Vision Research Vol. 19, pp. 515 to 522
Pergamon Press Ltd 1979. Printed in Great Britain

PROBABILITY SUMMATION OVER TIME

ANDREW B. WATSON[1]
Department of Psychology, University of Pennsylvania, Philadelphia, PA 19174, U.S.A.

(*Received* 21 *June* 1978)

Abstract—Frequency-of-seeing and sensitivity–duration curves were collected for temporal signals of limited spectral extent. A comparison of the two sorts of data suggests that a stimulus is detected whenever the excursions of its linearly filtered, noise-perturbed temporal waveform exceed some fixed magnitude.

INTRODUCTION

The sensory response to light persists in time. This fact obliges us to understand how the temporally distributed effects of visual stimulation collectively determine threshold. A classical solution to this problem has been to suppose that the eye integrates the light signal over some interval of time, the *critical duration*. The light is visible whenever the integral exceeds some fixed value. The finding which motivated this idea was the apparent reciprocity between intensity and time for stimuli at threshold (Bloch, 1885; Graham and Margaria, 1935), but this result is equally well accommodated by a more general model. This model supposes the eye to operate in the vicinity of threshold as a linear temporal filter followed by a threshold mechanism, which responds only when the excursions of the filtered signal exceed some fixed magnitude (De Lange, 1952; Kelly, 1961; Sperling and Sondhi, 1968; Roufs, 1972. In this *peak-or-trough detector, temporal summation* occurs as a consequence of the integrating action of the filter. In yet another model the linear filter is followed by a squaring device, and the filtered, squared signal is integrated over some finite interval. In the version of this *power integrator* proposed by Rashbass (1970) the integration interval is about 200 msec; in that of Koenderink and van Doorn (1978) it is about 500 msec.

A feature common to all of these proposals is that they are deterministic: they do not attempt to represent the probabilistic nature of detection. We may imagine the variability of detection to be due to "noise" in the visual process. If the time scale of the noise fluctuations is long relative to that of the model, then it can properly be neglected. If the two scales are comparable, however, then the intrusion of noise may render the above models incapable of correctly predicting the visibility of long as well as short stimuli.

One way in which noise may influence the visibility of a temporally extended signal is through *probability summation over time*. To describe this process it is convenient to treat the continuous time interval containing the signal as a sequence of brief, finite intervals, or *instants*. In order to properly represent the signal waveform, the duration of each instant must be less than half the period of the highest frequency present in the signal. The presence of noise ensures that within each instant there is some probability that threshold will be exceeded, and so the overall probability that the signal is detected must take into account all of these momentary probabilities. Let P_i be the probability that threshold is exceeded during the instant beginning at time t_i. If the fluctuations of the noise are sufficiently rapid (if the autocorrelation function of the noise process is sufficiently narrow), then these probabilities will be independent from instant to instant. Then supposing that the signal is detected if and only if threshold is exceeded in at least one instant, the probability of detection, P, will be

$$P = 1 - \prod_i (1 - P_i). \tag{1}$$

One consequence of equation 1 is that for a stimulus for which at least some P_i are not zero, the probability of detection will continue to rise with increases in duration for as long as the signal may be prolonged. The same cannot be said for plausible versions of the models described above.

It will be useful to express this argument in quantitative terms. To do this we may make use of a model elaborated by Watson and Nachmias (1977). It resembles in its essentials models suggested by Stotter and Robson (1978), Graham (1977) and Tolhurst (1975) to describe probability summation over various dimensions of visual and auditory stimuli. The model, as sketched in Fig. 1, consists of a linear filter, an additive noise source, a threshold device, and a guess generator. An OR gate combines the outputs of the guess generator and threshold device. The output of the linear filter to an input $f(t)$ is $g(t)$. The threshold device is a non-linear element which responds with the value 1 whenever its input exceeds some threshold value, positive or negative. The guess generator responds with a 1 with probability γ on each trial, independent of what happens in the rest of the model. The OR gate responds with a 1, and the observer reports "yes, I detect the stimulus", whenever either of its inputs is a 1.

Two obstacles must be passed before the model may be used to predict the visibility of actual temporal waveforms. First, the manner in which the

[1] Present address: Physiological Laboratories, University of Cambridge, Cambridge CB2 3EG, England.

