations between sensory stimulus and neural response," Internal Report (Lab. of Medical Physics, University of Nijmegen, the Netherlands).

- Kiang, N. Y.-S., Watanabe, T., Thomas, E. C., and Clark, L. F. (1960). *Discharge patterns of single fibers in the cat's auditory nerve* (MIT, Cambridge, MA).
- Kiang, N. Y.-S., and Moxon, E. C. (1974). "Tails of tuning curves of auditory nerve fibers," *J. Acoust. Soc. Am.* 55, 620–636.
- Pfeiffer, R. R. (1970). "A model for two-tone inhibition of single cochlear nerve fibers," *J. Acoust. Soc. Am.* 48, 1373–1378.
- Pfeiffer, R. R., and Kim, D. O. (1973). "Considerations of nonlinear response properties of single cochlear nerve fibers," *Basic mechanisms in hearing*, edited by A. R. Møller (Academic, New York, London), pp. 555–591.
- Price, R. (1959). "A useful theorem for nonlinear devices having Gaussian inputs," *IRE Trans. Inf. Theory* IT-4, 69–72.
- Rhode, W. S. (1971). "Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique," *J. Acoust. Soc. Am.* 49, 1218–1231.
- Rhode, W. S., and Robles, L. (1974). "Evidence from Mössbauer experiments for nonlinear vibration in the cochlea," *J. Acoust. Soc. Am.* 55, 586–596.
- Ross, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (1967). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey," *J. Neurophysiol.* 30, 709–722.
- Sachs, M. B. (1969). "Stimulus-response relation for auditory-nerve fibers: two-tone stimuli," *J. Acoust. Soc. Am.* 45, 1035–1036.
- Sachs, M. B., and Kiang, N. Y.-S. (1968). "Two-tone inhibition in auditory-nerve fibers," *J. Acoust. Soc. Am.* 43, 1120–1126.
- Siebert, W. M. (1966). "Stimulus transformations in the peripheral auditory system," in *Recognizing Patterns*, edited by P. A. Kolers and M. Eden (MIT, Cambridge, MA), pp. 104–133.
- Smith, R. L. (1975). "Short-term adaptation and incremental responses of single auditory-nerve fibers," Special report IAS-8-11, Syracuse, NY.
- Weiss, Th. F. (1964). "A model for firing patterns of auditory nerve fibers," MIT, RLE, Tech. Rep. No. 418.
- Weiss, Th. F. (1966). "A model of the peripheral auditory system," *Kybernetik* 3, 153–175.
- Wilson, J. P., and Evans, E. F. (1971). "Grating acuity of the ear: Psychophysical and neurophysiological measures of frequency resolving power," in *Proceedings of the 7th International Congress on Acoustics* (Akademiai Kiado, Budapest), Vol. 3, pp. 397–400.
- Wilson, J. P., and Johnstone, J. R. (1972). "Capacitive probe measures of basilar membrane vibration," in *Proceedings IPO Symposium on Hearing Theory*, edited by B. L. Cardozo (Eindhoven, the Netherlands), pp. 172–181.

Chapter III. MODELS OF THE AUDITORY PERIPHERY

III-1. Introduction

Several models of the peripheral hearing organ exist; of the older qualitative models (which date from the time before electrophysiological data from single auditory nerve fibres were obtainable) the one of Davis (1957, 1958) is still of interest. The most important quantitative models are those of Siebert (1965, 1968, 1970), Weiss (1966), Duifhuis (1970, 1972) and Johannesma (1971).

Because Duifhuis' model is an extended and improved version of Siebert's model, the latter is not considered here. The supplements and improvements to Weiss' model made by Geisler (1968) and De Boer (1967) will be considered with Weiss' model in III-4.

The anatomical and physiological aspects relevant to modelling were discussed in chapter I. Some of the aspects are quite accurately known; others, however, can at the moment only be surmised. Thus models of the auditory periphery cannot as yet be very detailed and also give very different representations of the various stages of the transformation.

Three models are discussed in this chapter: those of Duifhuis, Johannesma and Weiss. The models are presented in their original forms as found in the mentioned publications; only for the sake of uniformity components were sometimes given new symbolic names. The chapter ends with a critical evaluation emphasizing the physiological aspects of each model. Deeper exploration, possible after a discussion of the analysis method and its results, is found in chapters V and VI.

III-2. Duifhuis' model

Fig. 3-1 depicts the model as described in Duifhuis (1972). It consists of two complexes: a filter bank followed by a firing model. In several respects this model is equal to the one described in the preceding chapter.

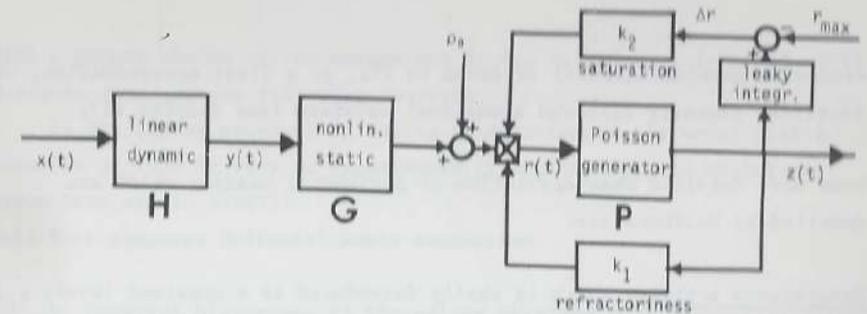


Fig. 3-1. Duifhuis' model. For the sake of uniformity the filter bank of the original model (Duifhuis (1972)) was replaced by one linear filter H.

The filter bank consists of a series of linear filters the properties of which agree with the frequency selectivity derived from electrophysiologically measured pure tonetuning curves. Thus the linearly filtered signal $y(t)$ does not represent the movement of the basilar membrane, but is a more sharply filtered hypothetical signal (see III-5).

The firing model is of the stochastic type. (The objections to a trigger model which led Duifhuis to this choice are discussed in section III-5.) The basic assumption incorporated into this model is that at each instant the firing probability density per unit of time is proportional to the amplitude of the stimulating waveform when the latter is positive; the probability is zero when the waveform is negative or zero. This firing model, mathematically a nonhomogeneous Poisson process, was introduced by Siebert (1965, 1968, 1970). The justification for a neural firing model consisting of a half-wave rectifier and a Poisson generator, P, is that the firing probability of a neuron in response to a stimulus (click,

sinusoid, complex stimuli) is found to fit, to a first approximation, the rectified linearly filtered acoustical waveform (see chapter II).

Some more detailed characteristics of peripheral hearing which are modelled by Duifhuis are:

Spontaneous activity which is easily introduced as a constant level, ρ_0 , added to the input of P (see Fig. 3-1). When no input signal is present, a Poisson process with rate, ρ_0 , (an excellent description of spontaneous activity) results.

Refractoriness which is modelled as a factor k_1 , by which the input to the Poisson process (the rate) is multiplied (when a spike occurred on $t = 0$):

$$k_1(t, \theta, \tau_1) = U(t - \theta) e^{-\tau_1/(t - \theta)} \quad (3-a)$$

in which $U(\cdot)$ is the Heaviside step function.

Duifhuis took the absolute component of the refractoriness, θ , to be 0.8 msec.; the time constant of the relative component, τ_1 , was assigned values between 0.1 and 1 msec.

Saturation which is introduced by a feedback system which controls the number of firings of a fibre. A "leaky count" of the output of the Poisson generator, P, is compared to a maximum rate to determine a factor, k_2 , which affects the input to P. The factor k_2 depends on Δr (the difference between "the desired rate" and the current rate) as follows:

$$k_2 = \frac{1}{1 + e^{\beta \cdot \Delta r}} \quad (3-b)$$

When $\beta = 1$ sec. and τ_2 , the time constant of the leaky integrator, equals 250 msec., the model shows an adaptation with a time constant of about 30 msec. (Duifhuis (1972)).

The complete model of Fig. 3-1 was realised as a software model. Simulations were done on sinusoidal, click and periodic pulse tone pip stimuli.

With a proper choice of the parameters θ , τ_1 , τ_2 and r_{max} (see Fig. 3-1) Duifhuis obtained the following results:

a) The firing probability during stimulation of the model with a sinusoid of 1000 Hz fits an experimentally obtained period histogram (from Rose et al. (1967)).

This fact supports Duifhuis' basic assumption.

b) Interval histograms of the pulses obtained with sinusoidal stimulation show good qualitative fits to the experimental data over a large dynamic range.

c) With the proper choice of parameters, PST histograms of click responses can be brought into good agreement with experimental results.

Adaptation occurs as a result of the adaptive behaviour of the saturation mechanism (see Duifhuis (1972)). Duifhuis' model may be characterized as a phenomenological model. The firing model is of a stochastic type.

III-3. Johannesma's model

We will deal with two versions of it; Fig. 3-2 shows the model as described by Koldewijn (1973). It consists of two complexes; the first complex is, as in Duifhuis' model, a linear filter H with sharp tuning curve properties. The second stage is explicitly supposed to model the properties of the hair cell plus the synapse and the primary afferent dendrite; it is a no-memory nonlinearity, G, followed by a lowpass linear filter K. An important difference between this and Duifhuis' model lies in the firing mechanism: Johannesma used the SILIT model (Stochastic Input Leaky Integrator Threshold model, Johannesma (1969)). The model implies an integration process presumed to lie in the peripheral part of the primary auditory neuron. When the integrated input signal passes the firing threshold, a spike is generated. An earlier version of the model (Johannesma (1971)) contained the SIPIT firing model (with a Perfect Integrator instead of a Leaky one). In that case the probability density of the firings

per unit of time is proportional to the input signal value (Johannesma (1969)). In the Leaky Integrator model this will only approximately be true.

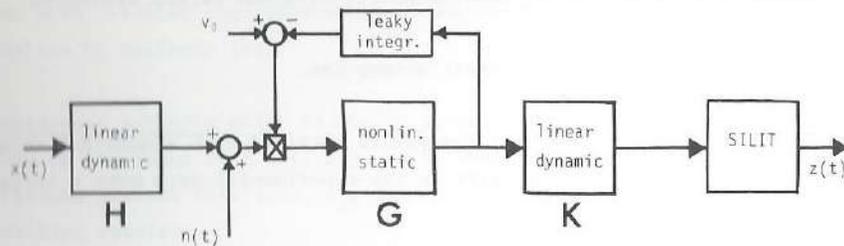


Fig. 3-2 Johannesma's model.

Here follow some detailed characteristics of peripheral hearing which are realised by the model:

Spontaneous activity is realised by an extra noise input just prior to the multiplicative action of the adaptation loop.

Refractoriness is present in this version of the model only as an inherent attribute of the SILIT component.

Saturation arises from the saturating characteristic of the rectifier. Some properties and simulation results of Johannesma's model have been published by van Gisbergen et al. (1971), Johannesma (1972) and Kolde-wijn (1973). Two properties are of special interest:

First, for units with low CF (<1 kHz), the firing probability is proportional to the rectified bandpass filtered input signal. This is in agree-

ment with Duifhuis' basic assumption.

The second property is concerned with synchrony. As already mentioned in chapter II, the phase-locking of neural spikes with the stimulus waveform gradually disappears for high stimulus frequencies. In Johannesma's model the phase-locking occurs as a direct consequence of the overlapping spectral characteristics of the band- and lowpass filters H and K (Fig. 3-2). When H and K do not overlap, no synchrony occurs. A proper choice of the cut-off frequency of K leads to good agreement between the model's loss of phase locking and experimental observations.

Adaptation is not realised by a sensitivity change of the pulse generating mechanism as in Duifhuis' model, but as a dynamic excitation change in the hair cell-synapse component. This adaptation occurs in an earlier stage of processing.

Johannesma's model may be characterised as a physiologically oriented model; the firing model is of the integrate and fire type.

Now we proceed with the third, much older, model.

III-4. Weiss' model

After the typical stochastic model of Duifhuis and the threshold model of Johannesma we treat the simpler and older model of Weiss (1966). Weiss' model (Fig. 3-3) was conceived shortly after the first electrophysiological single fibre data were available. Again, the first element is a linear filter, H, but here it represents the mechanical properties of the basilar membrane, the most plausible data available to Weiss - those of von Békésy (1950). This is in contrast with the other models in which H represents the sharp tuning curve filter.

The second element, G, is a transducer, modelled as a no-memory nonlinearity. The shape of the nonlinearity has a great influence on the simulation results of the model.

The last part is the firing model. It consists of a linear element, K, (which Weiss only casually mentions), followed by a trigger. When the input signal to the trigger passes a certain level (the threshold) in

upward direction, a spike is generated. After a firing, the threshold is reset to a maximum value which decays exponentially to a resting value with a time constant τ_R .

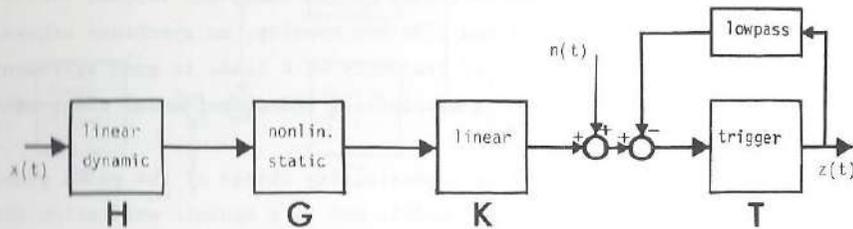


Fig. 3-3 Weiss' model.

The stochastic character of the firings results from a low-pass noise, $n(t)$, added to the trigger's input signal. This noise signal also produces spontaneous activity in the absence of a stimulus.

In considering the results of the simulation of Weiss' model, let us first look at spontaneous activity. The properties of spontaneous activity are governed by the time constant of the reset mechanism τ_R , and by the cut-off frequency of the low-pass noise. A choice of 0.3 msec., for τ_R and a bandwidth of 2-5 kHz for $n(t)$ will yield a satisfactory spontaneous activity.

The evoked activity cannot be simulated successfully, however. The failure is mainly caused by Weiss' choice of linear filter H, which is not sharp enough to predict the pure-tone tuning curves. Weiss partially solved this problem; with a proper choice for the transducer and firing model parameters, a sharpening can be obtained for frequencies up to 2 kHz.

This solution does not hold for the higher frequencies, however. The most simple way out of this problem was suggested by de Boer (1967). De Boer's argument is based on the fact that his receptor functions show very sharp tuning that cannot be explained by any instantaneous nonlinear action (see also chapter II). The solution is to take a sharp linear filter instead of the "basilar membrane filter".

Refractoriness again is an inherent property of the firing model. Geisler (1968) adapted the threshold reset mechanism so as to take care of an absolute refractory period.

Saturation results from the instantaneous nonlinearity, G. As for adaptation, this was not explicitly included by Weiss in this version of his model.

Weiss (1966) mentions several failures of his model, which would disappear with the introduction of the sharper linear filter. We will not discuss the results of the original Weiss model further. In the sequel we adopt de Boer's improvement and use a "modified model" which has a sharp linear filter as its first element, like the other two. Weiss' firing model could be characterized as a trigger model.

III-5. Discussion

The components of the models are now discussed in relation with their physiological correlates. Each section begins with a short description of the component and a brief exposé of the relevant physiological processes. Then an attempt is made to establish the link for each of the three models if applicable. During the presentation comparisons with the properties of the preceding model are given.

All three models are built according to the following general scheme: an initial linear filter is followed by a static nonlinearity which in turn is followed by a firing model. The output of the linear filter is a hypothetical cochlear signal which is supposed to mirror the receptor potential of the hair cell. The nonlinearity models the properties of the synapse between the hair cell and the primary auditory neuron. The

output of the nonlinearity thus mimicks the generator potential. The firing model, of course, reflects the properties of the dendrite and the "initial segment" of the primary auditory neuron.

1) The linear filter H - Basilar membrane filter, "second filter".

The mechanical properties of the basilar membrane as revealed by direct measurement of its displacement and velocity are linear over a large dynamic range (Von Békésy (1960), Johnstone et al. (1970), Wilson and Johnstone (1972)). Only one experimenter finds clear nonlinear behaviour at high intensities (Rhode, (1971, 1973)). However the mechanical filtering cannot explain the much sharper frequency selectivity found in single fibre responses, e.g., pure tone tuning curves and revcor functions. There is considerable evidence of a "cochlear sharpening"; Evans (1974) has suggested a "second filter" which is strongly dependent on metabolic processes and appears to be physiological vulnerable. However the nature of this filtering action is still very unclear, and it seems at the moment not justified to try to include it in a model. Besides, the total frequency selectivity of the auditory periphery (at least in cat) as deduced from neuronal responses in the auditory nerve is quite convincingly demonstrated to be linear over a large dynamic range (chapter II). Thus at the moment it seems adequate to describe the frequency selectivity, up to at least 60 dB above the threshold of a fibre, by means of one, sharp, linear filter.

All three models (when we use de Boer's modification of Weiss' model) do contain it, so on this point no controversy exists.

2) The rectifier - The sensory synapse.

The models contain a rectifier in the form of a static nonlinearity. The physiological mechanisms which underlie this element are supposed to reside in the hair cell - primary afferent synaptical junction. The properties of these synapses are unknown, but an analogy with the well known neuromuscular junction may be supposed for operational purposes. At the neuromuscular junction in the frog (del Castillo and Katz (1954)) - and later in many other excitatory synaptical junctions - it was found that during depolarisation of the presynaptic element, an enhanced tendency exists for the synaptical vesicles to empty into the synaptical

cleft. During hyperpolarisation the opposite occurs. This rectification effect is depicted in Fig. 3-4 for the squid giant synapse: the abscissa gives the presynaptical polarisation, the ordinate gives the postsynaptical polarisation as the result of the transmitter action. We can assume that the linearly filtered acoustical stimulus modulates the hair cell membrane potential, and that in the depolarised phase of the modulation synaptical vesicles empty into the cleft in amounts that depend on the extent of the depolarisation. Postsynaptically, EPSP's (excitatory postsynaptical potentials) are generated.

