Biol. Cybern. 59, 137–147 (1988)

Biological Cybernetics

Visual Contrast Detection by a Single Channel Versus Probability Summation Among Channels

U. Mortensen

Westfälische Wilhelms-Universität Münster, FB 8, Institut III, Fliednerstrasse 21, D-4400 Münster, Federal Republic of Germany

Abstract. It is now generally accepted that the human visual system consists of subsystems ("channels") that may be activated in parallel. According to some models of detection, detection is by probability summation among channels, while in other models it is assumed that detection is by a single channel that may even be tuned specifically to the stimulus pattern (detection by a matched filter). So far, arguments in particular for the hypothesis of probability summation are based on plausibility considerations and on demonstrations that the data from certain detection experiments are *compatible* with this hypothesis. In this paper it is shown that *linear* contrast interrelationship functions together with a property of a large class of distribution functions (strict log-concavity or log-convexity on the relevant set of contrasts/intensities) uniquely point to detection by a single channel. In particular, models of detection by probability summation based on Quick's Model are incompatible with linear contrast interrelationship functions. Sufficient (and observable) conditions for the strict log-concavity/log-convexity of distribution functions are presented.

1 Introduction

While it appears to be a well established fact that the visual system is a network of subsystems [1] the role these subsystems play in the detection process is not generally agreed upon. Some authors, e.g. Shapley and Tolhurst (1973), Kulikowski and King-Smith (1973), Hauske et al. (1976) proposed models of the detection process, and provided data supporting them, according to which a stimulus pattern is detected by a single, linear subsystem that may even be specifically tuned to the stimulus pattern. Other authors, e.g. Sachs et al. (1971), Mostafavi and Sakrison (1976), Graham (1977), Wilson and Bergen (1979) proposed models of detection by probability summation among subsystems. Here the subsystems are conceived as being activated stochastically independent of each other and consequently as signalling the presence of a stimulus pattern independent of each other. The stimulus is detected if it is signalled by at least one subsystem. Graham (1977) in particular claimed that the data of Shapley et al. and of Kulikowski et al. can equivalently be explained in terms of her model of detection by probability summation among spatial frequency channels.

The interpretation of detection data in terms of models of probability summation on the one hand and by single channel models on the other leads clearly to different evaluations of system properties. Now an assumption made by all authors mentioned above is that the detection process can be formulated without reference to its temporal aspects. The models may be called *static*; they may be viewed as special cases of the general class of dynamic models. The advantage of restricting one's attention to static models is the direct interpretability of threshold data with respect to the spatial transfer characteristics of the detecting (sub-)systems. Take for instance the *c*ontrast-*i*nterrelationship-*f*unctions (CIFs) reported by Kulikowski and King-Smith (1973): a pattern $m_1 s_1$, e.g. a thin, vertical line (s_1) with contrast m_1 is superimposed on a second pattern $m_2 s_2$, e.g. a sinusoidal grating with spatial frequency w_0 (s_2) with contrast m_2. For various values of m_2 the corresponding value of m_1 is determined such that the probability of

[1] Subsystems may be a "detectors" in the sense of Shapley and Tolhurst (1973) or Kulikowski and King-Smith (1973), which are meant to be single receptive fields, or a "channels" in the sense of Graham (1977), which are assumed to be made up of many receptive fields without probability summation among the receptive fields forming a channel, or "mechanisms" as discussed by Wilson and Bergen (1979). These are classes of receptive fields, and there is probability summation among the members of such a class (i.e. "spatial probability summation")

detection remains constant. According to the static model of a linear line detector the response of the detector is given by $m_1h_1+m_2h_2$ with h_1 the "unit"-response of the detector to the line and h_2 the unit response of the detector to the grating and the unit responses are constants for all values of m_1 and m_2. The CIF is given by $m_1=c_0/h_1-m_2(h_2/h_1)$, c_0 some constant, and the slope of the CIF may be taken to represent the relative sensitivity of the line detector to the spatial frequency w_0. In a general dynamic model the invariance of the unit responses need not hold and there is no simple and direct interpretation of CIFs. The advantage of static models carries over to the case where detection by probability summation among subsystems is considered.