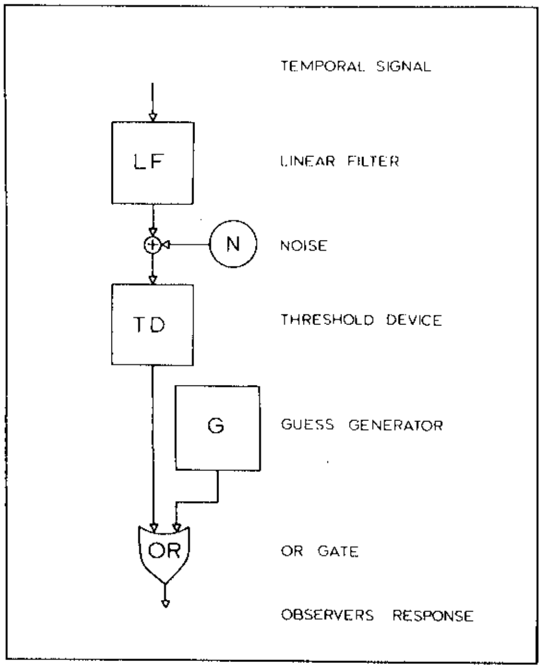

Fig. 1. Diagram of a model for the detection of temporal waveforms. Its components and properties are described in the text.

noise, threshold device, guess generator and OR gate act to determine the detectability of a filter response $g(t)$ must be quantitatively described. Second, the nature of the linear filter must be defined, so that $g(t)$ may be obtained from $f(t)$.

The first task is simplified by adopting the formula

$$P_i = 1 - \exp[-|g(t_i)|^\beta] \tag{2}$$

to describe the probability that the threshold device has an output of 1 within the instant beginning at time t_i. This expression is one version of a distribution function studied extensively by Weibull (1951). Several of its practical and theoretical virtues have been noted by Quick (1974) and Green and Luce (1975). Some examples of the function are shown in Fig. 2. Each incorporates a different value of β, the parameter which controls the slope of the function. This particular function has been chosen both because it is mathematically convenient and because, as will be demonstrated, it provides a very good description of experimentally determined psychometric functions.

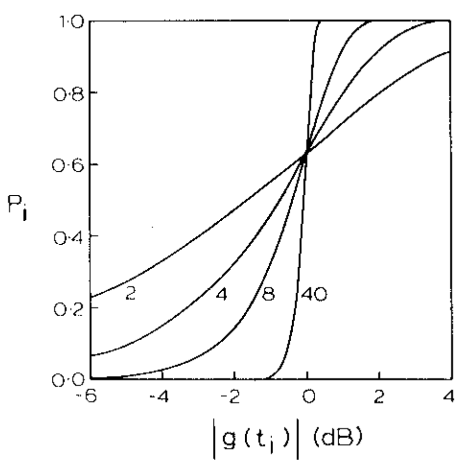

Fig. 2. Psychometric function described by equation 2. Each curve describes for the value of β indicated the probability that the threshold device will respond as a function of the instantaneous magnitude of the response $g(t)$.

The observer will respond "no" only if both the threshold device and the guess generator fail to respond during the observation interval. If the values assumed by the noise are independent from instant to instant, then the probability of a "yes", P, may be written

$$\begin{aligned} P &= 1 - (1-\gamma)\prod_i \exp[-|g(t_i)|^\beta] \\ &= 1 - (1-\gamma)\exp\left[-\sum_i |g(t_i)|^\beta\right]. \end{aligned} \tag{3}$$

It is sometimes convenient to consider the instants to be of infinitely small duration, in which case we may pass to the integral

$$P = 1 - (1-\gamma)\exp\left[-\int_{-\infty}^{\infty} |g(t)|^\beta \, dt\right]. \tag{4}$$

The first obstacle has been cleared. The second problem, of deriving the filter response $g(t)$ from the input waveform $f(t)$, amounts to the problem of determining the transfer function of the filter. For the eye, this problem has as yet no general solution. The function, particularly its phase component, cannot be measured without elaborate assumptions. When measurements are obtained, they are found to vary with the spatial waveform and adapting luminance. However, a solution that is sufficient for the purposes of these experiments may be obtained by assuming that the transfer function has *constant gain and linear phase* over the region of frequency occupied by the input waveform $f(t)$. Where this is so, the filter output will simply be the input delayed in time by some interval τ and scaled in amplitude by some constant s. That is:

$$g(t) = sf(t + \tau). \tag{5}$$

This assumption can approximate the truth only if the signal $f(t)$ is designed in such a way as to occupy a very narrow frequency region. An unfortunate characteristic of such a signal is that it will occupy an extended region of time, and thus not be amenable to presentation in discrete trials. We may compromise with a signal which is relatively localized in both time and frequency:

$$f(t) = a\exp(-t^2/\sigma^2)\sin(2\pi f_t t). \tag{6}$$

This *frequency burst* is the product of a Gaussian of amplitude a and time constant σ and a sinusoid of carrier frequency f_t. Several examples are shown in Fig. 3. For each, the frequency spectrum is approximately Gaussian, and declines to $1/e$ (about 37%) of its maximum in $1/(\pi\sigma)$ Hz.