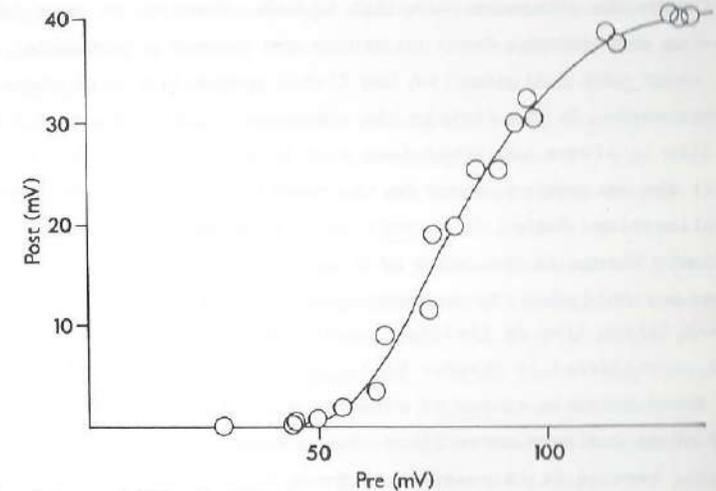


Fig. 3-4 From Katz and Miledi (1966).

If we assume that in the hyperpolarised phase no EPSP's are generated, then a rectification of the hair cell membrane potential results. (In this very simple description we ignore nonlinear effects which certainly exist in the mechano-electrical transduction, but these might be of minor importance in the low frequency encoding which we study here.)

It is well known that synaptic transmission may show both saturation and adaptation (Auerbach (1972)). Indeed Johannesma allocates these properties at the rectifier site (synapse).

Envisaging the processes, there are at least three possible sources of saturation. The first one is the receptor potential. At this level it

seems hard to conceive of a dynamic range of even 60 dB. For example, the maximum modulations of the hair cell membrane potential must be of the order of 20 mV. With a 60 dB dynamic range this requires the modulations at threshold of the order of 20 μ V- an unrealistic value in view of spontaneous membrane potential fluctuations, which might be of the order of 2 mV.

The second possible source of saturation is the generator potential. The effect of nonlinear summation of EPSP's (Martin (1966)) will tend to saturate the generator potential as well. However, we must face the following experimental fact: no matter how intense a sinusoidal stimulus is, we never meet "clipping" of the firing probability (see chapter II). In other words, in the firings the sinusoidal distribution of firing probability is always preserved (see also Rose et al (1971)). This would mean that the saturation cannot be the result of a saturating instantaneous nonlinearity. Rather it appears to be a dynamic sensitivity change - a dynamic change in the slope of a nonsaturating rectifier. (This argument does not hold when the instantaneous saturation is followed by a band pass filter like in Pfeiffer's model (Pfeiffer (1970)). This possibility is reconsidered in chapter VI.)

A third possible source of saturation is the limited firing rate capacity of an auditory nerve fibre. This, however, does not seem to play a role, because it is possible to force an acoustical nerve fibre into firing rates by means of electrical stimulation (Kiang and Moxon (1972)) much higher than those found in a normal, acoustical, stimulus situation.

Though all three models contain a rectifier, only Johannesma used it as a dynamic element by modelling the adaption and saturation at that site.

In view of the results concerning adaption of Eggermont (1972), Johannesma's modelling seems the most plausible. In view of the remarks about saturation, Weiss' solution of attributing the saturation to the static nonlinearity cannot be correct.

3) The firing model - The primary afferent, the "initial segment". The action potentials that travel along a primary auditory neuron are generated at the "initial segment" of its ganglion cell. When at that site the membrane potential surpasses a certain depolarisation level, a spike is initiated. In other words, spike initiation is the result of a threshold crossing. Consequently, a model of the auditory periphery should contain a threshold mechanism.

This does not necessarily mean that a non threshold model like the one of Duifhuis should not be able to simulate very accurately the input-output relations. It is even conceivable that the threshold character of the spike generation is indeed a detail which is irrelevant for an adequate description of the firing model. The following reasoning for instance eliminates subthreshold synaptic events and would justify such a view. The transmitter release from the presynaptic element, the hair cell, can be described as a Poisson process, the rate of which depends on the membrane potential of the hair cell. If, first, the size of an EPSP in the post synaptic element is relatively large (several mV's is conceivable in the very thin nerve endings involved), second, the cable properties of the primary dendrites are such that the EPSP's arrive relatively unaffected at the initial segment and, third, the firing threshold is low compared to the size of the EPSP's, then every transmitter quantum which is released will give rise to a spike (unless refractoriness interferes). In this way the Poisson character of transmitter release is translated into a point process-like generator potential (the single EPSP's are the "events") which in its turn is translated into the same Poisson process but now in the form of action potentials. The threshold character of the spike generation need not show up in the model in this case.

In view of our general knowledge of synapses and receptors however, (Papas (1972), Loewenstein (1971)) it is more likely that the generator potential is far from resembling a point process like the one described above. Very probably the generator potential of the primary afferents consists of many summed EPSP's together with a very substantial intrinsic membrane noise. According to Verveen (1963) the relative spontaneous

fluctuations in nerve fibres are related to the diameter of the fibre by: $\sigma/E = 0.03 d^{-0.8}$, in which σ is the s.d. of the fluctuations, E the mean membrane potential and d the diameter of the fibre in μm . Near the base of a hair cell we may expect σ/E to be about 0.2! In this case the generation of action potentials can not be described without the introduction of threshold crossings, which are, unfortunately, mathematically very difficult to handle (see e.g. Slepian (1962)). On the surface, the firing moments resemble an inhomogeneous Poisson process, but the threshold character of the generation of the spikes would be suppressed by such a description. In this thesis it will be shown that the incorporation of a threshold in the model enables us to solve a discrepancy between experimental facts and model predictions.

We now proceed with Johannesma's firing model. The lowpass filter, K , models the passive conduction of the generator potential along the postsynaptic element to the "initial segment". In our opinion such an element must be included in a model of the peripheral auditory system because of well-known properties of neuronal membranes.

Now returning to the firing model we must consider that there are indications suggesting, both indirect - on anatomical grounds (Spoendlin (1970))- , and direct (Robertson (1975)), that the spike originating site of the primary auditory afferent is not localised at the site of origin of the axon (in analogy with the motoneurone) but more peripheral, and that the spikes pass the soma via nodes of Ranvier. Thus the soma is probably not involved in an integration process connected to the firing mechanism. This makes it, in our opinion, less probable that the integration time is several milliseconds (Koldewijn (1973) uses 3.3 msec in his simulation). For a node of Ranvier of a frog nerve fibre, Tasaki (1955) measures a capacity of the node of about 1 pF and a resistance of about 50 M Ω . This means an RC-time of about 50 μsec . Of course we do not want to imply a strict resemblance between the initial segment of our interest and this frog node, we want to stress the possibility

that the integration time in the initial segment could be very short indeed. We now touch upon a correspondence which exists between Johannesma's SILIT model and Weiss' firing model: when the integration time gets very short, Johannesma's model degenerates to Weiss' (see VI-4).

Duifhuis' firing model takes care of the saturation and adaptation. Though these properties can be observed at the initial segment (Granit et al. (1963), Nakajima (1964)), the results of Ishii (1971) (in goldfish) and Eggermont (1972) (in guinea pig) suggest strong synaptic adaptation in the peripheral hearing organ.

Let us now consider the simplest threshold firing model - that of Weiss. In discussions of the complete model some rather serious shortcomings are often mentioned:

a) The internal noise in the model poses some difficulties (Duifhuis (1972)). A large dynamic range can only be obtained when the internal noise level in the model increases constantly with increasing input level. This problem does not arise when, as described above, we assume that the noise is added to the signal after adaptation and saturation - in physiological terms, when the membrane noise of the afferent fibre is the main noise source, and when the main adaptation and saturation effects are the result of dynamic sensitivity changes in the hair cell. In this case the signal to noise ratio of the generator potential remains constant when saturation is complete. (In principle this happens in Weiss' model, but with the type of saturation used the result is unsatisfactory, see chapter VI.)

b) For clicks of high intensity the Weiss model predicts PSTH's which are much narrower than those found experimentally. This difficulty also disappears when the argument in a) is applied.

c) With a constant trigger level and sinusoidal stimulation we would expect threshold crossings to occur closer to zero crossings of the stimulating signal as the stimulus level is raised. There should be a shift of the period histogram with increasing intensity. This shift is not found in primary auditory impulses evoked by sinusoids at CF (Andersen et al (1971)). This difficulty will be discussed after the Weiss model has been considered in detail (chapter VI).

Geisler's attempt to improve Weiss' model by changing the properties of the noise and the behaviour of the threshold failed because responses to stimulation at high intensity levels could not properly be simulated (Geisler (1968)). Again, the main reason for misfit with experimental data, apart from the wrong linear filter, lies in the kind of adaptation/saturation mechanism used in the model.

Though as a whole Weiss' model shows bad simulation results, the firing model may be more or less correct. In the sequel it will be shown that Weiss' firing model has an important property which gives a possible explanation for certain experimental findings.

Chapter IV. MATHEMATICAL ANALYSIS OF THE MODELS

IV-1. Introduction

In the following mathematical analysis of the models that were introduced in chapter III, the models of Duifhuis and Johannesma are treated together; the analysis is quite straightforward for the simplified versions used. Weiss' model, notwithstanding its simple appearance, is much more difficult to analyse. A simplified version of Weiss' model is analysed in this chapter, but that takes considerably more effort than the former analysis. In the following paragraphs the shapes of the revcor functions for the various models are calculated and it is attempted to elucidate the properties of the simulated signals $y^*(t)$ in relation to the firing probabilities $p(t)$ in these cases. As explained in chapter II, $y^*(t)$ is the linearly filtered input signal to the model, where the revcor function is taken as the impulse response of the linear filter. Looked upon in a more general way, the revcor function is proportional to the first order Volterra-Wiener kernel, which implies that $y^*(t)$ is the best (in mean square error sense) linear approximation to the output signal $p(t)$. (See for an extensive review of the Volterra-Wiener approach Hung and Stark (1977).)

IV-2. The analysis of the models of Johannesma, Duifhuis and Siebert

In this chapter we analyse the models without taking into account the adaptation mechanism. This mechanism is virtually inoperative in the case of Gaussian White Noise stimulation: attempts to demonstrate its effect failed. Another simplification reduces the problem of analysis considerably. The spike generating mechanism may be ignored because its transparency for the correlation technique used. This is true for these models because the spike generating mechanisms are instantaneous; i.e. the firing probability at any moment is proportional to the input signal (as for Johannesma's model, this is true for the SIPIT firing model (Johannesma (1969)), but only approximately for the SILIT model).

For the moment, then, we shall ignore the spike generator and concentrate on the analysis of the simple system of Fig. 4-1 (in the case of Duifhuis' and Sieberts' models we simply take $\delta(\tau)$ for the impulse response of the second filter, $k(\tau)$).

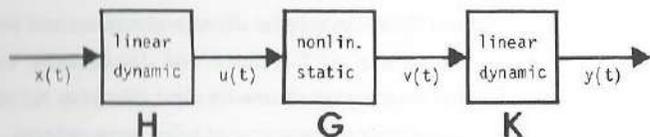


Fig. 4-1

The result of the first order crosscorrelation procedure:

The same definition of the crosscorrelation function used in chapter II will be used here. In order to obtain this function we start with the crosscorrelation function between $x(t)$ and $v(t)$:

$$R_{xv}(\tau) = \overline{x(t-\tau) v(t)} = c_1 h(\tau) \tag{4-a}$$

(where the bar means time-averaging: $\overline{x(t-\tau) v(t)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t-\tau) v(t) dt$)

with c_1 depending on the nonlinearity. This follows immediately from the theorem of Price (Price (1958)). The crosscorrelation function between the input and output signals then becomes:

$$\begin{aligned} R_{xy}(\tau) &= \overline{x(t-\tau) y(t)} = \int_0^\infty \overline{x(t-\tau) v(t-\sigma)} k(\sigma) d\sigma \\ &= c_1 \int_0^\infty h(\tau-\sigma) k(\sigma) d\sigma = (h * k)(\tau) \end{aligned} \tag{4-b}$$

where $*$ stands for the operation of convolution.

In the Fourier domain this reads:

$$\tilde{R}_{xy}(\omega) = c_1 \tilde{H}(\omega) \tilde{K}(\omega) \tag{4-c}$$

where $\tilde{}$ means the operation of Fourier transforming:

$$\tilde{Y}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt.$$

In this case the crosscorrelation function $R_{xy}(\tau)$ is proportional to the impuleresponse of the linear part (or cascade of linear parts) of the system.

Now we treat a useful case in which the amplitude characteristics of the linear filters H and K can be obtained via a second order analysis. When the nonlinearity $\phi(\cdot)$ is a function which can be expanded in a power series, the second order crosscorrelation function reads (for the case of a quadratic nonlinearity see Johannesma (1976))

$$\begin{aligned} R_{xxy}(\sigma, \tau) &= \overline{x(t-\sigma) x(t-\tau) y(t)} \\ &= c_2 \int_0^\infty h(\sigma-v) h(\tau-v) k(v) dv + c' \delta(\sigma-\tau), \end{aligned} \tag{4-d}$$

which in the Fourier domain becomes:

$$\tilde{R}_{xxy}^v(\omega, \nu) = c_2 \tilde{H}(\omega) \tilde{H}(\nu) \tilde{K}(\omega+\nu) + c'. \tag{4-e}$$

On the ancillary diagonal ($\nu = -\omega$) this function has the form:

$$\tilde{R}_{xxy}^v(\omega, -\omega) = c_2 \tilde{K}(0) |\tilde{H}(\omega)|^2 + c'. \tag{4-f}$$

From this relation the amplitude characteristic of H can be recovered. Then the amplitude characteristic of K follows from:

$$|\tilde{K}(\omega)| \approx \frac{|\tilde{R}_{xxy}(\omega)|}{|\tilde{H}(\omega)|} \tag{4-g}$$

Thus when the second order V-W kernel can be measured accurately and the nonlinearity is such that the amplitude of the kernel is not too small, the linear parts of the system of Fig. 4-1 can be determined. With the aid of a least squares method the nonlinearity can then be estimated.

Returning to the first order kernel, in the case of Johannesma's model we find:

$$R_{xy}(\tau) \approx (h * k)(\tau) \tag{4-h}$$

and Duifhuis' or Siebert's model yields:

$$R_{xy}(\tau) \approx h(\tau) \tag{4-i}$$

where \approx means proportional to.

We are now in a position to elaborate on the simulation result relating the linearly filtered stimulus, $y^*(t)$, and the firing probability.

According to (4-i), - the Duifhuis and Siebert case - the simulation result $y^*(t)$ should be proportional to an instantaneous nonlinearly distorted version of the input to the spike generator, and thus to the firing probability:

$$y^*(t) \approx u(t) \quad (4-j)$$

and

$$p(t) \approx y(t) = \phi\{y^*(t)\} \quad (4-k)$$

In these models the firing probability and the simulated signal are instantaneous nonlinearly related. For Johannesma's model, no such simple correspondence exists, although for very low frequency units an approximate relation similar to (4-k) holds. In Johannesma's model the simulated signal has the form:

$$\begin{aligned} y^*(t) &= \int_0^{\infty} x(t-\tau) h^*(\tau) d\tau \\ &= \int_0^{\infty} x(t-\tau) \int_0^{\infty} h(\tau-v) k(v) dv d\tau \\ &= \int_0^{\infty} u(t-v) k(v) dv \end{aligned} \quad (4-l)$$

while the firing probability is:

$$p(t) \approx y(t) = c \int_0^{\infty} \phi\{u(t-v)\} k(v) dv \quad (4-m)$$

Obviously there is no instantaneous relation between $p(t)$ and $y^*(t)$ here. Only when the filters H and K are such that the pass band of H falls completely within the flat pass band of K, and also that several higher harmonics of the narrow band signal $u(t)$ pass K relatively unaffected, will $p(t) \approx \phi\{u(t)\} = v(t)$ and $y^*(t) \approx u(t)$. Indeed in this idealised case the relation (4-k) holds, and thus for Johannesma's model we should expect an instantaneous nonlinear relation between the firing probability and the simulated signal. This would hold at least for low frequency units.

In all these cases, the simulation method or first order V-W approximation yields a signal, to which the actual firing probability has an instantaneous nonlinear relation. This means that a comparison of that signal, $y^*(t)$, and the corresponding $p(t)$ can serve to test the validity of the models. This is done in the next chapter. Now we will discuss the analysis of Weiss' model which is a much more difficult problem.

IV-3. The analysis of Weiss' model

The original Weiss model (Weiss 1963, 1966) was altered in Chapter III, where it was shown that the linear band pass filter with "Békésy" characteristics should be replaced by one with tuning curve characteristics. When we ignore the adaptation mechanism for the reason mentioned in IV-2, we arrive at a simple version of the modified Weiss model shown in Fig. 4-2. Note that in this case the spike generator can not be discarded because, as we shall now see, the transparency principle does not hold. In this paragraph we assume G to be an ideal halfwave rectifier. This is for the ease of the notation but does not present a restriction.