It will be shown in this paper that with respect to the class of static models under very general and at least partially testable conditions CIFs of the sort determined by Abadi and Kulikowski (1973), Kulikowski and King-Smith (1973), Shapley and Tolhurst (1973) cannot be interpreted in terms of models of detection by probability summation among independent subsystems. This result appears to be of relevance in so far as the implication of the data of e.g. Kulikowski and King-Smith (1973) and Hauske et al. (1976) may have to be re-considered: while models of detection by probability summation like Graham's (1977) and Wilson and Bergen's (1979) are considered to be parsimonious with respect to the assumed types of subsystems, models like that of Hauske et al. (1976) are not seen as parsimonius since the existence of pattern specific detectors is postulated, and it is not clear why the visual system should contain special detectors for a variety of different and very specific patterns. However, if data do indeed force us to accept single-channel models we may have to think about the genesis of such pattern-specific detectors in the visual system; self-organisation processes induced by the experimental situation may for example be responsible for Hauske et al.'s data.

The proof of the unique meaning of linear CIFs presented in this paper does not carry over to general dynamic models that do not allow for the invariance of the unit responses mentioned above. So if one considers static models as being inadequate one gains the possibility of interpreting linear CIFs in terms of models of detection by probability summation, but looses the possibility of interpreting the data with respect to invariant unit responses; spatial and temporal transfer characteristics have to be evaluated simultaneously. It is not difficult to see that against the background of the hypothesis of detection by probability summation among subsystems the simultaneous evaluation of spatial and temporal transfer characteristics represents a very intricate problem. If, on the other hand, static models are considered to be adequate we will have to deal with the questions arising from Hauske et al.'s (1976) work.

The general assumptions from which our results are derived are given in Sect. 1.1; although we will not consider the general case of detection by temporal probability summation (TPS), the results partially carry over to this case as well. Detection by TPS is a lengthy topic by itself and will be discussed in a separate paper.

According to the result presented in 1.1 the interpretation of threshold curves is tied to a property of the distribution functions of the random variables representing the effect of "noise" in the subsystems. They have to be strictly log-concave (log-convex) on the set of stimulus parameters considered in the experiment. In Sect. 1.2 examples of distribution functions exhibiting this property are given; it turns out that practically all distribution functions employed in psychophysics share this property. Still, the random variables cannot be directly observed and so we have to consider properties of threshold curves and psychometric functions that allow to *infer* whether the corresponding distribution functions actually are strictly log-concave or log-convex. Therefore in Sect. 2 a test is suggested that allows us to decide whether strict log-concavity/convexity holds provided detection by probability summation among independent subsystems is assumed. Section 3 contains a summarising discussion of the results.

1 Linear Contrast-Interrelation Functions and Detection by PSS

1.1 Definitions, Assumptions and a Basic Implication

The stimulus pattern is defined by a luminance distribution, and this distribution may again be specified in terms of a function characterising the luminance value at each retinal coordinate and each point of time within the (open) time interval $J=(0, T)$ representing a trial. Further stimulus patterns may be generated from a given one by varying the value of some attribute of the stimulus, e.g. spatial or temporal frequency, duration, size etc. The set Σ^* of stimulus patterns thus generated may be seen as a class of patterns, e.g. the class of all sinusoidal gratings presented with identical time course but having different spatial frequencies. A pattern from such a class is then characterised by a real number $r \in Q=(q_1, q_2)$, $q_1<q_2$ some reals, representing the value of the attribute with regard to which the patterns of the class differ, $-r$ may thus represent the value of a spatial frequency. Let m denote intensity or contrast of a pattern, $m \in M$ an interval from $\mathbb{R}$, and let $\theta=\mathbb{R}X\mathbb{R}$ represent retinal coordinates. The stimulus pattern s may then be represented by $s=m\sigma$, $m \in M$,

where σ is a mapping

$\sigma: \theta \times J \times Q \to (-1, 1)$.

σ is the normalised spatio-temporal luminance distribution defining the pattern. Then

$\Sigma^* = \{m\sigma(\theta, t, r); m \in M, t \in J, r \in Q\}$.

Let σ_1 and σ_2 be two normalised luminance distributions that do not necessarily belong to the same class, i.e. σ_1 may represent a normalised grating and c_2 a normalised disc. Let $m_1 \in M$, $m_2 \in M$; the pattern $m_1\sigma_1 + m_2\sigma_2$ will be called a *superposition* in the following. Here we may choose for instance m_2 to be the parameter of interest, i.e. $r = m_2$. In this case Σ^* consists of all superpositions $s = m_1\sigma_1 + r\sigma_2$, $r \in Q$.