The behavior of the model may be compared to data in two quite different ways. First, by suitable combination of equations 4, 5 and 6 we may easily calculate the probability of a "yes" response to any

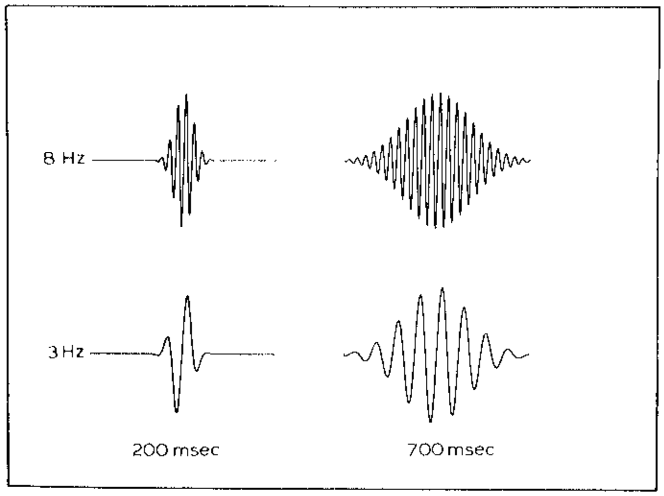

Fig. 3. Examples of the temporal waveforms used in the experiments. Each frequency burst is the product of a sinusoid whose frequency is noted at the left, and a Gaussian, whose time constant is indicated below.

given signal amplitude, a. This probability is a function of the parameters s, γ and β, whose values may be manipulated in order to maximize the correspondence of predicted and obtained psychometric functions.

The second comparison touches more directly upon temporal summation. Examination of equation 4 reveals that, for all stimuli at some fixed probability of detection,

$$c = \int_{-\infty}^{\infty} |g(t)|^{\beta} \, \mathrm{d}t \qquad (7)$$

where c is a constant. In words, equation 7 asserts that for all stimuli at threshold, the integral of the absolute value of the response, raised to the power β, will be a constant.

A remarkable and fortunate feature of this expression is that it includes as special cases three of the deterministic models described above. The *peak-or-trough detector* is quantitatively equivalent to the probability summation model when there is no noise. This latter circumstance may be represented by a psychometric function that is infinitely steep, that is, by setting β equal to ∞ (in practice a β of 40 is sufficient). The *power integrator* is described by equation 7 with a β of 2, and with finite limits to the integral. In Rashbass' version, the limits are plus and minus 100 msec; in Koenderink and van Doorn's they are plus and minus 250 msec.

Since the models which equation 7 may represent are varieties of temporal summation, it is natural to ask how they behave as a signal is lengthened in time. The models are restricted to signals described by equation 6, so the duration of the signal must be varied by manipulating σ, the time constant of the Gaussian envelope. Furthermore, σ cannot be made too small, or equation 5 will again be violated. Subject to these constraints, we may calculate the variation in sensitivity that each model predicts for variations in σ. Examples of such calculations are shown in Fig. 4. In (a) the integration has been performed without limit, in (b) the limits are plus and minus 100 msec.

For finite integration, sensitivity increases appreciably only for values of σ ranging up to the duration of the integration interval. For integration without limits, sensitivity continues to increase without asymptote as σ is made larger. Furthermore, this in-

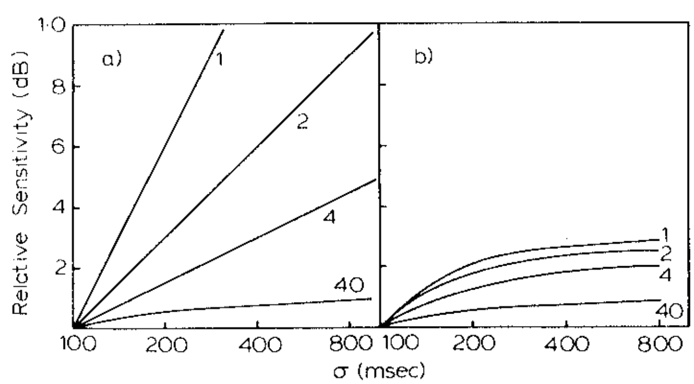

Fig. 4. Sensitivity to a frequency burst as a function of the time constant σ. Calculations are based on Eqn 7 without integration limits (a), or with limits of plus and minus 100 msec (b). The parameter is the exponent β.