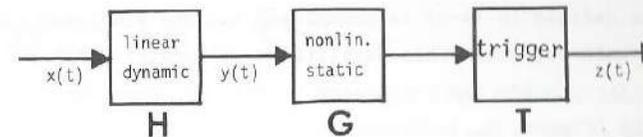


Fig. 4-2

The result of the first order crosscorrelation procedure:

The crosscorrelation function between a GWN input signal $x(t)$ and the resulting spike signal $z(t)$ (treated as a series of dirac δ -functions) for the system of Fig. 4-2 is given by

$$R_{xz}(\tau) = b h(\tau) + \sqrt{\frac{2}{\pi}} h'(\tau) \quad (4-n)$$

where $x(t)$ and $z(t)$ are assumed to be of unit variance and the quadratic content of $h(\tau)$ is equal to 1 (de Boer and Kuyper (1968)). The quantity b is dimensionless and is equal to the threshold of the trigger relative to the s.d. of $y(t)$; $h'(\tau)$ is the normalised derivative of $h(\tau)$:

$$h'(\tau) = c_0 \frac{d}{d\tau} h(\tau) \text{ with } c_0 = \frac{1}{\left[\int_0^\infty \left(\frac{d}{d\tau} h(\tau) \right)^2 d\tau \right]^{1/2}}$$

$$\text{thus } \int_0^\infty h'^2(\tau) d\tau = \int_0^\infty h^2(\tau) d\tau = 1.$$

We now define the function $h^*(\tau)$ for this chapter:

$$h^*(\tau) = \frac{1}{\sqrt{b^2 + 2/\pi}} \{ b h(\tau) + \sqrt{\frac{2}{\pi}} h'(\tau) \}. \quad (4-0)$$

This function is the normalised revcor function.

Because $h(\tau)$ and $h'(\tau)$ are approximately orthogonal in view of the narrow band frequency selectivity (see Chapter II for examples), the revcor function as defined in (4-0) is normalised and the simulated signal, $y^*(t)$, has unit variance. This facilitates mathematical manipulations. The result (4-0) holds for the system of Fig. 4-2 when the threshold level of the trigger, b , is fixed.

A fluctuating threshold:

It is not realistic to assume a fixed threshold. We must certainly expect stochastic fluctuations in the level (membrane noise, synaptic noise) and more or less deterministic fluctuations due to refractoriness. A fluctuating threshold is introduced as follows;

When $y(t)$ passes the threshold level and $y'(t)$ is positive, a spike is generated as in the original case, but now the threshold level is not fixed. The probability that the threshold level will have a certain level b at any given time, is governed by a probability density function $f_{th}(\cdot)$:

$$f_{th}(b) = p\{y(t) = b \mid t \text{ is a spike moment}\} = p\{y(t_n) = b\} \quad (4-p)$$

where t_n is a spike moment. For the time being we will not be concerned with problems in relation with the temporal behaviour of the threshold fluctuations; these will be discussed with the generalisation of Weiss' firing model (chapter V).

In the following treatment we shall simply assume that the time range of the auto-correlation function of the threshold fluctuations is sufficiently short compared to the time between the spike moments (threshold crossings).

For the calculation of the first order crosscorrelation function, we rewrite (4-n) as follows:

$$R_{xz}(\tau) = E\{x(t-\tau) \mid y(t) = b, y'(t) > 0\} \\ = b h(\tau) + \sqrt{\frac{2}{\pi}} h'(\tau) \quad (4-q)$$

where E means mathematical expectation.

With the aid of (4-p) we find for the fluctuating threshold case:

$$R_{xz}(\tau) = \int_{-\infty}^{\infty} f_{th}(b) \{ b h(\tau) + \sqrt{\frac{2}{\pi}} h'(\tau) \} db = \bar{b} h(\tau) + \sqrt{\frac{2}{\pi}} h'(\tau) \quad (4-r)$$

where \bar{b} is equal to the mean value of $y(t)$ on spike moments: that is, \bar{b} is the mean threshold level. Again, the function $h^*(\tau)$ is defined as the normalised crosscorrelation function given by

$$h^*(\tau) = \frac{1}{\sqrt{\bar{b}^2 + 2/\pi}} \{ \bar{b} h(\tau) + \sqrt{\frac{2}{\pi}} h'(\tau) \} \quad (4-s)$$

Thus in the case of an arbitrary threshold distribution, where spikes are generated when $y(t)$ crosses the threshold and $y'(t)$ is positive, the function $h^*(\tau)$ has the same shape as in the case of a fixed threshold level. The only difference is that the coefficient of $h(\tau)$ is now equal to the mean threshold \bar{b} instead of the fixed level b .

Clearly, the crosscorrelation procedure does not yield the impulse characteristic of the linear part in Weiss' model, contrary to the findings with the other models. In Weiss' case there is a perturbation term containing the derivative of $h(\tau)$. We should not expect a simple proportional relation between the simulated signal, $y^*(t)$, and the firing probability $p(t)$ in Weiss' case, as we should expect to find in the other cases (see 4-k). Before we proceed with this relation, we might ask if the perturbation term might be directly visible from the shape of the revcor function itself. Unfortunately, this is not the case; the linear filter H^* defined by the revcor function has slopes which can deviate no more than 6dB/oct. from the slopes of the actual filter H . In the case of the peripheral auditory system the slopes are too steep to perceive such a difference. So purely from the derived function $h^*(\tau)$ no conclusion can be drawn as to the presence or absence of a perturbation term which could indicate that a (fluctuating) threshold is present in the spike generating system.

The result of the simulation method:

The linear filter used in the simulation procedure, H^* , has the function given in (4-s) as its impulse response. Thus the simulated signal, $y^*(t)$, has the form:

$$y^*(t) = \frac{1}{\sqrt{b^2 + 2/\pi}} \{ \bar{b} y(t) + \sqrt{\frac{2}{\pi}} y'(t) \} \quad (4-t)$$

In this section we wish to clarify the correspondence between $y^*(t)$ and $p(t)$. Whenever the relation between $y^*(t)$ and $p(t)$ is quite different from that characteristic of the other models (4-k), it should be possible to infer from experimental data whether or not it is necessary to incorporate a threshold into the firing model.

In the fixed threshold case the character of the threshold does show up clearly in the simulated signal $y^*(t)$. The stochastic variable $y^*(t_n)$, where t_n , $n = 1, N$ are spike moments, is distributed as follows:

$$p\{y^*(t_n)=\eta\} = \begin{cases} \frac{(b^2+2/\pi)\pi\{\eta - b^2/\sqrt{b^2+2/\pi}\}^2}{4\sqrt{b^2+2/\pi}} e^{-\frac{b^2}{\sqrt{b^2+2/\pi}}\eta} & \eta > \frac{b^2}{\sqrt{b^2+2/\pi}} \\ 0 & \text{otherwise.} \end{cases} \quad (4-u)$$

This is easily found from:

$$y^*(t_n) = \frac{1}{\sqrt{b^2 + 2/\pi}} \{ b y(t_n) + \sqrt{\frac{2}{\pi}} y'(t_n) \} \\ = \frac{1}{\sqrt{b^2 + 2/\pi}} \{ b^2 + \sqrt{\frac{2}{\pi}} y'(t_n) \} \quad (4-v)$$

The constraint $y'(t_n) > 0$ together with the statistical independence of $y(t)$ and $y'(t)$ leads to the half Gaussian distribution to the right of the value $b^2/\sqrt{b^2 + 2/\pi}$. In Fig. 4-3a, b and c, these distributions are displayed for the threshold values $b = 0, 1$ and 2 respectively, together with the a priori (Gaussian) distribution of $y^*(t)$. For the sake of clarity the distributions are all normalised to the same maximal value.

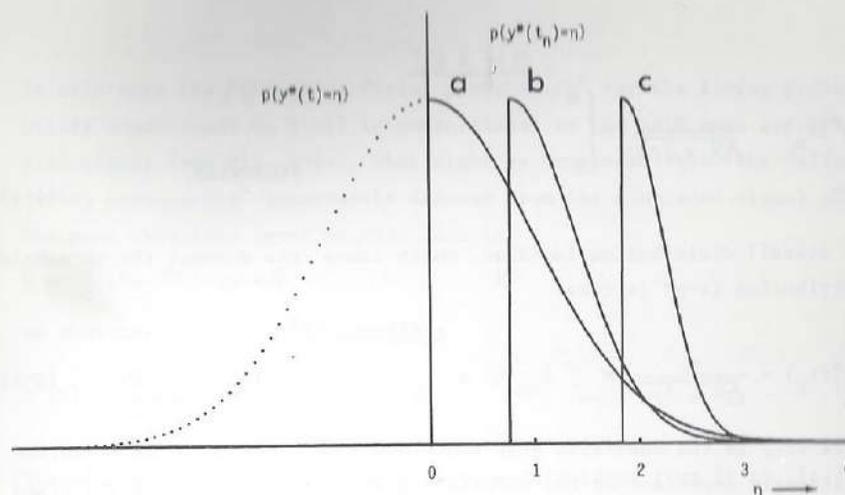


Fig. 4-3

The conclusion is that, although $y^*(t)$ is a perturbed version of $y(t)$ having the property that $y(t_n) = b$ for all t_n , the threshold character is still clearly preserved. In a histogram of values of the simulated signal at spike moments, all values lie to the right of the value $b^2 / \sqrt{b^2 + 2/\pi}$.

This is not the case for the fluctuating threshold where the threshold character need not be preserved in the $y^*(t)$ signal. At spike moments we now have:

$$\sqrt{b^2 + 2/\pi} y^*(t_n) = \bar{b} y(t_n) + \sqrt{\frac{2}{\pi}} y'(t_n) \quad (4-w)$$

in which the normalising factor $\sqrt{b^2 + 2/\pi}$ is introduced for the ease of calculation.

For a given threshold value, b , equation (4-w) yields the value:

$$\bar{b} \cdot b + \sqrt{\frac{2}{\pi}} y'(t_n) \quad (4-x)$$

which leads to the distribution function:

$$p\{y^*(t_n) = \frac{n}{\sqrt{b^2 + 2/\pi}}\} = \begin{cases} e^{-\frac{\pi(n - \bar{b}b)^2}{4}} & n > \bar{b}b \\ 0 & \text{otherwise} \end{cases} \quad (4-y)$$

The overall distribution function, which takes into account the threshold distribution (4-p) is thus:

$$p\{y^*(t_n) = \frac{n}{\sqrt{b^2 + 2/\pi}}\} = \int_{-\infty}^{\infty} f_{th}(b) e^{-\frac{\pi(n - \bar{b}b)^2}{4}} \cdot U(n - \bar{b}b) db \quad (4-z)$$

where $U(\cdot)$ is the Heaviside step function. The shape of the distribution is strongly dependent on the threshold distribution and unlike the fixed threshold case need not betray a threshold character at all.

If we go deeper into the problem of how far the threshold character of spike generation is preserved in the simulated signal $y^*(t)$, we must distinguish between two cases:

1. The threshold fluctuations are relatively small compared to the standard deviation of $y(t)$; that is, $f_{th}(\cdot)$ is a narrow distribution around a mean value \bar{b} .
2. The fluctuations are large, even a histogram of $y(t_n)$ values would not reveal the threshold character of the spike generator.

In case 1, the threshold character, evident in the values of $y(t_n)$ all being around \bar{b} , is more or less preserved in $y^*(t_n)$ values, because the latter are distributed according to a more or less smeared out version of the distribution of (4-u) (examples were shown in Fig. 4-3a, b and c). In case 2, the threshold character can no longer be deduced in the $y(t)$ domain from the $y(t_n)$ values, but rather from the fact that $y'(t_n) > 0$ for all spike moments t_n . This feature is completely lost in the $y^*(t)$ domain since $\frac{dy^*(t)}{dt} \Big|_{t=t_n}$ turns out to be essentially zero (see below).

To complete this section we give an example of a threshold distribution belonging to class two. Let us take:

$$f_{th}(b) = \begin{cases} b e^{-b^2/2} & b > 0 \\ 0 & b < 0 \end{cases} \quad (4-aa)$$

In this case the "effective firing probability" (or the firing probability conditional on $y(t)$) is proportional to the half wave rectified $y(t)$ signal (see Fig. 4-4a). What might be concluded about the "effective firing probability" incorrectly deduced from the simulated signal $y^*(t)$? The mean threshold level in this case is:

$$\bar{b} = \int_0^{\infty} b^2 e^{-b^2/2} db = \sqrt{\frac{\pi}{2}}, \quad (4-ab)$$

so that the function $h^*(\tau)$ would be

$$h^*(\tau) = \frac{1}{\sqrt{\pi^2/4 + 1}} \left\{ \frac{\pi}{2} h(\tau) + h'(\tau) \right\}. \quad (4-ac)$$

The distribution of $y^*(t_n)$ values can be calculated from (4-z). It is displayed in Fig. 4-4b together with the a priori distribution of $y^*(t)$ and the effective firing probability, which is the quotient of the two (all three are normalised to the same maximal value).

By comparing the firing probability to the simulated signal $y^*(t)$, we would conclude from Fig. 4-4b an expanding nonlinear dependence of $p(t)$ on the linearly filtered stimulus, but in reality the dependence is linear (for positive $y(t)$ values). This kind of problems we may meet when we are uncertain whether the spike generating mechanism is of a threshold type.

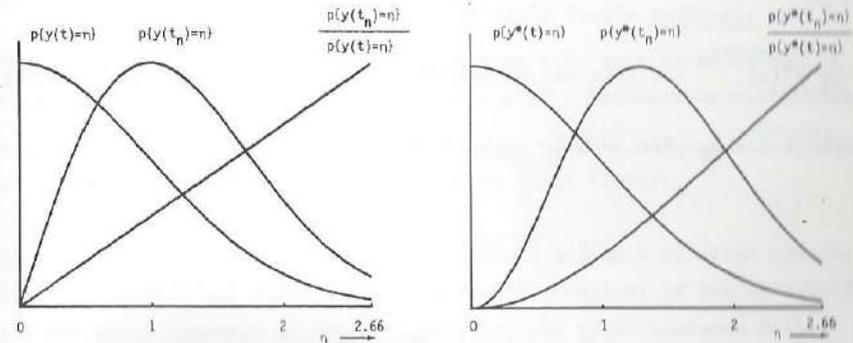


Fig. 4-4

It is clear from this section that the $y^*(t_n)$ histogram would be useful only in the case of a fixed or a relatively stable threshold; we must continue the analysis in order to find a better tool.

Simulated signal and firing probability:

An essential difference between the Weiss-like spike generators discussed in the last three sections and the other spike generators is that the time derivative of the input signal $y(t)$ plays a role in the first group but not in the second. Accordingly, in the second group (Johannesma, Duifhuis) we expect that the mean derivative of $y^*(t)$ on spike moments is approximately zero. This is indeed the case:

When we assume that (in the mean) no dependence exists between $p(t)$ and $y'(t)$, or $\overline{\frac{d}{dt}y(t)}^{t=t_n} = 0$, then, because $y^*(t)$ is functionally related to $y(t)$, also $\overline{\frac{d}{dt}y^*(t)}^{t=t_n} = 0$

For the Weiss-like models considered so far we have for the actual $y(t)$ signal the property:

$$\overline{\frac{d}{dt}y(t)}^{t=t_n} = E\{y'(t) | y'(t) > 0\} = \frac{\sqrt{2}}{\pi} \int_0^{\infty} n e^{-n^2/2} dn = \frac{\sqrt{2}}{\pi} \quad (4-ad)$$

but, surprisingly, we again have approximately zero mean time derivative of the simulated signal $y^*(t)$ at the spike moments:

$\overline{\frac{d}{dt}y^*(t)}^{t=t_n} \approx 0$. This can be seen from the following reasoning: (using (4-t))

$$\begin{aligned} \sqrt{b^2 + 2/\pi} \frac{d}{dt} y^*(t) &= \bar{b} \frac{d}{dt} y(t) + \sqrt{\frac{2}{\pi}} \frac{d}{dt} y'(t) \\ &= \bar{b} \sqrt{\int_0^{\infty} h'^2(\tau) d\tau} y'(t) + \sqrt{\frac{2}{\pi}} \frac{\int_0^{\infty} h''^2(\tau) d\tau}{\int_0^{\infty} h'^2(\tau) d\tau} y''(t) \\ &= \bar{b} c_1 y'(t) + \sqrt{\frac{2}{\pi}} c_2 y''(t) \end{aligned} \quad (4-ae)$$

It should be remembered that the notation ' means differentiation and normalising (see 4-n). The factor c_1 (for differentiating $y'(t)$) and c_2 (for differentiating $y''(t)$) are practically equal in the narrow band case, thus we find at the spike moments:

$$\sqrt{b^2 + 2/\pi} \frac{d}{dt} y^*(t) |^{t=t_n} \approx c \{ \bar{b} y'(t_n) + \sqrt{\frac{2}{\pi}} y''(t_n) \} \quad (4-af)$$

where \approx means: approximately equal to.

According to (4-ad): $\overline{y'(t)}^{t=t_n} = \frac{\sqrt{2}}{\pi}$. An accurate estimate of $\overline{y''(t)}^{t=t_n}$ is $-\bar{b}$, because $\rho_{y(t)y''(t)}(0) \approx -1$ and $\overline{y(t)}^{t=t_n} = \bar{b}$. Thus we can write for the mean value of the derivative of $y^*(t)$ on spike moments:

$$\overline{\frac{d}{dt}y^*(t)}^{t=t_n} \approx \frac{c}{\sqrt{b^2 + 2/\pi}} \{ \bar{b} \frac{\sqrt{2}}{\pi} - \sqrt{\frac{2}{\pi}} \bar{b} \} = 0 \quad (4-ag)$$

This result means that irrespective of the threshold distribution $f_{th}(\cdot)$, in the mean the spikes occur at maxima of $y^*(t)$ (see also de Jongh (1972)). An interesting feature is that in the fixed threshold case, the distribution of the spike moments around the maxima of $y^*(t)$ is very narrow when compared with the mean period of $y^*(t)$. This will be indicated for the case $b = 0$ (when spikes are generated on zero crossings of $y(t)$) but holds equally well for arbitrary b -values. When b equals zero, then, using (4-af), we get

$$\frac{d}{dt}y^*(t) \approx c y''(t) \quad (4-ah)$$

which on spike moments yields

$$\frac{d}{dt}y^*(t) |^{t=t_n} \approx c y''(t_n). \quad (4-ai)$$

The signal $y''(t)$ has a Gaussian distribution with unit variance but, since $y(t_n) = b = 0$ and $\rho_{y(t)y''(t)}(0) \approx -1$, $y''(t_n)$ is narrowly distributed around zero. From this it follows that spike moments can occur only near a maximum of $y^*(t)$ (see for examples de Jongh (1973)).

In the case of a fluctuating threshold the coincidence of spike moments with maxima of $y^*(t)$ disappears; when the fluctuations of the threshold are not small compared to the s.d. of $y(t)$, the spike instants will spread over as much as half the period of $y^*(t)$, as in the other models.

At this stage of the analysis we can only state that a fixed threshold model should reveal itself because the firing probability will be narrowly distributed around the maxima of $y^*(t)$, but the distinction between a fluctuating threshold model and a nonthreshold model on the basis of

a comparison between $p(t)$ and $y^*(t)$ escapes detection. A still more refined analysis is required to distinguish between the latter two cases. Before considering such an analysis we will, in the next chapter, introduce an experimental result and attempt to explain it, first in terms of a fixed threshold model, and then in terms of a fluctuating threshold model and the non-threshold models.

The results of this comparison lead to new insight into the relation between $y^*(t)$ and $p(t)$.

Chapter V. EXPERIMENTAL DATA AND CHARACTERISTICS OF THE MODELS

V-1. Introduction

In this chapter experimental data are discussed; the data are, like in chapter II, from single auditory nerve fibres in cat. We accentuate the behaviour of one unit: nr. 750712. This unit had a CF of about 600 Hz, and therefore is a good representative of low frequency units. Other units, forming a pool of about 15, corroborated the findings presented below. On the other hand their diversity was not large enough to permit a systematic analysis of the effect of stimulus parameters and unit characteristics on an experimentally observed particularity.

It is attempted to explain the experimental result introduced in V-2, with: a) a fixed threshold trigger, b) a fluctuating threshold trigger and c) the non-Weiss models, in V-3, V-4 and V-5 respectively. The data given in V-2 are only a starting point; it turns out that a more intricate characteristic, introduced in V-4, is crucial for detailed decisions about the models. The conclusion is then, that none of the models can explain the experimental result.

In V-6 a more thorough investigation of an experimental result leads to the idea that a time derivative is involved in the spike generation procedure. This idea leads to a generalisation of the analysis method: the revcor function has to be calculated for this more complex situation (V-7). In V-8 the results of V-7 are applied to the experimental result: the conclusion is, that if a dependence of $p(t)$ is postulated of both a signal $y(t)$ and its time derivative $y'(t)$, the experimental result can be explained in a surprisingly simple way.

V-2. An experimental result

Fig. 5-1 shows a plot of $p(t)$ against $y^*(t)$ obtained from unit 750712.

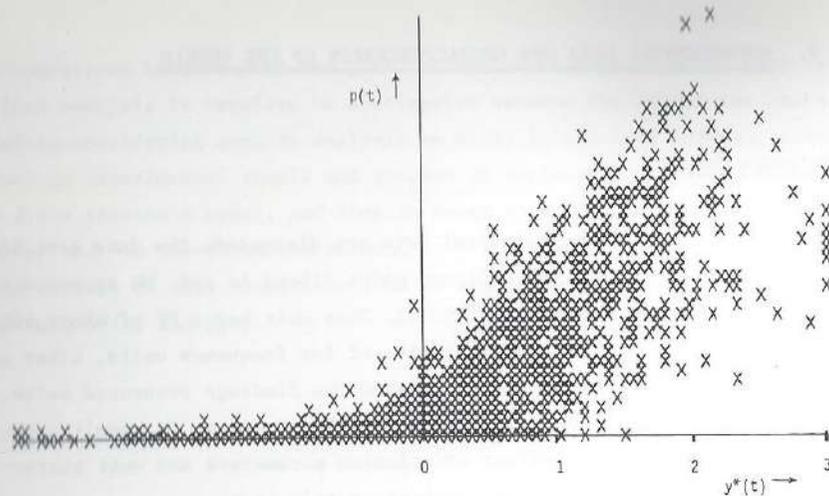


Fig. 5-1 horizontal scale: units s.d. of $y^*(t)$
vertical scale: arbitrary

Clearly a rectification effect can be seen; however there is considerable scatter of the data points.

This most simple characteristic in principle could be the result of a fluctuating threshold trigger or a non Weiss firing model. However, the fixed threshold could already be discarded when comparing Fig. 5-1 with Fig. 4-3, which displays the theoretical curves. Still, for the sake of completeness, the next section is devoted to the fixed threshold trigger case. The other firing models are treated in subsequent sections.

V-3. The experimental result and a fixed threshold

In this section we assume that the model of Fig. 4-2 holds true with a fixed threshold level for the trigger. If we assume b to be near zero then the most favourable correspondence exists between the theoretical and the experimental $y^*(t_n)$ plots of Figs. 4-3a and 5-1 respectively (which is however still poor).

There are two discordances in connection to the fixed threshold assumption, however. The first concerns the number of firings and their regularity. A fixed threshold near zero must give rise to a very regular firing pattern forced by the regularity of the zero crossings of the narrow band signal driving the spike generator. In contrast, spike trains in auditory nerve fibres show irregular firing reminiscent of Poisson processes (Gray (1967), Siebert (1972)). A plausible cure for this discrepancy is to add an extra noise to the signal before the spike generator (Weiss (1966)). This is equivalent to introducing a fluctuating threshold, however, and will be discussed in the next section. Nevertheless if we want to keep the principle of a fixed threshold we shall have to introduce the concept of firing probability so that when the threshold level is crossed by $y(t)$, a firing may result with a certain probability less than 1. In principle it is possible to generate pulse trains with the desired stochastic temporal firing pattern in this way. A second problem associated with the fixed threshold can not be cured. A fixed threshold implies perfect synchrony or phase locking of the spikes to a sinusoidal stimulus. Even at the highest sound levels no such perfect synchrony is found. Firings are distributed over at least one quarter of a cycle of the stimulus (Kiang (1965), Anderson et al. (1971)). Again, internal noise or jitter of the spike moments might solve this problem, but these are antagonistic to the fixed threshold assumption. Hence we must conclude that a fixed threshold trigger cannot describe the spike generating mechanism.

V-4. The experimental results and a fluctuating threshold

In this section we consider spike moments t_n , $n = 1, 2, \dots$ on which $y(t)$ has the properties: $y(t_n) = b$, $y'(t_n) > 0$, where b is a stochastic variable having a distribution function, $f_{th}(b)$, as introduced in section IV-3. We have shown that when \bar{b} is the mean threshold level that is $\bar{b} = \overline{y(t_n)} = \int_{-\infty}^{\infty} b f_{th}(b) db$, then the revcor function is given by (4-8) and the simulated signal will be

$$y^*(t) = \frac{1}{\sqrt{b^2 + 2/n}} (\bar{b} y(t) + \sqrt{2/n} y'(t)). \quad (5-a)$$

It was suggested in section IV-3 that an experimental result like the one of Fig. 5-1 can be explained in principle by a fluctuating threshold firing model. Let us see whether we can indeed do this!

We want to recover the threshold distribution $f_{th}(\cdot)$ from the properties of the simulated signal $y^*(t)$ at the spike moments, when we assume that the threshold fluctuates as defined in IV-3.

When $y(t)$ crosses the threshold on the level b , causing a spike at t_n , the value of the simulated signal is given by

$$y^*(t_n) = \frac{1}{\sqrt{b^2 + 2/\pi}} \{ \bar{b} b + \sqrt{2/\pi} y'(t_n) \}. \quad (5-b)$$

The mean value is given by

$$\overline{y^*(t_n)} = \frac{\bar{b}^2 + \sqrt{2/\pi} \sqrt{2/\pi}}{\sqrt{b^2 + 2/\pi}} = \sqrt{b^2 + 2/\pi}. \quad (5-c)$$

Thus the mean value of the simulated signal on the spike moments yields the mean threshold level directly:

$$\bar{b} = \sqrt{\overline{y^*(t_n)}^2 - 2/\pi}. \quad (5-d)$$

With the value of \bar{b} , the differential equation (4-s) can be solved:

$$\sqrt{b^2 + 2/\pi} h^*(\tau) = \bar{b} h(\tau) + \sqrt{2/\pi} h'(\tau). \quad (5-e)$$

The Fourier transform yields:

$$\sqrt{b^2 + 2/\pi} \tilde{h}^*(\omega) = \bar{b} \tilde{h}(\omega) + \sqrt{2/\pi} i\omega \tilde{h}(\omega) \quad (5-f)$$

and thus

$$\tilde{h}(\omega) = \frac{\sqrt{b^2 + 2/\pi}}{\bar{b} + \sqrt{2/\pi} i\omega} \tilde{h}^*(\omega). \quad (5-g)$$

With the aid of the inverse Fourier transform $h(\tau)$ can be found. It is thus possible to determine the desired impulse response $h(\tau)$ from the measured kernel $h^*(\tau)$. Once we have $h(\tau)$ we can reconstruct the signal $y(t)$ which, with the fluctuating threshold model, gives the threshold distribution $f_{th}(b)$ from the threshold values $y(t_n)$ directly.

This scheme is now applied to data obtained from unit nr. 750712. The mean value of $y^*(t)$ at the spike moments was: $\overline{y^*(t_n)} = 1.0$. Equation (5-c) gives the mean threshold level: $\bar{b} = \sqrt{(1.0)^2 - 2/\pi} = 0.6$. In Fig. 5-2 we see the revcor function $h^*(\tau)$ together with the solution of the differential equation (5-e) (the latter is called $h^{**}(\tau)$ because it is another, supposedly better, estimate of the impulse response $h(\tau)$).

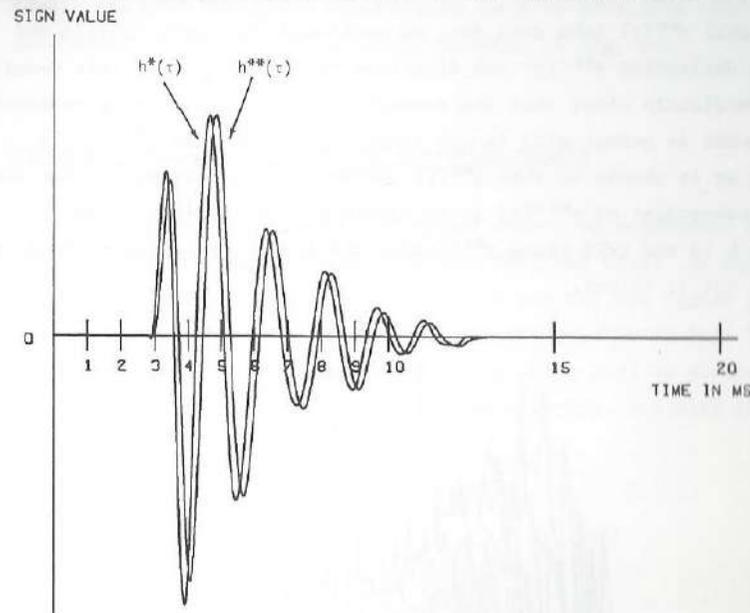


Fig. 5-2

With the aid of $h^{**}(\tau)$ another simulation was performed, which yielded the signal $y^{**}(t)$. This signal should be a better estimate of $y(t)$ than $y^*(t)$, because under the assumption of a (fluctuating) threshold mechanism the latter signal was contaminated with the time derivative $y'(t)$. Now obviously the justification of the assumptions should follow from the properties of $y^{**}(t)$. At the spike moments t_n , $n = 1, 2, \dots$, the values of $y^{**}(t)$ should be distributed according to the threshold distribution $f_{th}(\cdot)$. This fact does not yield a test for the validity of the assumptions, however, because there were no a priori assumptions with respect to

$f_{th}(\cdot)$; on the contrary, the $y^{**}(t_n)$ distribution should yield an estimate of the unknown threshold distribution.

The property which characterizes a threshold model, namely that $y'(t_n) > 0$ for all spike moments t_n , should also apply to $y^{**'}(t_n)$, however.

This provides a test for the validity of the threshold model defined in this section. The dependence of the firing probability $p(t)$ on the simulated signal $y^{**}(t)$ (the data are, as mentioned from unit 750712) and its time derivative $y^{**'}(t)$ are displayed in Fig. 5-3. From this result it is immediately clear that the assumption of the fluctuating threshold firing model is wrong: $p(t)$ is not identically zero when $y^{**'}(t)$ is negative as it should be when $y^{**}(t)$ correctly estimates $y(t)$. Also $p(t)$ is not independent of $y^{**'}(t)$ as it should be; on straight lines $y^{**}(t) = b$ in the half plane $y^{**'}(t) > 0$, the firing probability should be constant but it is not.

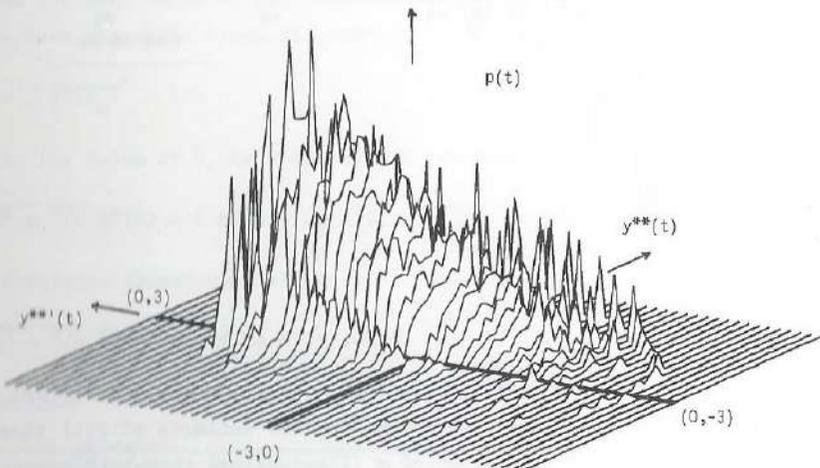


Fig. 5-3 scales in the horizontal plane: units s.d. of $y^{**'}(t)$
vertical scale: arbitrary

Thus from Fig. 5-3 we can see that a firing model with a threshold, even a fluctuating threshold, as defined in this section, is not compatible with the experimental data.

In the next sections a generalisation of the simple trigger firing model compatible with the given experimental data is developed. But first we will reconsider the non-Weiss models in relation to Fig. 5-3.

V-5 The experimental result and the non-Weiss models

For the non-Weiss models we should expect no dependence of $p(t)$ on $y'(t)$ at all, or only a slight one (see IV-2 and VI-4). It was shown in section IV-2 that in the case of units of low CF, the signal $y^*(t)$ is an estimate of $y(t)$. So, in order to get the pendant of Fig. 5-3, for this case we need to consider a figure in which $p(t)$ is plotted as a function of $y^*(t)$ and $y^{*'}(t)$. This is done in Fig. 5-4 with the same experimental data.

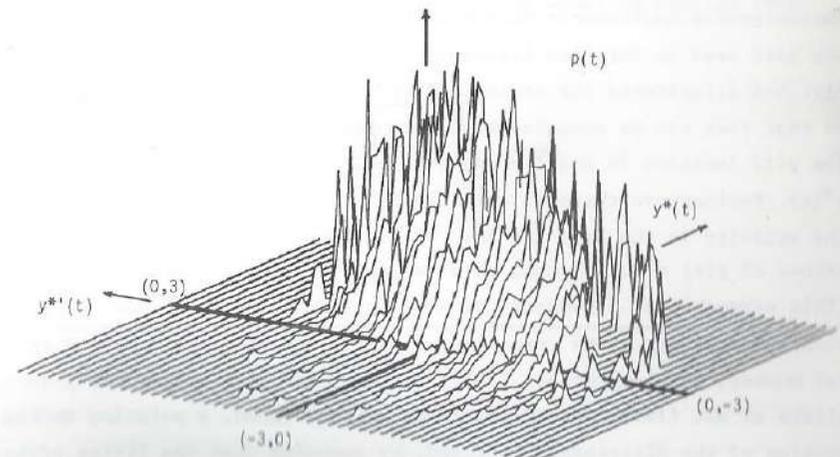


Fig. 5-4 scales in the horizontal plane: units s.d. of $y^*(t)$
vertical scale: arbitrary

Obviously $p(t)$ is dependent on $y^{*'}(t)$; $p(t)$ is clearly not an instantaneous function of $y^{*}(t)$ as it should be if a non-Weiss model were correct.

Thus it seems that the non-Weiss models are also incompatible with this experimental result (see chapter VI for further discussion). We are forced to extend the analysis to the case where $p(t)$ depends on both $y(t)$ and $y'(t)$. This is done in the next sections.