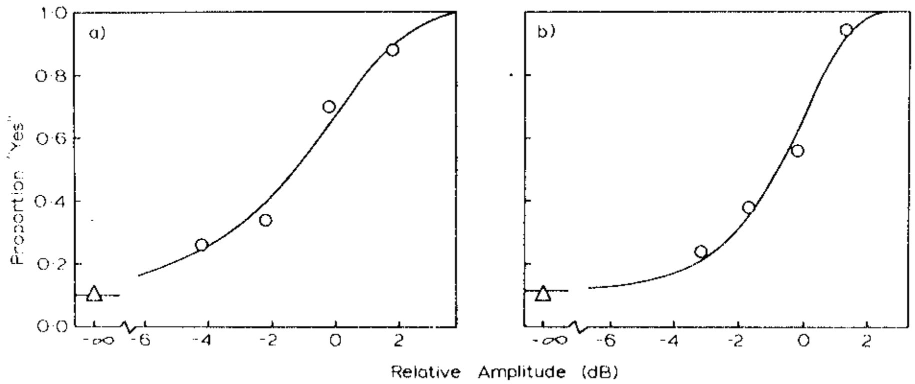

Fig. 5. Selected results of experiment 1. The ordinate expresses proportion of "yes" responses, the abscissa, amplitude of the frequency burst relative to its threshold. The triangles show the false alarm rate. The continuous curves are the best fitting versions of equation 9. Panel (a): observed AM; panel (b): observer RP.

crease is very nearly linear when expressed on logarithmic axes, the slope being equal to $1/\beta$ (the curves depart from linearity only when $\sigma f_t/\beta$ is small). Notice that the technique of limiting the frequency extent of the signal effectively bypasses the integrating action of the linear stage of the model, so that what temporal summation does occur is due to probability summation. Thus in the case of $\beta = 40$, where probability summation does not occur, very little temporal summation takes place.

These results encourage us to perform the following sequence of experiments. First, thresholds may be measured for frequency bursts varying in σ. Examination of these results will indicate whether any of the curves of Fig. 4 are an adequate description of the data. If the relation between the logarithms of sensitivity and σ appears linear, then the most appropriate value of β may be estimated. The estimate of β obtained in this way may then be compared to that yielded by fitting the psychometric function expressed by equation 4 to the data.

METHODS

Two experiments were performed. In both, the spatial pattern was a 4 c/deg sinusoidal grating, generated by a PDP 11/10 computer on the face of a Tektronix 604 CRT with a frame rate of 200 Hz by z-axis modulation of a high frequency raster. In each frame, a pre-calculated list of 12 bit numbers representing the spatial waveform was read from the computer to a digital-analog-converter (DAC) at a rate of about 6 μsec/number. The unattenuated contrast of the pattern in each frame was determined by a number from a second pre-calculated list of 10 bit numbers, representing the temporal waveform as described by equation 6, sampled at intervals of 5 msec, read to the digital input of a second, multiplying DAC. The analog input to the multiplying DAC was provided by the output of the first DAC. The output of the multiplying DAC was applied to the z-axis of the CRT, after passing through a computer controlled attenuator.

The screen was viewed binocularly with natural pupils and a chin rest from a distance of 228 cm. It subtended 2.5° horizontally and 1.9° vertically, and was surrounded by an 8° diameter circular surface of about the same color and luminance (approximately 15 cd/m^2). Two observers were used. Both were emetropic and both were naive as to the purposes of the experiment.

In experiment 1, psychometric functions were obtained for temporal waveforms in which σ was constant at 400 msec while the temporal frequency ranged from 1 to 20 Hz. Each waveform lay within an interval of 2 sec, centered on the midpoint, and was accompanied by a tone of equal duration. Within a session, a minimum of 50 presentations were made at each of four amplitudes of two temporal frequencies. (Stimuli containing both temporal frequencies, not relevant to the present experiment were also present in each session.) An equal number of blank trials were also included.

In experiment 2, a staircase procedure was used to determine thresholds concurrently for temporal waveforms whose value of σ ranged from 100 to 700 msec in steps of 100 msec. The temporal frequency was fixed within a session at either 3 or 8 Hz. Each waveform occupied an interval of 3 sec. The staircase proceeds for each individual stimulus by reducing its amplitude by one logarithimic "step" following each "yes" response, and increasing it by one step following each "no". A step size of 3 dB is used until the first instance in which the response changes from "yes" to "no", or "no" to "yes" (a reversal).[2] Subsequently, a step size of 1 dB is used. Upon completion of at least 30 reversals, threshold for each stimulus is computed by taking the average, in dB, of the midpoints between all amplitude pairs resulting in reversals, excepting the first reversal. Two sorts of catch trial, a blank and a stimulus 10 dB above the current estimate of threshold, each occur with a probability of 0.05. An incorrect response to either is signaled by a sequence of feedback tones. Responses to catch trials have no effect upon the progress of the staircase and do not enter into the calculation of threshold.

RESULTS

Experiment 1

Two data sets selected from the results of Experiment 1 are shown in Fig. 5. The abscissa in each

[2] By convention, decibels of contrast are given by the formula $dB(x) = 20 \log_{10}(x)$.

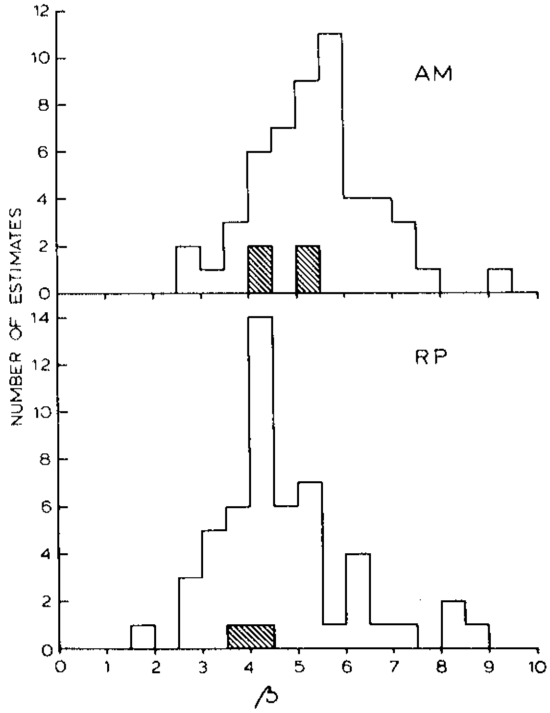

Fig. 6. Distribution of estimates of the parameter β. Results for the two observers are shown separately. See text for details.

panel indicates amplitude of the temporal waveform in dB, relative to an amplitude at which approximately 63% are reported to be seen. The triangular symbols show false alarm rates.

We wish to compare these data to the behavior of the model, as described by equation 4. To express the sensitivity parameter of the model in familiar units, we substitute

$$\alpha = s^{-1}\left[\int_{-t}^{t} |\exp(-t^2/\sigma^2)\sin(2\pi f_t t)|^{\beta}\, dt\right] \quad (8)$$

to arrive at

$$P = 1 - (1 - \gamma)\exp\left[-(a/\alpha)^{\beta}\right]. \quad (9)$$

The parameter α, the "threshold amplitude" is now the amplitude at which the probability of a response from the threshold device is 0.63. The continuous curves in Fig. 5 are described by equation 9 and are fitted to the data by a procedure described in the appendix. This procedure yields maximum likelihood estimates of the parameters α, β and γ, as well as a χ^2 statistic (d.f. = 2) describing the quality of fit. The average χ^2 for AM was 1.57, for RP, 1.82. The data sets of Fig. 5 were selected in order to show the quality of fit indicated by these averages. In (a) the statistic is 1.59, in (b) it is 1.82. In each figure the fit approximates the average of all fits for that observer. In general, the fits are quite good. Only two out of 104 are rejected at the 0.05 level.

The distribution of estimates of the parameter β which are obtained from the fitting procedure is shown in Fig. 6. Each unshaded bar indicates the number of data sets yielding an estimate of β between the specified limits. The majority of both observers' estimates lie between 4 and 6. The distribution for AM appears shifted to somewhat higher values than that for RP. The shaded bars in Fig. 6 will be described later.

In Fig. 7 sensitivity (the (inverse of the estimate of the threshold amplitude α) is plotted as a function of temporal frequency for both observers. For both, sensitivity is approximately uniform until it begins to decline in the neighborhood of 8 Hz. The various arrows inset into Fig. 7 will be discussed below.

Experiment 2

The results of Experiment 2 are plotted in Fig. 8. Each panel contains the amplitude thresholds collected during a single session. Observer and temporal frequency are indicated.