-6. An influence of $y'(t)$ on the firing probability

It appears that none of the models is compatible with the data presented. What is the reason of the discrepancy? In chapter II several examples of comparisons between $p(t)$ and $y^{*}(t)$ were displayed (Fig. 10). When we look at these we observe a "very good correspondence" between the two; in each positive going half period of $y^{*}(t)$ we find enhanced firing probability; in the other half period the probability is zero or very small. There are, however, systematic deviations of $p(t)$ from the instantaneous nonlinearly distorted (rectified) $y^{*}(t)$ waveform, which are best seen in very low frequency units (CF less than 1500 Hz).

Fig. 5-5 illustrates the deviations (the data are again from unit 750712 so that they can be compared with the preceding figures). Within each lobe the $p(t)$ function is clearly asymmetrical in contrast with the shape of $y^{*}(t)$. Furthermore there is a systematic relation between the place of the activity in the lobe and the "character" of the lobe: the smaller values of $p(t)$ occur as much as a quarter lobe later relative to $y^{*}(t)$. (This asymmetry or "skewness" of $p(t)$ is readily found in many publications on auditory nerve responses: e.g. in Rose et al. (1967) for squirrel monkey; in Kiang et al. (1965) for cat, and also very clearly in Klinke et al. (1977) for caiman!). As will be shown, a solution to the problem of the discrepancy is found, by assuming that the firing probability is determined not only by $y(t)$ but also by $y'(t)$.

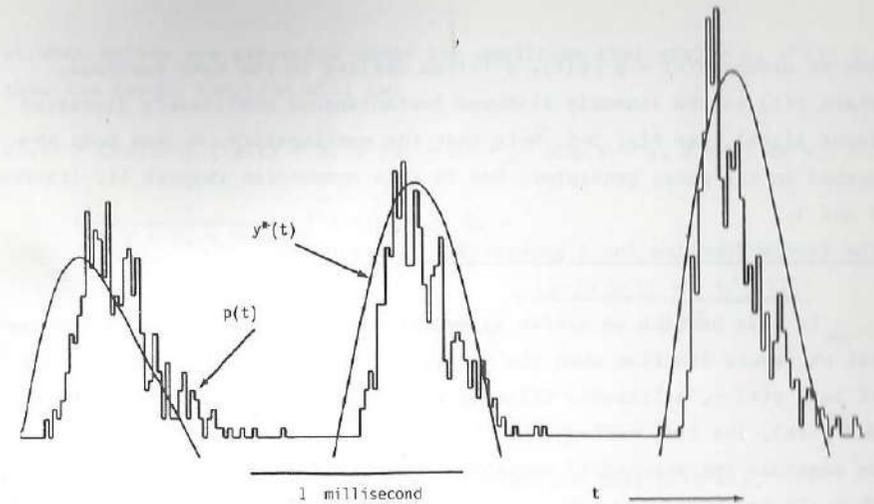


Fig. 5-5

We first solve the problem of calculating the revcor function (or first order V-W kernel) for a generalised model. In the generalised model the firing probability is determined by the pair $(y(t), y'(t))$ in a more complex way than in the fluctuating threshold model of section IV-3.

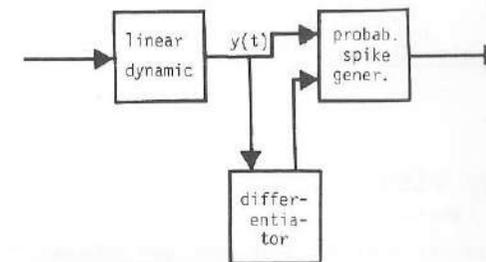


Fig. 5-6

Now we assume $p(t) = p(y(t), y'(t))$ as defined in the next sections, where $y(t)$ is the linearly filtered instantaneous nonlinearly distorted input signal, see Fig. 5-6. Note that the nonlinearity, G , has been absorbed in the pulse generator. See in this connection chapter II, figures 2 and 3.

7. The revcor function for a generalised firing model

In this section we derive an expression for the first order V-W kernel or revcor function when the firing probability, $p(t)$, is a function of both $y(t)$ -, a linearly filtered version of the input signal $x(t)$ - and $y'(t)$, its time derivative.

We consider the stochastic variables $x = x(t-\tau)$, $y = y(t)$ and $z = y'(t)$. They all are Gaussian with zero mean and unit variance: $x, y, z \sim N(0,1)$. The correlation coefficients are: $\rho_{xy} = h(\tau)$, $\rho_{xz} = h'(\tau)$ and $\rho_{yz} = 0$. Hence, the covariance matrix is:

$$\Lambda = \begin{pmatrix} 1 & h(\tau) & h'(\tau) \\ h(\tau) & 1 & 0 \\ h'(\tau) & 0 & 1 \end{pmatrix} \quad (5-h)$$

The joint probability density function of x, y and z can now be calculated (Peller (1950)):

$$p(x, y, z) = \frac{1}{(2\pi)^{3/2} \Delta} e^{-\frac{x^2 + (1-h^2(\tau))y^2 + (1-h^2(\tau))z^2 - 2h(\tau)xy + 2h(\tau)h'(\tau)yz - 2h'(\tau)xz}{2\Delta^2}}$$

$$= \frac{1}{(2\pi)^{3/2} \Delta} e^{-\frac{y^2}{2}} e^{-\frac{z^2}{2}} e^{-\frac{\{x - (yh(\tau) + zh'(\tau))\}^2}{2\Delta^2}}$$

(5-i)

where $\Delta^2 = \det(\Lambda) = (1-h^2(\tau) - h'^2(\tau))$.

The revcor function of the generalised model can be calculated with this function. For reasons of clarity this is done in four steps.

a) When spikes are generated under the condition that $y(t) = b$, $y'(t) = c$, then the revcor function will be:

$$h^*(\tau) = E\{x(t-\tau) \mid y(t) = b, y'(t) = c\} = \int_{-\infty}^{\infty} x p\{y = b, z = c\} dx =$$

$$= \frac{1}{p\{y = b, z = c\}} \int_{-\infty}^{\infty} x p\{x, b, c\} dx =$$

$$= \frac{e^{-b^2/2} e^{-c^2/2} e^{-b^2/2} e^{-c^2/2}}{\sqrt{2\pi} \Delta} \int_{-\infty}^{\infty} x e^{-\frac{\{x - (b h(\tau) + c h'(\tau))\}^2}{2 \Delta^2}} dx$$

$$= b h(\tau) + c h'(\tau) \quad (5-j)$$

b) When spikes are generated under the condition that $y(t) = b$, $y'(t) > 0$, (the problem of de Boer and Kuyper (1968)), the revcor function is:

$$h^*(\tau) = E\{x(t-\tau) \mid y(t) = b, y'(t) > 0\} =$$

$$= \frac{1}{p\{y = b, z > 0\}} \int_{-\infty}^{\infty} \int_0^{\infty} p\{x, b, z\} dz dx =$$

$$= \frac{\sqrt{2}}{\pi} \int_0^{\infty} e^{-z^2/2} \frac{1}{\sqrt{2\pi} \Delta} \int_{-\infty}^{\infty} x e^{-\frac{\{x - (b h(\tau) + z h'(\tau))\}^2}{2 \Delta^2}} dx dz$$

$$= \frac{\sqrt{2}}{\pi} \int_0^{\infty} e^{-z^2/2} \{b h(\tau) + z h'(\tau)\} dz$$

$$= b h(\tau) + \sqrt{\frac{2}{\pi}} h'(\tau). \quad (5-k)$$

c) A more general problem occurs when spikes can be generated under the conditions that $b \in B$, $c \in C$ with B, C measurable sets in $(-\infty, \infty)$. First we consider the case of a constant firing probability in $B \otimes C$ (\otimes means direct product). When $(y, z) \in B \otimes C$ then the firing probability is a positive constant, otherwise it is zero. In this case the revcor function can be calculated as follows:

$$h^*(\tau) = E\{x(t-\tau) \mid y(t) \in B, y'(t) \in C\} =$$

$$= \frac{1}{p\{y \in B, z \in C\}} \int_B \int_C p\{x, y, z\} dy dz dx$$

$$\begin{aligned}
 &= \frac{1}{p(y \in B, z \in C)} \frac{1}{2\pi} \int_B \int_C \frac{e^{-y^2/2} e^{-z^2/2}}{\sqrt{2\pi} \Delta} \int_{-\infty}^{\infty} x e^{-\frac{(x - (yh(\tau) + zh'(\tau)))^2}{2 \Delta^2}} dx dy dz \\
 &= \frac{1}{p(y \in B, z \in C)} \frac{1}{2\pi} \int_B \int_C (yh(\tau) + zh'(\tau)) e^{-y^2/2} e^{-z^2/2} dy dz \\
 &= \bar{b} h(\tau) + \bar{c} h'(\tau) \tag{5-1}
 \end{aligned}$$

where \bar{b} and \bar{c} are the mean values of $y(t)$ and $y'(t)$ respectively, at the spike moments. The results of problems a) and b) follow immediately from this result. We are now in a position to treat the most general case:

d) For each pair (y, z) a firing probability density $p_f(y, z)$ is defined. The reasoning now proceeds as follows: a point (b, c) in the $(y(t), y'(t))$ plane yields a contribution to the revcor function of the shape $b h(\tau) + c h'(\tau)$ (c.f. problem a).

The relative amplitude of the contribution is $p_f(b, c) p(y(t) = b, y'(t) = c) \stackrel{\text{def}}{=} p(b, c)$

The revcor function is now given by:

$$\begin{aligned}
 h^*(\tau) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (b h(\tau) + c h'(\tau)) p(b, c) db dc = \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b p(b, c) db dc h(\tau) + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c p(b, c) db dc h'(\tau) \\
 &= \alpha h(\tau) + \beta h'(\tau) \tag{5-m}
 \end{aligned}$$

where α and β are constants.

Again this result is a generalisation of the results from problems a), b) and c). For instance the solution (5-i) from c) follows easily from (5-m) when $p_f(b, c)$ is constant for $(b, c) \in B \otimes C$ and otherwise $p_f(b, c) = 0$.

The conclusion from the result (5-m) is that also in this generalised case, the revcor function is a linear combination of $h(\tau)$ and $h'(\tau)$.

The coefficients of $h(\tau)$ and $h'(\tau)$ can be interpreted as the mean values of $y(t)$ and $y'(t)$ respectively, taken at the spike moments.

Thus for any spike generating system where the firing probability is an instantaneous function $p_f(y(t), y'(t))$ of both $y(t)$ and $y'(t)$, where $y(t)$ is the linearly filtered input signal, the normalised first order V-W kernel or revcor function will be a linear combination of the impulse response $h(\tau)$ and its time derivative $h'(\tau)$:

$$h^*(\tau) = \frac{1}{\sqrt{\bar{b}^2 + \bar{c}^2}} (\bar{b} h(\tau) + \bar{c} h'(\tau)) \tag{5-n}$$

This means, in general, that the statement: the first order V-W kernel is the best estimate of the impulse response of "the linear part" of a system is not correct. A correct statement is: the convolution of the input signal and the first order V-W kernel does supply the best approximation, in mean square error sense, to the nonlinear transformed output signal.

V-8. The experimental data and the generalised firing model

The result of section V-7 is now used to explain the experimental data of unit nr. 750712 in terms of a generalised firing model.

We assume that the function $p(t)$ that controls the firing probability instantaneously depends on both $y(t)$ and $y'(t)$ as introduced in V-6. When the "firing probability function" as defined in section V-7d is $p_f(y(t), y'(t))$, the twodimensional threshold distribution density function is $p(b, c) = p_f(b, c) \cdot p(y(t) = b, y'(t) = c)$, as in (V-7), and we find the mean $y(t)$ -value \bar{b} and $y'(t)$ \bar{c} at the spike moments from:

$$\begin{aligned}
 \bar{b} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b p(b, c) db dc \\
 \bar{c} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c p(b, c) dc db \tag{5-o}
 \end{aligned}$$

According to the result of V-7 the revcor function is:

$$h^*(\tau) = \frac{1}{\sqrt{\bar{b}^2 + \bar{c}^2}} (\bar{b} h(\tau) + \bar{c} h'(\tau)) \tag{5-p}$$

and the simulated signal is:

$$y^*(t) = \frac{1}{\sqrt{\bar{b}^2 + \bar{c}^2}} (\bar{b} y(t) + \bar{c} y'(t)) \tag{5-q}$$

with a mean value at the spike moments (since $\overline{y(t)}^{t=t_n} = \bar{b}$ and $\overline{y'(t)}^{t=t_n} = \bar{c}$):

$$\overline{y^*(t)}^{t=t_n} = \sqrt{\bar{b}^2 + \bar{c}^2} \quad (5-r)$$

This mean value of the simulated signal at the spike moments can easily be obtained from the experimental data and enables us to choose pairs (\bar{b}, \bar{c}) with the proper squared sum. If the model is correct, one of these pairs is the "right" one; the problem now is to recognize this pair when no a priori information about the firing probability is available, other than the assumption that it depends on both $y(t)$ and $y'(t)$.

In our search for the "right" (\bar{b}, \bar{c}) pair we will use plots like those in Fig. 5-3 and Fig. 5-4. For any (\bar{b}, \bar{c}) combination such that (5-r) is fulfilled, the differential equation (5-n) can be solved. Now such a solution $h^{**}(\tau)$ (as it is called in V-4) can be used to generate a signal $y^{**}(t)$ which, in the case of the correct (\bar{b}, \bar{c}) pair, should be an uncontaminated estimate of $y(t)$. A plot of the firing probability $p(t)$ in the $y^{**}(t)$ - $y^{**'}(t)$ plane should reveal whether the chosen (\bar{b}, \bar{c}) pair is plausible.

We proceed with the data from unit nr. 750712. As mentioned in V-4, the mean value of the simulated signal at the spike moments was 1.0. We start with the choice $(\bar{b}, \bar{c}) = (1.0, 0)$. But then, according to (5-n) because \bar{c} is zero, the estimate $h^*(\tau)$ is proportional to $h(\tau)$ and the solution to the differential equation (5-n), $h^{**}(\tau)$, is the original $h^*(\tau)$ function. In this case Fig. 5-4 applies. As already said in V-5 such a dependence of $p(t)$ on $y(t)$ and $y'(t)$ is unlikely. We cannot possibly interpret this result and so we conclude that $\bar{c} \neq 0$.

To gain insight in this problem consider an arbitrary pair (\bar{b}, \bar{c}) such that $\sqrt{\bar{b}^2 + \bar{c}^2} = 1.0$: $(\bar{b}, \bar{c}) = (0.46, 0.89)$. With the aid of $h^{**}(\tau)$, the solution of (5-n), a signal $y^{**}(t)$ is generated. In Fig. 5-7, $p(t)$ is plotted as a function of $y^{**}(t)$ and $y^{**'}(t)$ for this choice. Evidently this solution is as inappropriate as the previous one. The problem now is how to obtain a more sensible choice for (\bar{b}, \bar{c}) . Every choice (\bar{b}, \bar{c}) leads, via the differential equation (5-n) to a function $h^{**}(\tau)$ which is a linear combination of $h(\tau)$, which we wish to determine, and $h'(\tau)$. A

similar statement holds for $y^{**}(t)$. This signal is a linear combination of $y(t)$ and $y'(t)$, the coefficient of the latter being zero for the correct (\bar{b}, \bar{c}) pair. This means that the $p(t)$ plots in the $y^{**}(t)$ - $y^{**'}(t)$ plane for all (\bar{b}, \bar{c}) pairs are identical in shape but show different rotation angles with respect to the $y(t)$ and $y'(t)$ axes. In other words, it should be possible to obtain the plot which we seek from each of the previous shown plots (5-3, 5-4, 5-7) simply by a rotation. To be precise: assume $y(t)$ to be transformed by $y^{**}(t)$ by

$$y^{**}(t) = y(t) \cos\phi + y'(t) \sin\phi.$$

Then, under the condition that $y(t)$ is a normalised narrow band signal, $y^{**'}(t) = y'(t) \cos\phi + y''(t) \sin\phi = -y(t) \sin\phi + y'(t) \cos\phi$.

This shows that under this condition the transformation of $y(t)$ and $y'(t)$ is orthogonal.

The most promising situation, plotted in Fig. 5-8, results from a choice $(\bar{b}, \bar{c}) = (.92, .42)$.

Of course the simplicity of this plot is no proof that these parameters are correct and that a firing model as defined in this section applies; but if such a dependence of the firing probability on $y(t)$ and $y'(t)$ is the case, this choice of (\bar{b}, \bar{c}) is much more likely than any other choice. If this is the case, then the resulting $y^{**}(t)$ estimates $y(t)$ without contamination with $y'(t)$, and the dependence of $p(t)$ on both $y(t)$ and $y'(t)$ can be determined from the plot 5-8. Clearly we find that:

- a) the firing probability $p(t)$ is virtually zero when $y(t) < 0$;
- b) $p(t)$ increases monotonically with $y(t)$ when $y(t) > 0$;
- c) $p(t)$ increases monotonically with increasing $y'(t)$.

A few more units were analysed in this way; they gave qualitatively the same result. As said, not enough different units and stimulus situations were met for a complete analysis. It would for instance be very interesting to know the dependence of the skewness on parameters like: the intensity of the stimulus, the CF of the unit and the spontaneous activity. However, the result reported in this chapter is interesting on its own; it is discussed in the next chapter.