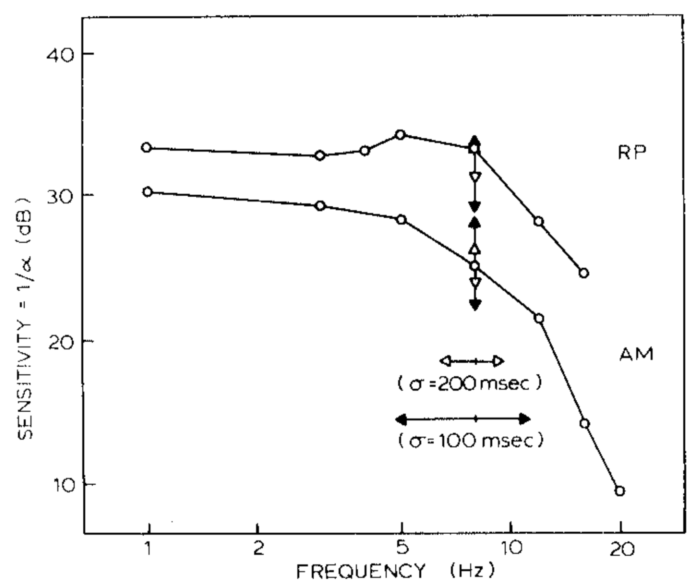

Fig. 7. Sensitivity to frequency bursts as a function of temporal frequency for the two observers. The horizontal arrows show the approximate spectrum width of a burst of the specified time constant, centered about 8 Hz. The vertical arrows show the variation in sensitivity that might be expected over these frequency ranges.

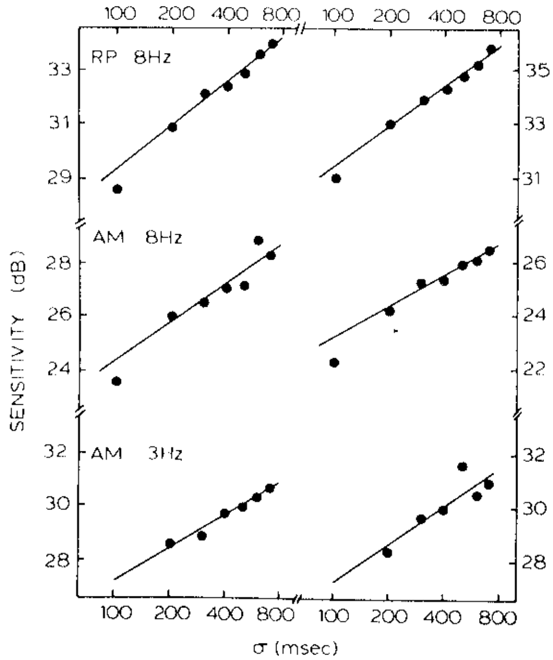

Fig. 8. Sensitivity to frequency bursts as a function of the time constant σ. The straight lines are linear regressions on all points for $\sigma > 100$ msec. Observer and temporal frequencies are: first row, RP, 8 Hz; second row, AM, 8 Hz; third row, AM, 3 Hz.

It is apparent that sensitivity continues to improve up to the largest value of σ used. This result is inconsistent with a *power integrator* having an integration interval of 200 msec (Rashbass, 1970). Furthermore, with the exception of the points at $\sigma = 100$ msec for a frequency of 8 Hz, the thresholds appear to lie upon a straight line as prescribed by equation 7 without integration limits.

Equation 7 may only be properly applied where the transfer function has constant gain over the spectrum of the signal. It is plausible to suppose that the thresholds for stimuli in which $\sigma = 100$ msec depart from the model predictions because their spectra extend too far in a region in which the transfer function gain is not in fact constant. The data of Fig. 7 demonstrate that sensitivity does indeed vary in the neighborhood of 8 Hz. The horizontal arrows show the approximate extent over which the spectra for $\sigma = 100$ msec (filled arrows) and 200 msec (open arrows) fall to 37% of their maxima. The vertical arrows show the variation in sensitivity which may be expected over this frequency range. For the longer signal, the variation is about 3 dB, for the briefer, about 6 dB.

To the extent that the above hypothesis is correct we are justified in fitting a straight line to the points for $\sigma > 100$ msec. The inverse of the slope of this line, which by the arguments above may be taken as an estimate of β, can be simply obtained by linear regression in the log–log space. These estimates are indicated by the shaded bars in Fig. 6. Clearly, the estimates obtained from the two experiments agree very well.

DISCUSSION

To summarize the argument: a model has been constructed to describe the effects of probability summation over time. The essence of the model is contained in two equations: equation 4, which describes the psychometric function, and equation 7, which describes the influence of duration upon sensitivity. Equation 7 also describes as special cases three deterministic models: the *peak-or-trough detector* and the *power integrator* with long or short integration interval. In the probability summation model the parameter β in equation 7 must have the same value as the β in equation 4. In the *peak-or-trough detector*, the value of β in equation 7 must be, in theory, infinite, though in practice a value greater than 40 is sufficient. The *power integrator* requires that the value of β in equation 7 be 2, and that limits to the integral be introduced.