At this point, in connection with Fig. 5-8, a reference should be made to the work of Grashuis (1974). In his analysis of unit responses in the cochlear nucleus, he concluded that the firing probability, in primary like units anyway, could be described with the aid of two

orthogonal linear functionals on the input signal $x(t)$. A favourable approximation seems to be provided by $p(t) = e^{Q(w,w')}$, where Q is a polynomial of the second degree and w, w' are the above mentioned functionals. In this way, accurate approximations to both our plot (5-8), and those of Grashuis, should be possible. Grashuis uses for w a convolution with the revcor function, and for w' a convolution with the Hilbert transform of the revcor function. Though mathematically this is attractive, interpretation is difficult. We rather use the time derivative of the revcor function for w' , the difference however, is minute for the narrow band signals involved. Still Grashuis' plots look different from ours (they seem to be more symmetrical); this is probably due to the fact that we recorded from auditory nerve fibres, while Grashuis used data from the more central auditory station of the cochlear nucleus.

In the next chapter we proceed with our idea of the involvement of a time derivative in the firing probability.

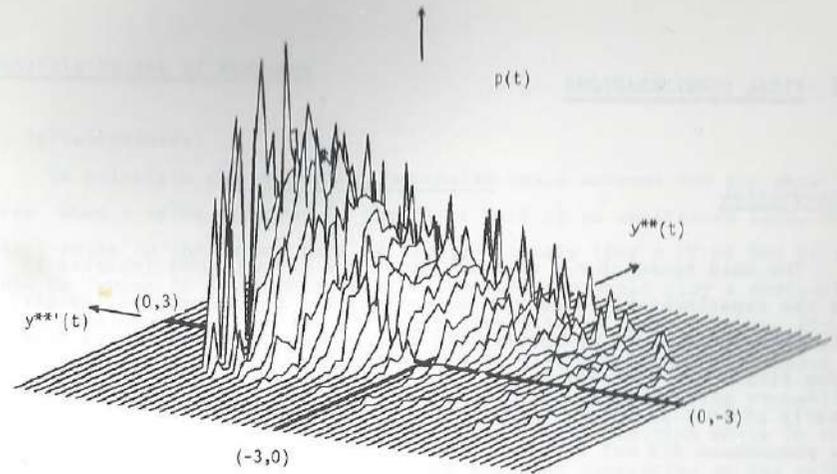


Fig. 5-7 scales as in Fig. 5-3

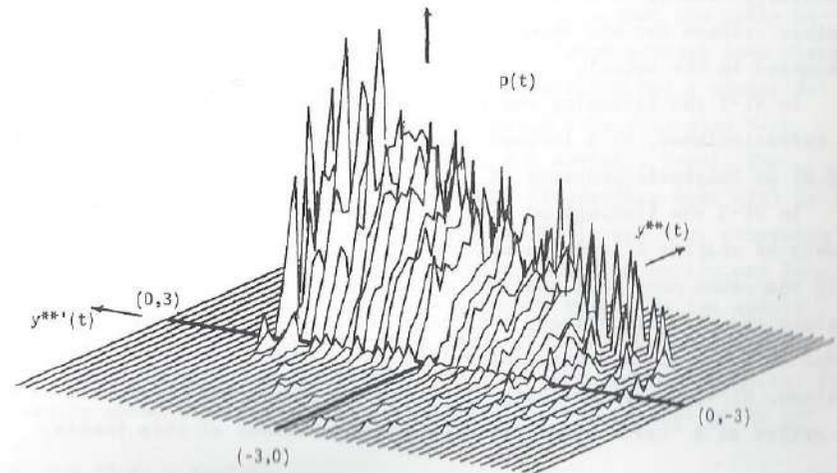


Fig. 5-8 scales as in Fig. 5-3

apter VI FINAL CONSIDERATIONS

-1. Introduction

The main reason for a need to revise the firing model (chapter V) was the experimentally observed skewness of the firing probability $p(t)$. This phenomenon could be observed in many publications concerning auditory nerve fibre data (see for some examples V-6), though it was not always as clearly present as in our result given in chapter V. As far as we know, the phenomenon did not attract systematic attention. That skewness is not always present (even if we restrict ourselves to low frequency stimulus situations) may be due to imperfect quality of the recorded impulses, which leads to smearing out of the histogram that estimates the firing probability. For instance, our own observation is that often during the recording, the physical shape of the spikes changes gradually or abruptly, inevitably leading to a smeared out estimate of the firing probability. Another reason for the absence of skewness, or only a small effect, is discussed in the sequel.

In VI-2 the following possible causes of the skewness are discussed: a) refractoriness, b) a lowpass filtering effect, c) a quick adaptation and d) an intrinsic property of the firing model.

In VI-3 the findings of VI-2 are discussed in connection with the models of chapter III. The conclusions are that for the non-Weiss models only the cause mentioned under c, a hypothetical fast adaptation, could explain the skewness effect, but that for a properly defined Weiss firing model d) applies. This surprising finding is speculated upon in the last section, VI-4, where remaining questions are raised and a proposal is made to arrive at a "best" model based on the conclusions of this thesis.

VI-2. Possible causes of skewness

a) Refractoriness.

In principle the refractory mechanism could account for the skewness. When a spike occurred in the first half of an excitatory lobe, another spike in the second half is highly unlikely (for a CF of 500 Hz a lobe is 1 msec in duration so that refractoriness would play a dominant role). Consequently, the firing probability will be greater in the first half lobe than in the second. We could call this a "virtual" influence of the time derivative of the excitatory signal on this firing probability since in the first half lobe the derivative is positive while in the second it is negative. As a consequence the mean derivative at the spike moments will be positive, though the derivative is not actively involved in the spike generation process. However, the shape of the revcor function is not given by (5-p) in this case. We will come back to this later, first we want an estimate of the skewness effect.

We will derive a quantitative measure of the amount of skewness.

For ease of calculation we use Duifhuis' assumption that the spike moments can be interpreted as an inhomogeneous Poisson process (see chapter III-2). First we calculate the firing probability for a sinusoidal input signal. Without refractoriness the firing rate will also vary sinusoidally; in normalised notation: $r(t) = c \sin(\pi t)$, $0 < t < 1$. The property of refractoriness can be introduced by assuming that $r(t)$ is equal to $c \sin(\pi t)$ when no spike occurred in $[0, t)$, and that otherwise $r(t) = 0$. This means absolute refractoriness during an excitatory lobe and complete recovery during the non excitatory phase. We can easily calculate the rate corresponding to a sinusoidal stimulus $c \sin(\pi t)$, $0 < t < 1$. Let $p_0(0, t)$ be the probability that no spike occurs in $[0, t)$. It is easily shown that the firing probability or rate is then given by:

$$r(t) = p_0(0, t) c \sin(\pi t). \tag{6-a}$$

Because of the Poisson assumption, $p_0(0, t)$ depends on the sinusoidal input signal in the following way:

$$p_0(0, t) = e^{-\int_0^t c \sin(\pi \tau) d\tau} \tag{6-b}$$

Thus the rate $r(t)$ is given by

$$r(t) = c e^{-c \int_0^t \sin(\pi\tau) d\tau} \sin(\pi t) = c e^{-(c/\pi)(\cos(\pi t) - 1)} \sin(\pi t) \quad (6-c)$$

The distortion factor $e^{-(c/\pi)(\cos(\pi t) - 1)}$ decreases monotonically in $[0, 1]$; thus the effect of absolute refractoriness is indeed a skewness of the firing probability relative to the excitatory signal. Is this effect large enough to explain the experimentally observed skewness? In order to answer this question we need to determine c in (6-c).

Let us first reconsider the problem: we wish to determine if the skewness of $p(t)$ in the narrow band noise case (namely the data of unit nr. 750712) can be explained by refractory effects. When the upward going deflections of the excitatory narrow band signal are replaced by purely sinusoidal deflections of amplitudes equal to those of the lobes, then with the aid of (6-c) we should obtain a fairly accurate idea of the "refractory skewness". The most pronounced skewness would be expected for the greatest lobes; the distortion factor becomes more effective with increasing c . In the data from unit nr. 750712 the maximum firing probability per lobe is about 24% (in a continuous sinusoidal stimulus situation this would mean about 145 spikes per second). Then $p_0(0,1) = 0.76$

$$e^{-c \int_0^1 \sin(\pi\tau) d\tau} = e^{-2c/\pi} = 0.76, \text{ which}$$

so that from (6-b) we find: $c = 0.43$. In Fig. 6-1, the excitatory sinusoidal wave form is depicted together with the corresponding firing probability.

Comparing Fig. 6-1 with Fig. 5-5, it is evident that the skewness is much smaller than the experimental one. The fact that $p_0(0,1) = 0.76$ means that the firing probability in the tail of a lobe is at least 0.76 times that in the front. Fig. 5-8 clearly shows a much greater depression of $p(t)$ for the highest $y^{**}(t)$ values when going from large positive $y^{**}(t)$ values to large negative ones (corresponding with fronts and tails of the lobes respectively). In order to explain the experimentally observed skewness by refractory effects, one would require unrealistic large firing rates. For instance, a value of $p_0(0,1) = 0.2$ means a firing probability per lobe of 80% which means 400 spikes per second

in a unit with a CF of 500 Hz.

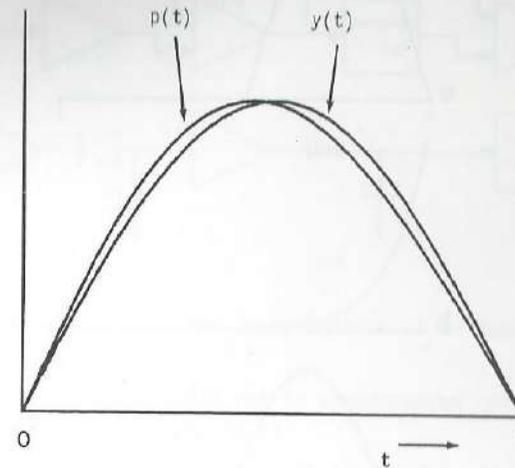


Fig. 6-1 vertical scale: arbitrary

The more realistic assumption of relative refractoriness, rather than absolute refractoriness, will further weaken the effect of this kind of skewness. We must conclude that, though refractoriness plays a role, it can not account for all of the observed skewness of $p(t)$. Also, though "refractory skewness" can qualitatively be described as "an influence of $y'(t)$ on the firing probability", this does not mean that, even approximately, a functional relationship $p(t) = F(y(t), y'(t))$ holds; the relative depression of $p(t)$ in the second half of a large lobe is greater than in a small one. Thus the results of V-7 are not applicable to this case, but anyway, the effect cannot explain the skewness phenomenon, and we leave it here.

- b) Skewness may result from a low pass filtering after the rectifier. So in Johannesma's model it will be present. In order to get an idea of the extent of this kind of skewness, we first look at Fig. 6-2. It shows
 - a) one period of a rectified sinusoid with a frequency of 1 kHz,
 - b) the impulse response of a simple RC low pass filter having a time constant of 100 μ sec and
 - c) the result of filtering the signal of a).

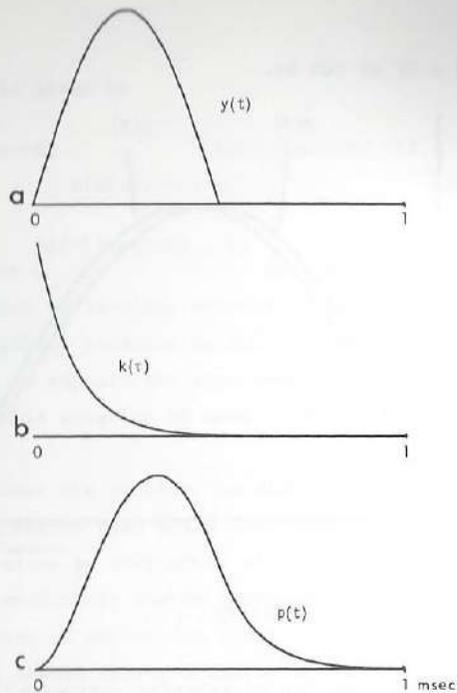


Fig. 6-2

Clearly from Fig. 6-2, we expect a $p(t)-y(t)-y'(t)$ plot to display a smearing out of the firing probability towards negative $y'(t)$ values. Is the resulting asymmetry visible in Fig. 6-2c sufficient to explain the skewness of the $p(t)$ versus $y^*(t)-y^{*'}(t)$ plot of Fig. 5-8?

To demonstrate the effect a hardware model was constructed, see Fig. 6-3.

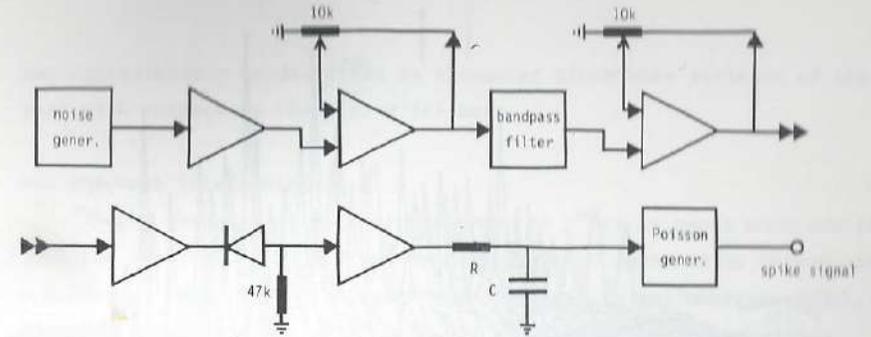


Fig. 6-3 Hardware model for the investigation of a lowpass filtering effect.

noise generator: H.P. type 3722A set to pseudorandom noise with a bandwidth of 5 kHz, $n=13$

bandpass filter: third of octave at 1008 Hz. Wandel und Goltermann
 amplifiers: operational Manifold RP-F. Philbrick.

poisson generator: designed by C. Kruidenier. Consists of a maximum length sequence generator driven by a clock pulse generator. The frequency of the clock was proportional to the input value; for negative values of the input, no pulses were generated.

In the RC part, the C was given different values in order to study the effect of the time constant on the skewness. The bandpass filter was a $\frac{1}{3}$ -oct filter with a central frequency of 1008 Hz. Since the output of the model consists of pulses, the simulation procedure as developed for the experimental data could directly be applied to this situation. In Fig 6-4 we see the result of the procedure applied to the model with $C = 0 \mu F$.

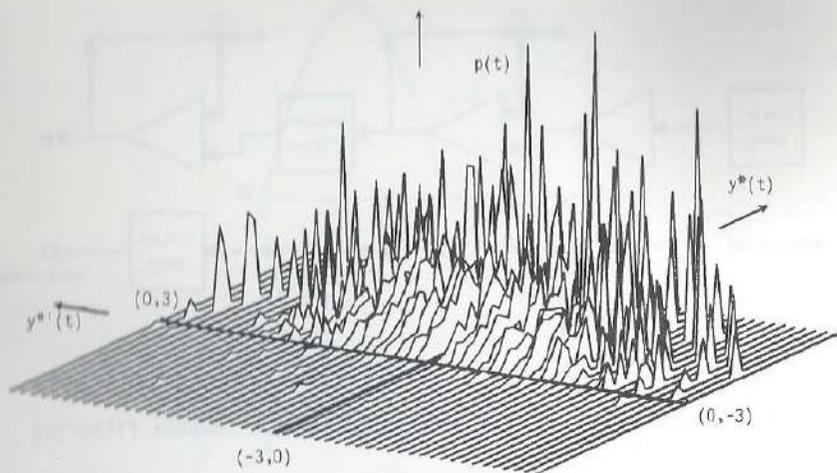


Fig. 6-4 same scales as in Fig. 5-4

As expected, no systematic dependence of $p(t)$ on $y^{*'}(t)$ is present. Refractory- nor low pass filtering effects nor adaptation plays a rôle, thus the only non linearity is the instantaneous one which, as has been discussed in chapters IV and V, has no effect on the shape of the cross correlation function. (Indeed, a direct check showed that $h^{*}(\tau) \equiv h(\tau)$.)

In the next figure, Fig. 6-5, however, we clearly see the effect of low pass filtering. The time constant of the RC part was taken 33 μ sec. The firing probability $p(t)$ differs from the rectified band pass filtered input signal to the effect that individual lobes are delayed and smeared out. The effect is somewhat less than that displayed in Fig. 6-2; there the time constant was 100 μ sec. When looking at Fig. 6-5, three points of interest are readily seen:

- 1) a smearing out of $p(t)$ at negative $y^{*'}(t)$ values is clearly visible.
- 2) though $y(t)$ leads with respect to $p(t)$, from this plot we can deduce that $y^{*}(t)$ lags: this phase shift of $y^{*}(t)$ with respect to $y(t)$

can approximately be described as a counter clock wise rotation of the plot with respect to the $y(t)-y'(t)$ axes.

But the most interesting is:

3) the skewness of $p(t)$ with respect to $y^{*}(t)$ is again much smaller than the experimental one. The skewness could be enhanced by introducing a different RC part with a longer time constant in the hardware model, but than the smearing out of $p(t)$ gets too pronounced to be comparable with the experimental result of Fig. 5-8.

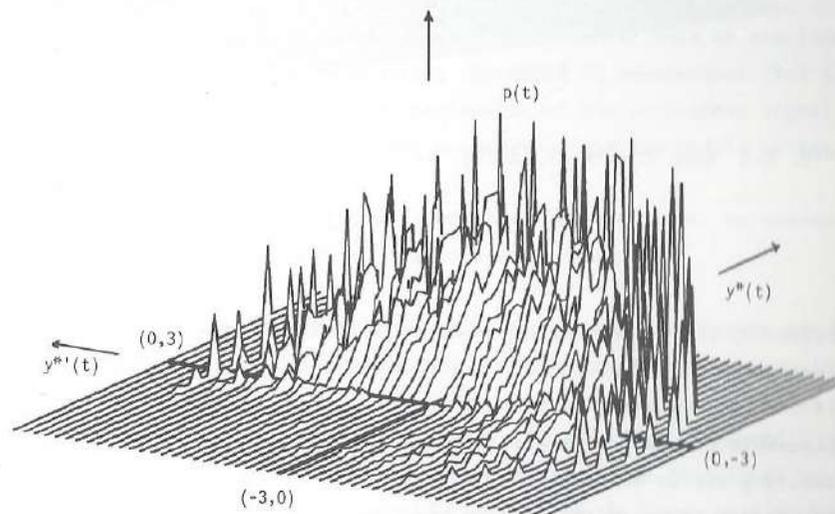


Fig. 6-5 same scales as in Fig. 5-4

Now it is time for a new experimental result: in Fig. 6-6 the result of the analysis procedure for unit nr. 750873 is displayed. In this plot $p(t)$ is plotted in the $y^{*}(t)-y^{*'}(t)$ plane; no rotation was done, so direct comparison with Fig. 6-5 is possible.

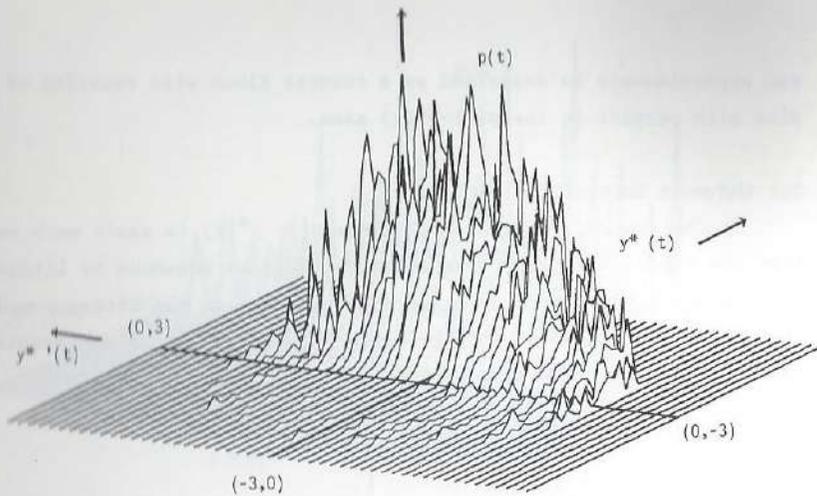


Fig. 6-6 same scales as in Fig. 5-4

Evidently the experimentally observed skewness is much more pronounced. Apart from that discrepancy, the similarity is striking. The conclusion is that Johannesma's lowpass filter K just before the firing model is a plausible element. (See also Eggermont (1972) in this context.) That this smearing out is less evident in the plot of unit nr. 750712 (Fig. 5-8) may be due to the difference in resonance frequencies: 580 Hz for the latter, compared to about 800 Hz for unit nr. 750873. As for the skewness, the conclusion is that a low pass filtering effect cannot explain it.

e) The third possibility is a fast adaptation.

If, during each excitatory lobe, the "effectiveness" of the stimulating wave form continuously decreases, then a skewness of $p(t)$ will result. With properly chosen parameters it should be possible to explain the experimentally observed skewness effect completely. Although such a rapid effect is as yet unknown, it cannot be excluded. (The time constant associated with fast adaptation as found in chapter II is too large.)

d) The last possibility that the skewness be an inherent property of the firing model is discussed in the next section.

VI-3. The skewness effect and the models

The conclusion of VI-2 was, that only either a fast adaptation, or an inherent property of the firing model could explain the skewness. We will only treat the latter possibility, because this adaptation is as yet unknown.

Since the firing model of Duifhuis and the SIPIT model of Johannesma are instantaneous devices (see chapter IV), they can be excluded as candidates for the skewness effect and we will concentrate on Weiss' firing model.

Concerning Weiss' firing model, from experimental data we concluded that neither a fixed nor a fluctuating threshold is appropriate ($V=3$ and $V=4$). The problem is that the time derivative of the excitatory signal evidently plays a role which can not be explained by the simple trigger as defined in those sections.

If we define a "threshold crossing" more carefully, however, we automatically (see below) arrive at the desired dependence.

Let us assume that in the process of spike generation two signals are involved: the excitatory signal $y(t)$, and a random fluctuating threshold signal $b(t)$. A spike is initiated when $y(t)$ crosses $b(t)$ as specified below. We arrive at a threshold model with a fluctuating threshold, which, however, is significantly different from the one introduced in V-6. There we described the threshold fluctuations in terms of a probability distribution function; only the probability density of a pertinent threshold value was given. The omission of not defining the temporal behaviour of the threshold value, or the autocorrelation properties of the fluctuations, leads to conceptual problems with respect to defining the threshold crossings. If we introduce a threshold signal, $b(t)$, and assume it to be a sample function of a differentiable stochastic process, however, then we have no difficulty defining crossings and thus spike moments: a moment, t , is a spike moment when the excitatory signal $y(t)$ crosses the threshold signal $b(t)$ "in upward direction" that is $y(t) = b(t)$ and $y'(t) > b'(t)$. The properties of the threshold signal $b(t)$ completely determines the "character" of the spike generator. It is not the aim of this thesis to further explore this connection; we will con-

tent ourselves with showing that $b(t)$ signals exist which, in a hardware model (see below), lead to results which very satisfactory simulate our auditory nerve data.

What properties should this threshold signal, $b(t)$, have? First $b(t)$ must be positive. From Fig. 5-8 we see that spikes do not occur when $y(t) < 0$. Thus in terms of the firing model introduced in this section this means that $y(t)$ only meets $b(t)$ in the upper half plane, so clearly $b(t) > 0$ must hold. It is readily seen that the required dependence of $p(t)$ on $y'(t)$ is qualitatively present: when $y(t)$ moves in positive direction ("towards $b(t)$ ") a crossing will be more probable than when $y'(t) < 0$. In the latter case $b(t)$ must "surpass" $y(t)$ in order to cause a crossing. In order to illustrate the principle a hardware model (Fig. 6-7) was used and $b(t)$ was taken to be low pass filtered noise (see Fig. 6-7 for details).

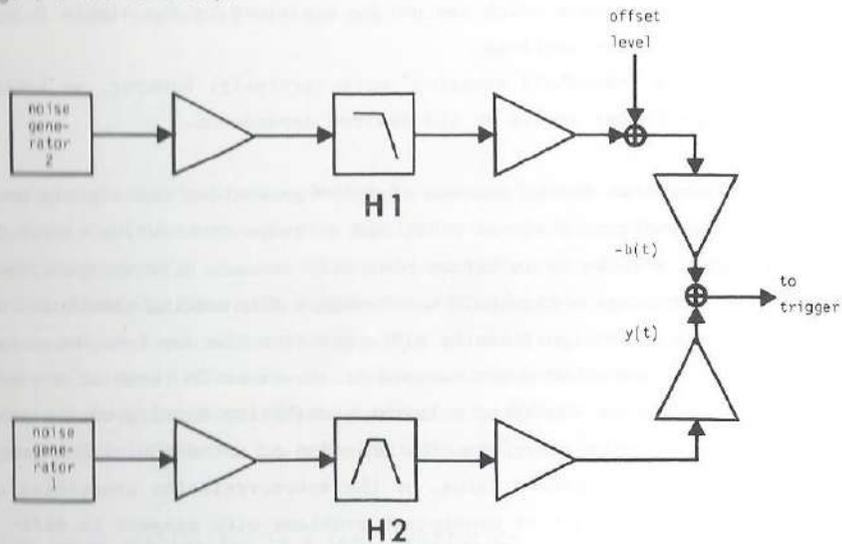


Fig. 6-7 noise generator 1: H.P. type 3722A. Bandwidth 5 kHz, $n=13$.
 noise generator 2 generated pink noise.
 H1: Lowpass filter; set to 2 kHz cut off (designed in our lab. rejection: 60 dB/oct over 400 Hz)
 H2: third of octave filter at 449 Hz. Wandel und Goltermann.
 The signal to the trigger had an r.m.s. value of 1V and had an offset of -1.2V. The level of the trigger was set at 0V.

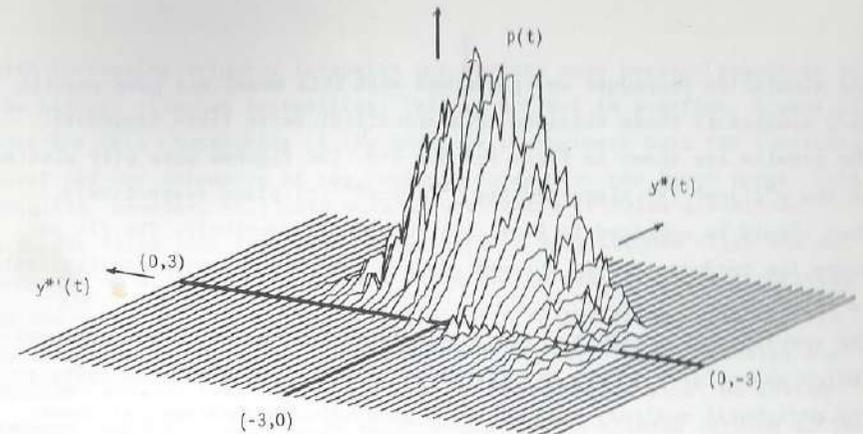


Fig. 6-8 same scales as in Fig. 5-4

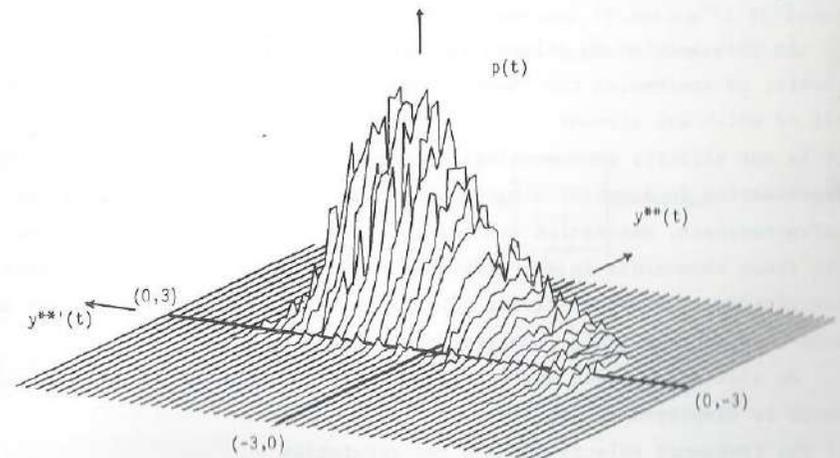


Fig. 6-9 same scales as in Fig. 5-4

The simulation procedure was performed with this model and gave results very similar to those obtained from acoustical nerve fibre responses. The results are shown in Figs. 6-8 and 6-9. The figures show $p(t)$ plotted in the $y^*(t)-y^{*'}(t)$ plane and the $y^{**}(t)-y^{**'}(t)$ plane respectively. They should be compared to Figs. 5-4 and 5-8 respectively. The fit between the hardware models outcome and the experimental data qualitatively is very good.

The conclusion is, that the fluctuating threshold model introduced here (which essentially is the trigger firing model of Weiss) could apply to the peripheral auditory systems' spike generating mechanisms, at least when an appropriate threshold behaviour is inserted. The experimentally observed skewness can completely be attributed to an inherent property of that firing model.

The Peripheral Auditory System model

In this section an attempt is made, applying the conclusions of this thesis, to synthesize the "best" model. The model consists of complexes all of which are already introduced in the preceding chapters. It is not strictly phenomenological because it consists of building blocks representing fundamental sensory mechanisms like spontaneous activity, refractoriness, saturation and adaptation. Therefore, the integration of all these components in the ultimate PAS-model should provide a complete description of neuronal firing behaviour to stimulus situations that are more complex than those investigated in this thesis.

We will now proceed discussing the various complexes of the PAS-model, which is displayed in Fig. 6-11.

1) The frequency selectivity and the saturation

In chapter III it was concluded that modelling the frequency selectivity via one sharp linear filter is a quite satisfactory solution. When, however, we want to include a Weiss-like firing model in the PAS-model, we run into trouble as already mentioned in III-5. This trouble can be cured by means of a different modelling of the frequency selective mechanism, as will be shown below. The problems with Weiss' firing model are all connected to the signal to noise ratio at the input of the trigger. If, namely, this ratio increases

with increasing stimulus intensity one expects near perfect synchrony at the highest stimulus intensities. This is not met in practice. A possible cure for this shortcoming of the model is to increase both the threshold level and the intensity of the internal noise with the input level. This solution, however, will only work in stationary stimulus situations; e.g. the first peak in a PSTH belonging to a very intense click can not properly be simulated. As already discussed in chapter III a saturation in the sense that the signal to noise ratio at the trigger input stays constant above a certain stimulus level above threshold eliminates the problem. However this saturation must be of a special kind. An instantaneous nonlinear saturation which would yield a clipped version of the linearly filtered input signal at the highest stimulus levels is unacceptable. Even at the highest intensities such a clipping can not be observed in actual firing patterns (Kiang (1965), Rose et al. (1967), Gray (1967)).

An instantaneous saturation mechanism which does not display this undesired clipping phenomenon is provided by Pfeiffer (Pfeiffer (1970)). It consists of a linear filter ("the basilar membrane filtering"?) followed by a saturating instantaneous nonlinearity and again a linear filter ("the second filter"?) (Fig. 6-10).

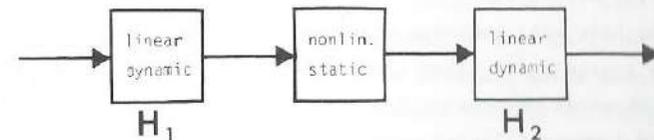


Fig. 6-10

This nonlinear filter was meant by Pfeiffer to model the phenomenon of two-tone suppression, but it can be used as a valuable extension of Weiss' model.

The properties of this nonlinear filter are subject of our present research. Preliminary results indicate that when, in the models, the linear filter H (Figs. 2-1, 2-2 and 2-3) is replaced by Pfeiffer's nonlinear filter of Fig. 6-10, the simulated signal still has the desired properties. The receiver function will be proportional to the impulse response of the cascade of the filters H_1 and H_2 (in Johannesma's model the lowpass filter K is a slightly complicating factor) and the simulating signal still is a remarkably good approximation to the nonlinearly filtered signal. As said in chapter III, however, the "second filter" situation is unclear and consequently this solution is speculative. The resulting saturation is favourable because of two reasons. First, Weiss' model can be incorporated in the PAS-model. Second, saturation in this early stage of the model is in accordance with the physiologically plausible assumption that already the receptor potential must have undergone the main saturation (III-5).

2) The noise inputs $n_1(t)$ and $n_2(t)$

The noise source $n_1(t)$ takes care of the spontaneous activity; the motivation for its location just before the feed back loop is that in this way it provides a correct modelling of post stimulatory depression of the spontaneous activity. See Koldewijn (1973) for simulation results. The physiological correlate of the noise signal $n_1(t)$ is: spontaneous fluctuations of the hair cell membrane potential plus synaptical noise.

The noise source $n_2(t)$ takes care of the threshold fluctuations; it can cause the skewness phenomenon described in this thesis. In our simulations we used lowpass filtered pink noise. The simulation results of Figs. 6-8 and 6-9 were obtained with a cut off frequency of the lowpass filter H_1 in Fig. 6-7 of 2 kHz. The physiological correlate is the membrane potential fluctuations in the primary auditory nerve fibre at the spike originating site.

3) The nonlinearity G_2 and the feedback loop

This complex takes care of the adaptation. As a part of Johannesma's model, simulations were performed by Koldewijn (1973). The time constant in the feedback loop was taken 33 msec. This value probably is too long; it is even questionable whether a simple exponentially decaying function adequately describes the adaptation. However, the principle of this mul-

tiplicative feedback loop is supported by our own unpublished observations. As for the physiological aspects, there is considerable evidence (Eggermont (1972), Prijs (1979), in guinea pig) that adaptation of the peripheral hearing is mainly situated in the hair cell - primary afferent synapse. According to Eggermont it should be modelled as a dynamic decrease in excitation during stimulation. The solution adopted here from Johannesma's model is a simplified version of Eggermont's model.

4) The firing model

In the PAS-model of Fig. 6-11 the firings are generated by a firing model consisting of a lowpass filter, K , followed by a trigger, T . The firings are fed back to T via a lowpass filter and subtracted from its input in order to take care of refractory properties. A noise source, $n_2(t)$, causes random fluctuations at the input of T .

In VI-3 a simulation was described in which only the trigger and the noise source were involved. This simulation proved successful and was the reason for incorporating this firing model in the PAS-model. The influence of K was also studied in VI-3. As for the influence of the refractoriness mechanism, this was not studied. In the case of low firing rates, no influence should be expected: when the value for the time constant of the lowpass filter in the feedback loop is taken to be 0.3 msec (Weiss (1966)) then the "refractoriness effect" is virtually zero in 1 msec.

In performing simulations with the hardware Weiss model of Fig. 6-7, it was observed that the resulting skewness of $p(t)$ strongly depends on the properties of the threshold signal $b(t)$, on the relative amplitude of $y(t)$ and on the central frequency of the bandpass filter. A systematic study of these dependencies together with systematic investigation of the dependency of the experimental skewness on the CF, the stimulus intensity, the spontaneous activity and the firing rate might give conclusive evidence about the validity of this solution. The physiological considerations for the firing model are: K represents the lowpass filtering effect of electrotonic conduction along a nerve fibre; the trigger, T , represents the neuronal action potential triggering; the refractory mechanism takes care of the temporarily suppressed excitability after a firing. The noise source, $n_2(t)$, mirrors the spontaneous potential fluctuations that exist across neuronal membranes.