The data collected demonstrate that the value of β in equation 7 does correspond to that in equation 4, that it does not equal 2 and is less than 40, and that equation 7 best describes the results when the integration interval is at least 700 msec in duration. In short, the data reported here are consistent with a model incorporating probability summation over time, and are inconsistent with several models which do not incorporate this process.

The weight of this conclusion is borne by the predicted and obtained correspondence between the two estimates of β. Each estimate, however, is subject to error. If the noise is correlated from instant to instant, then less advantage will be gained by extending the signal in time. The β estimated by experiment 2 would then exceed the true value. If the noise consists, in part, of variations so slow as to occur between individual trials or groups of trials, as suggested by Hallet (1969), then the estimate of β that is derived from experiment 1 will be less than that obtained from experiment 2. It is interesting that both of these departures from the simple predictions of the model suggest that the β of experiment 2 should exceed that of experiment 1, whereas the data, if anything, tend to indicate the opposite. It is possible that the methods of data collection and analysis used in experiment 1 produce an upwardly biased estimate of β (as is suggested by simulations of a related psychophysical procedure), but an understanding of this and other detailed predictions of the model may require a more sophisticated treatment of the time properties of the noise.

It must also be recognized that the correspondence between the two estimates of β may be fortuitous. A model in which the filter is followed by a power law transformation and integration over an extended period would provide a comparable result (Rashbass, 1976).

Previous data which reveal a continuing increase in sensitivity as a signal is prolonged beyond the presumed critical duration have been collected by Nachmias (1967), Tolhurst (1975), and Roufs and van Stuyvenberg (1977). Despite differences in spatial and temporal waveforms, in each case the results display a rate of increase of sensitivity comparable to that obtained here. In the two more recent works it is asserted that the result is due to probabilistic effects, and Tolhurst supports the assertion with a qualitative argument.

In past treatments of probability summation over time the calculation has been performed over a finite collection of discrete events (Roufs, 1974). One virtue of the present approach is that it allows an approximate description of probability summation over a continuous waveform. This makes it possible to apply the model to the visibility of more complex temporal stimuli; for example, waveforms which are not adequately described in terms of the number of their peaks. One interesting stimulus of this sort to which the model has been applied is that containing several quite different frequencies. The results of this investigation will be reported in a future publication.

Acknowledgements—I thank Jacob Nachmias for the very substantial contributions that he has made to this research. I also thank Norma Graham, John Robson, Dave Heeley and Denis Pelli for critical and productive readings of an earlier draft. This research was supported by NSF Grant BMS75-07658 to Jacob Nachmias, and formed a part of a doctoral dissertation submitted to the University of Pennsylvania.

REFERENCES

Bloch A. M. (1885) Experience sur la vision. *C. r. Séanc. Soc. Biol.* **37**, 493–495.

Chandler J. P. (1965) STEPIT. *Quantum Chemistry Program Exchange*, Department of Chemistry, Indiana University, Bloomington, Indiana, U.S.A.

De Lange H. (1952) Experiments on flicker and some calculations on an electrical analogue of the foveal systems. *Physica* **18**, 935–950.

Graham C. H. and Margaria R. (1935) Area and intensity time relation in the peripheral retina. *Am. J. Physiol.* **113**, 299–305.

Graham N. (1977) Visual detection of aperiodic stimuli by probability summation among narrow band channels. *Vision Res.* **17**, 637–652.

Green D. M. and Luce R. D. (1975) Parallel psychometric functions from a set of independent detectors. *Psychol. Rev.* **82**, 483–486.

Hallet P. E. (1969) The variations in visual threshold measurement. *J. Physiol., Lond.* **202**, 403–419.

Hoel P. G., Port S. C. and Stone C. J. (1971) *Introduction to Statistical Theory*. Houghton Mifflin, Boston.

Kelly D. H. (1961) Visual responses to time-dependent stimuli. II. Single-channel model of the photopic visual system. *J. opt. Soc. Am.* **51**, 747–754.

Koenderink J. J. and van Doorn A. J. (1978) Detectability of power fluctuations of temporal visual noise. *Vision Res.* **18**, 191–195.

Nachmias J. (1967) Effect of exposure duration on visual contrast sensitivity with square-wave gratings. *J. opt. Soc. Am.* **57**, 421–427.

Quick R. F. (1974) A vector magnitude model of contrast detection. *Kybernetik* **16**, 65–67.

Rashbass C. (1970) The visibility of transient changes of luminance. *J. Physiol., Lond.* **210**, 165–186.