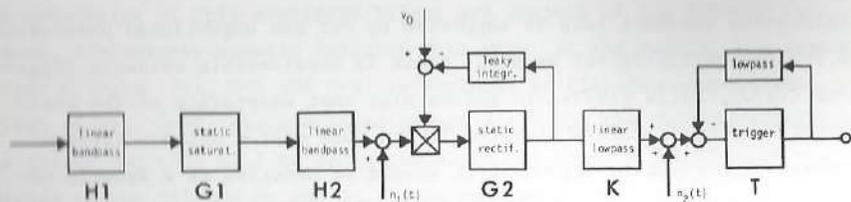


Fig. 6-11 the PAS-model

Now we end the discussion of the PAS-model with an important conclusion about Johannesma's SILIT-model. It is readily seen that the SILIT firing model and Weiss' firing model are functionally equivalent. The SILIT-model consists of a lowpass filter followed by a trigger, with a feedback loop across the two. When the lowpass filter is put in the feedback loop, Weiss' model emanates. Thus these two models are functionally equivalent (i.e. cannot be distinguished via their input-output conduct) with the proper noise source adjustment. The conclusion must be that also the SILIT firing model can explain the skewness effect. According to us, though, the time constant of the leaky integrator should be taken much shorter than the 3.3 msec which Koldewijn (1973) used in his simulation, say 0.3 msec. Only physiological measurement at the site of the spike generation could clarify which of the two models is closer to reality.

VI-5 Concluding remark about the Volterra-Wiener approach

As mentioned in chapter III, the revcor function and the first order V-W kernel are essentially the same. One might ask why the seemingly logical next step in the V-W approach, of including the second order kernel in the analysis, has not been taken. We did measure a few second order kernels, but the result was disappointing. The reason for that is the following. As has been shown in chapter II, a very reasonable model for the peripheral auditory system (in the low CF case) when using GWN as the input signal and considering as the output signal the firing probability in one nerve fibre, is simply the cascade of a bandpass filter and an instantaneous nonlinearity, a rectifier. For that simple system, the second order V-W kernel is (via (4-d) with $k(\tau) = \delta(\tau)$) proportional to $h(\tau) \cdot h(\sigma)$, where $h(\tau)$ is the impulse response of the bandpass filter. Indeed, no systematic difference could be demonstrated between the measured estimate of the second order kernel and this predicted function. So the analysis according to the V-W approach was abandoned and the less formal approach (using only the first order kernel) as found in this thesis, was followed.

SUMMARY

In this thesis we are concerned with the P(eripheral) A(uditory) S(ystem). This system transforms a sound stimulus into trains of nervous impulses in the ensemble of single auditory nerve fibres. Several models exist describing this transformation. This thesis attempts to probe the applicability of three of the most important models. The models are analysed mathematically using a correlation method. Two restrictions are set by the method: first, only the responses of fibres with low characteristic frequency (<5 kHz) can be analysed; second, the stimulus involved is wideband and stationary. For verification purposes experimental data were collected from neuronal units of the cat's auditory nerve. It was demonstrated that none of the models correctly describes the experimental facts.

A synthesised model is proposed that should describe the experimental facts adequately. It is based on the following conclusions and considerations:

- a) The frequency selectivity of the PAS as derived from tuning curve properties can very well be described as a linear filtering for a dynamic range of at least 40 dB.
- b) A firing model should contain a threshold mechanism.
- c) A lowpass filtering is probably present just before spike initiation.
- d) The firing probability in a primary auditory fibre can, in first approximation, be described as a rectified linear transform of the input signal.
- e) In a closer analysis the firing probability appears to be governed by two linear functionals on the input signal wave form. In this most general case the first order Volterra-Wiener kernel, c.q. the revcor function, is a linear combination of the filter response function and its time derivative.
- f) The deviations from the linear transform approximation can largely be attributed to an inherent property of the firing model. Other factors are of minor importance in this respect.
- g) Probably saturation is not provided by an instantaneous nonlinearity after the frequency selectivity.
- h) The SILIT firing model of Johannesma (1971) and Weiss' (1966) firing model are functionally equivalent. Both models are plausible solutions for the spike generating mechanism of the PAS.

REFERENCES

- Anderson, D.J., Rose, J.E., Hind, J.E., Brugge, J.F.: Temporal position of discharges in single auditory nerve fibres within the cycle of a sine-wave stimulus: frequency and intensity effects. *J. Acoust. Soc. Amer.* 49, 1131-1139 (1971).
- Auerbach, A.A.: Transmitter Release at Chemical Synapses. In: Pappas and Purpura (1972).
- Békésy, G. von: Experiments in hearing. New York: McGraw-Hill 1960.
- Boer, E. de: Correlation studies applied to the frequency resolution of the cochlea. *J. Aud. Res.* 7, 209-217 (1967).
- Boer, E. de, Jongkees, L.B.W.: On cochlear sharpening and cross-correlation methods. *Acta Oto-Laryng.* (Stockh.) 65, 97-104 (1968).
- Boer, E. de, Kuyper, P.: Triggered correlation. *IEEE Trans. Biomed. Eng.* 15, 169-179 (1968).
- Boer, E. de, Jongh, H.R. de: Computer simulation of cochlear filtering. *Proc. 7th Int. Congr. on Acoustics*, 393-396. Budapest: Akademiai Kiado 1971.
- Boer, E. de: Spectral transformations by an infinite clipper. *J. Acoust. Soc. Amer.* 60, 960-963 (1976).
- Castillo, J. del, Katz, B.: Quantal components of the end-plate potential. *J. Physiol. (Lond.)* 124, 560-573 (1954).
- Dallos, P.: *The Auditory Periphery*. Acad. Press. New York 1973.
- Davis, H.: Biophysics and physiology of the inner ear. *Physiol. Rev.* 37, 1-49 (1957).
- Davis, H.: A mechano-electrical theory of cochlear action. *Ann. Otol. (St. Louis)* 67, 789-802 (1958).

Davis, H.: Hearing and Deafness. A guide for Laymen. Murray Hill Books.
New York. 1947.

Duifhuis, H.: A tentative firing model for the auditory receptor.
IPO-APR. 5, 18-24 Eindhoven (1970).

Duifhuis, H.: Perceptual Analysis of Sound (Thesis). Eindhoven: 1972

Eggermont, J.J.: De tijdsafhankelijke eigenschappen van cochleaire actie-
potentialen bij de cavia (Thesis). Leiden: 1972.

Evans, E.F.: The effects of hypoxia on the tuning of single cochlear
nerve fibres. J. Physiol. (Lond.) 238, 65-67 (1974).

Evans, E.F.: Cochlear Nerve and Cochlear Nucleus. In: Handbook of Sensory
Physiology Volume V/2, 1-108. Ed: Keidel, W.D. and Neff, W.D.
Springer Verlag Berlin 1975

Geisler, C.D.: A model of the peripheral auditory system responding to
low frequency tones. Biophys. J. 8, 1-15 (1968).

Gisbergen, J.A.M. van, Grashuis, J.L., Johannesma, P.I.M., Vendrik, A.J.H.:
Single cells in the auditory system investigated with stochastic
stimuli. Proc. XXV Congress on Physiol. Sciences, München (1971).

Gisbergen, J.A.M. van: Characterization of responses to tone and noise
stimuli of neurons in the cat's cochlear nuclei (Thesis). Nijmegen:
1974.

Granit, R., Kernell, D., Shortess, G.K.: Quantitative aspects of repeti-
tive firing of mammalian motoneurons, caused by injected currents.
J. Physiol. 168, 911-931 (1963).

Grashuis, J.L.: The Pre-Event Stimulus Ensemble. An analysis of the
stimulus-response relation for complex stimuli applied to auditory
neurons (Thesis). Nijmegen: 1974.

Gray, P.R.: Conditional probability analysis of the spike activity of
single neurons. Biophys. J. 7, 759-777 (1967).

Hung, G., Stark, L.: The Kernel Identification Method (1910-1977).
Review of theory, Calculation, Application, and Interpretation.
Math. Biosc. 37, 135-190 (1977).

Ishii, Y., Matsuura, S., Furakawa, T.: An input-output relation at the
synapse between hair cells and eighth nerve fibres in goldfish.
Jap. J. Physiol. 21, 91-98 (1971).

Johannesma, P.I.M.: Stochastic neural activity (Thesis). Nijmegen: 1969.

Johannesma, P.I.M.: Dynamical aspects of the transmission of stochastic
neural signals. First European Biophysics Congress. Baden near
Vienna. Austria: 1971

Johannesma, P.I.M.: The Pre-Response Stimulus Ensemble of Neurons in the
Cochlear Nucleus. In: Hearing Theory 1972, 58-69. Eindhoven: IPO
1972.

Johannesma, P.I.M.: Lecture Notes on Neural Information Processing.
Aarhus Universitet. Matematisk Institut. Danmark (1976).

Johnstone, B.M., Taylor, K.J., Boyle, A.J.: Mechanics of the guinea pig
cochlea. J. Acoust. Soc. Amer. 47, 504-509 (1970).

Jongh, H.R. de: About coding in the VIIIth nerve. In: Hearing Theory
1972, 70-77. Eindhoven: IPO 1972.

Jongh, H.R. de: Analysis of a spike generator. In: Proceedings of the
3d IPAC Symposium on Identification and System Parameter estimation,
Ed: Eyekhof, P. den Haag, 607-612 (1973).

Katz, B., Miledi, R.: Input output relations of a single synapse. Nature
212, 1242 (1966).

Kiang, N.Y.-s, Watanabe, T., Thomas, E.C., Clark, L.F.: Discharge patterns of single fibres in the cat's auditory nerve. Cambridge, Mass.: M.I.T. Press 1965.

Kiang, N.Y.-s, Moxon, E.C.: Physiological considerations in artificial stimulation of the inner ear. *Ann. Otol.* (1972).

Klinke, R., Pause, M.: The performance of a primitive hearing organ of the cochlea type: primary fibre studies in caiman. In: *Psychophysics and Physiology of Hearing*. Ed: Evans, E.F., Wilson, J.P. Academic Press, New York (1977).

Koldewijn, G.J.R.: Modelstudie aan het perifere auditief systeem. Internal report, Lab. Med. Physics and Biophysics. Nijmegen: 1973.

Loewenstein, W.R.: Principles of Receptor Physiology. Handbook of sensory physiology I. (1970) Springer verlag Berlin 1971.

Martin, A.R.: Quantal nature of synaptic transmission. *Physiol. Rev.* 46: 51-66 (1966).

Nakajima, S.: Adaptation in stretch receptor neurons of crayfish. *Science* 146, 1168-1170 (1964).

Pappas, G.D., Purpura, D.P.: Structure and Function of Synapses. Raven Press, New York (1972).

Pfeiffer, R.R.: A model for two-tone inhibition of single cochlear nerve fibres. *J. Acoust. Soc. Amer.* 48, 1373-1378 (1970).

Prijs, V.: Thesis. Leiden. to be published.

Price, R.: A useful theorem for nonlinear devices having Gaussian inputs. *IRE Trans. Inf. Theory* IT-4, 69-72 (1958).

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

Rhode, W.S.: An investigation of cochlear mechanics using the Mössbauer effect. In: *Basic Mechanisms in Hearing*, 49-63. Ed. Møller, A.R. New York: Academic Press 1973.

Robertson, D.: Possible relation between structure and spike shapes of neurones in Guinea Pig cochlear ganglion. *Brain Research* 109, 487-496 (1976).

Rose, J.E., Brugge, J.F., Anderson, D.J., Hind, J.E.: Phase-locked response to low-frequency tones in single auditory nerve fibres of the squirrel monkey. *J. Neurophysiol.* 30, 769-793 (1967).

Rose, J.E., Hind, J.E., Anderson, D.J., Brugge, J.F.: Some effects of stimulus intensity on response of auditory nerve fibres in the squirrel monkey. *J. Neurophysiol.* 34, 685-699 (1971).

Russel, I.J., Sellick, P.M.: The tuning properties of cochlear hair cells. In: *Psychophysics and Physiology of Hearing*. Ed: Evans, E.F., Wilson, J.P. Academic Press, New York (1977).

Siebert, W.M.: Some implications of the stochastic behaviour of primary auditory neurons. *Kybernetik* 2, 206-215 (1965).

Siebert, W.M.: Stimulus transformations in the peripheral auditory system. In: *Recognizing Patterns*. Ed: Kolers, P.A., Eden, M. 104-133. M.I.T. Cambr. Mass. (1968).

Siebert, W.M.: Frequency Discrimination in the Auditory System: Place or Periodicity Mechanisms? *Proc. IEEE* 58, 723-730 (1970).

Slepian, D.: The one-sided barrier problem for Gaussian-noise. *Bell System Tech. J.* 41, 463-501 (1962).

- Spoendlin, H.: Structural basis of peripheral frequency analysis. In: Frequency Analysis and Periodicity Detection in Hearing, 2-36. Leiden: Sijthoff 1970.
- Tasaki, I.: New measurements of the capacity and the resistance of the myelin sheath and the nodal membrane of the isolated frog nerve fibre. Amer. J. Physiol. 181, 639 (1955).
- Verveen, A.A.: Axon diameter and fluctuation in excitability. Acta Morph. Neerl. -Scand. 5, 79-85 (1963).
- Weiss, T.F.: A model for firing patterns of auditory nerve fibres. Quarterly Progress Report No. 69, Research Lab. of Electr. M.I.T. Chambr. Mass., 217-223 (1963).
- Weiss, T.F.: A model of the peripheral auditory system. Kybernetik 3, 153-175 (1966).
- Wilson, J.P., Johnstone, J.R.: Capacitive probe measures of basilar membrane vibration. In: Hearing Theory 1972, 172-181. Eindhoven: IPO 1972.

I

Het neuronale vuurmodel van Weiss en het SILIT vuurmodel van Johannesma zijn functioneel equivalent. (Dit proefschrift).

II

Het meten van RC-tijden behorende bij biologische membranen, met behulp van microelectrodes, is een moeilijke zaak. Het gebruik van de zgn. capaciteitscompensatie is één van de redenen waarom de resultaten van deze metingen met veel reserve moeten worden beschouwd.

P.M. Sellick and I.J. Russell: Intracellular studies of cochlear hair cells. In: Electrical evoked activity in the auditory nervous system. Ed.: R. Naunton. Acad. Press. In press.

III

De vuurkans van een primaire auditieve zenuwvezel kan, in het geval van een stationaire brede band stimulus, met behulp van twee lineaire functionalen, werkend op de acoustische ingangssignaalvorm, benaderd worden.

IV

De door Oetinger en Hauser voorgestelde verbetering van het een-dimensionale cochleamodel door middel van het begrip "toegevoegde massa" is onderhevig aan bedenkingen. Het verbeterde model geldt slechts onder de voorwaarde dat de mechanische impedantie van de toegevoegde massa veel kleiner is dan die van het basilaire membraan.

R. Oetinger und H. Hauser (1961). *Acustica* 11,161-177

V

De verdediging die Hubbard, Llinás en Quastel aanvoeren voor het gebruik van het woord quantum in verband met transmitter release is zwak.

J.I. Hubbard, R Llinás, D.M.J. Quastel: Electrophysiological analysis of synaptic transmission. (blz. 117). Publ. E.Arnold Ltd. London. 1969.

VI

De reclameslogan "bloemen houden van mensen" moet ontstaan zijn vanuit een gebrekkig biologisch inzicht gepaard aan een, op historische gronden falsifieerbaar, idee van menselijke beminnelijkheid.

VII

Gezien de afhankelijkheid die de externe, zogenaamd onafhankelijke, registeraccountant van een bedrijf heeft ten opzichte van zijn broodheer, is de accountantshandtekening onder een jaarrekening zonder waarde.

VIII

Er dient een deugdelijke kwaliteitscontrole op tandartswerk te komen. De negatieve consequentie van deze maatregel, een toename in het tandarts tekort, zou meer dan gecompenseerd worden door een aanmerkelijke verbetering in de kwaliteit van het werk.

IX

Gezien het element van spelbederf dat gewezen staatssecretaris Klein ziet bij de invoering van de prikklok en gezien Tinbergens ideeën over een reciprook verband tussen beloning en werkplezier, kunnen wetenschappelijke werkers binnenkort hun salaris met geruster hart incasseren.

X

Wie op grote schaal wetsovertredingen wenst te plegen en zich op onrechtmatige manier wenst te verrijken, doet er, in verband met de strafmoraal van de staande en zittende magistratuur verstandig aan deze delicten in vennootschapsverband te plegen.

XI

Wie op grote schaal zijn medemens directe en indirecte schade wil toebrengen, kan dit als automobilist volstrekt ongestraft doen, d.m.v. lawaai, stank, bedreiging met letsel, aanjagen van doodsangst etc., tegenover voetgangers, fietsers en binnenstadbewoners.

XII

Het is verbazingwekkend dat men in conservatieve kring gekant is tegen democratisering en popularisering van het instituut "Hoger Onderwijs". Immers, deze leiden tot achteruitgang van de kwaliteit van het onderwijs en daarmee tot het minder geschikt worden van het diploma als objectief selectie criterium. Dit leidt op zijn beurt tot een revival van de kruiwagen bij het bestijgen van de maatschappelijke ladder.

XIII

De gewoonte, ook van niet gepromoveerde dokters, om hun naambordje met dr. te beginnen heeft meestal niet met plaatsgebrek op dat bordje te maken.