Rashbass C. (1976) Unification of two contrasting models of the visual increment threshold. *Vision Res.* **16**, 1281–1283.

Roufs J. A. J. (1972) Dynamic properties of vision—II. Theoretical relationships between flicker and flash thresholds. *Vision Res.* **12**, 279–292.

Roufs J. A. J. (1974) Dynamic properties of vision—VI. Stochastic threshold fluctuations and their effect on flash to flicker sensitivity ratio. *Vision Res.* **14**, 871–888.

Roufs J. A. J. and van Stuyvenberg J. A. P. (1977) Gain curve of the eye to subliminal sinusoidal modulation of light. *IPO Annual Progress Report 1976* **11**, 56–63.

Sperling G. and Sondhi M. M. (1968) Model for visual luminance discrimination and flicker detection. *J. opt. Soc. Am.* **58**, 1133–1145.

Stotter A. T. (1975) *The detection of complex auditory signals*. Ph.D. Thesis, University of Cambridge.

Tolhurst D. J. (1975) Reaction times in the detection of gratings by human observers: a probabilistic mechanism. *Vision Res.* **15**, 1143–1149.

Watson A. B. and Nachmias J. (1977) Patterns of temporal interaction in the detection of gratings. *Vision Res.* **17**, 893–902.

Weibull W. (1951) A statistical distribution function of wide applicability. *J. appl. Mech.* **18**, 292–297.

APPENDIX

This appendix describes the procedure that has been used to fit frequency-of-seeing data to a psychometric function of the type proposed by Quick (1974), using a distribution function described by Weibull (1951). The procedure has been implemented in a computer program named QUICK.

The data are in the form of number triples (a_i, n_i, x_i), $i = 1, 2, \ldots, I$, where a_i is the amplitude of the ith stimulus, n_i the number of presentations, and x_i the proportion of yes responses. We suppose that the number of yes responses, $n_i x_i$, is sampled from a binomial distribution of parameters n_i and P_i, that is,

$$f(n_i x_i) = \binom{n_i}{n_i x_i} (P_i)^{n_i x_i} (1 - P_i)^{n_i(1-x_i)}, \quad n_i x_i = 1, \ldots, n_i$$
$$= \phi \tag{A1}$$

elsewhere.

We wish to maximize the likelihood function for the I data samples,

$$L = \prod_i f(n_i x_i \mid P_i). \tag{A2}$$

It is equivalent to maximize the logarithm of this function,

$$\log L = \sum_i \log \binom{n_i}{n_i x_i} + n_i x_i \log(P_i) + n_i(1 - x_i) \log(1 - P_i)]. \tag{A3}$$

Collecting separately those terms which depend upon the parameters P_i and those which do not, we write

$$\log L = \sum_i \left[\log \binom{n_i}{n_i x_i} \right] + M \tag{A4}$$

where

$$M = \sum_i n_i [x_i \log(P_i) + (1 - x_i) \log(1 - P_i)]. \tag{A5}$$

It is now sufficient to maximize M as a function of the I parameters P_i.

The maximization is performed under two hypotheses. In the first, H_0, we assume nothing about the P_i. Each is free to assume its most likely value relative to the x_i. Since the maximum likelihood estimate of the parameter P_i of the density described by equation A1 is simply x_i, the maximum of M under H_0 is given by

$$\hat{M}_0 = \sum_i n_i [x_i \log(x_i) + (1 - x_i) \log(1 - x_i)]. \tag{A6}$$

(In the evaluation of equation A6 it is occasionally necessary to make use of the fact that $\lim_{x \to 0} [x \log(x)] = 0$.)

In the second hypothesis, H_1, we assume that the P_i are given by

$$P_i = 1 - (1 - \gamma) \exp[-(a_i/\alpha)^\beta]. \tag{A7}$$

Note that under this constraint the I parameters P_i are reduced to the three parameters α, β and γ. An expression for M_1 may now be obtained by substituting equation A7 into equation A5. The maximum of M_1 is then found by means of STEPIT (Chandler, 1965), a general purpose minimization routine. The values of α, β and γ that provide the maximum value of M are then the maximum likelihood estimates of these parameters.

The degree to which equation A7 adequately represents the data is assessed by comparison of the maxima of M under the two hypotheses. Specifically, the statistic

$$-2\log[\hat{L}_1/\hat{L}_0] = 2[\hat{M}_0 - \hat{M}_1]$$

may be shown to be asymptotically χ^2 with degrees of freedom equal to the difference in the number of parameters in the two hypotheses, that is, $I - 3$ (Hoel, Port and Stone, 1971).