The probability of detection for a member of a class Σ^* depends then upon *(i)* the value of m (contrast or intensity), and *(ii)* the value r of the attribute with respect to which the members of the class differ. The following definition refers to such a class of stimulus patterns:

Definition 1. Let Σ^* be a class of stimulus patterns differing with respect to a certain attribute A. Let M and Q be open intervals of the real line representing the possible values of intensity/contrast and of A, respectively. For each $r \in Q$, the mapping $P_r: M \to (0, 1)$ will be called a *psychometric function*, and the set $H^*(\Sigma^*) = \{P_r; r \in Q\}$ will be called a *psychometric family*.

Comment. $P_r(m) \in (0, 1)$ is the probability of detection for a stimulus pattern from $H^*(\Sigma^*)$ if the pattern has intensity/contrast $m \in M$ and attribute value $r \in Q$. The notation $P_r(m)$ stresses the point that the psychometric function is a function of m for given value of r; r is a parameter of the psychometric function. However, for the following it is convenient to regard the probability of detection as a function of *m and r* and we will then write $P(m, r)$ instead of $P_r(m)$. We will write simply H^* instead of $H^*(\Sigma^*)$ if there is no possibility of confusion.

The notion of a psychometric family was introduced by Falmagne (1982); this concept turns out to be useful when properties of certain sets of psychometric functions are investigated with respect to the properties of detection mechanisms.

Definition 2. Let H^* be a psychometric family. Let $f: Q \to M$ be a mapping such that for given $r \in Q$ the value of $m = f(r)$ is such that for all $P_r(m) \in H^*$, $P_r(m) = P(m, r) = P(f(r), r) = p_0$ a constant (e.g. $p_0 = 1/2$). Then f will be called the *threshold curve*. If the stimulus pattern is defined by a superposition with $r = m_2$ and $f(r) = m_1$, $f(r)$ will be called the *contrast-interrelationship function* (CIF).

Definition 3. Suppose that in a given experiment the subsystems $C_1, \ldots, C_n$, $n \geqq 1$ may be activated by the stimulus pattern. Suppose further that the stimulus is detected if at least one of the C_j, $1 \leqq j \leqq n$, signals the stimulus pattern; then detection is said to be by *probability summation among subsystems* (PSS).

Assumption 1. Suppose that for stimulus pattern $s \in \Sigma^*$ subsystems from the set $C(\Sigma^*) = \{C_1, \ldots, C_n; n \geqq 1\}$ may be activated. The following conditions are supposed to hold:

(i) The activities of the C_j, $1 \leqq j \leqq n$ are representable by random variables X_j, having distribution functions $F_{X_j}(x; m, r)$, since $s \in \Sigma^*$ depends upon m and r.

(ii) The $F_{X_j}(x; m, r)$ are differentiable with respect to m and r at least twice.

(iii) The C_j signal the stimulus if $X_j > S_j$, S_j some constant representing an internal threshold.

(iv) The X_j are stochastically independent.

Assumption 2. Let $s = m\sigma \in \Sigma^*$ and let $J = (0, T)$ again represent a trial of duration T. Let the temporal course of the deterministic part of C_j's response to $s \in \Sigma^*$ be represented by the function $g_j: J \to \mathbb{R}, j = 1, \ldots, n$. Then:

(i) For all $m \in M$, the C_j are representable as linear systems; thus there exist functions $h_j: J \to \mathbb{R}$ such that $g_j(t; m, r) = mh_j(t; r)$. (The h_j will be called *unit responses* since they represent the response to the normalised pattern σ.)

(ii) The X_j are defined by

$$X_j = \max_{t \leqq T} g_j(t) + U_j, \tag{1}$$

with U_j being a random variable having constant value within $(0, T)$.

(iii) For $j = 1, \ldots, n$, U_j is independent of $g_j(t)$ for all $t \in (0, T)$.

Assumption 3. Suppose the stimulus pattern is a superposition; for the deterministic part of C_j's response we have from Assumption 2 and Definition 2

$g_j(t) = f(r)h_{j1}(t) + rh_{j2}(t)$.

It will be assumed that at least one of the following conditions holds:

(i) $h_{j2}(t) = a_j h_{j1}(t)$, $a_j =$ constant for all $t \in (0, T)$

(ii) For some subinterval $I_j \subset (0, T)$, $h_{j2}(t) =$ constant and $dh_{j1}(t)/dt = 0$ only if $t \in I_j$.

Comments. Assumption 1 implies that if detection is by PSS, then the probability of detection is given by $P_r(m) = P(m, r) = 1 - P\left(\bigcap_{j=1}^{n} \{X_j \leqq S_j\}\right)$, i.e. by

$$P(m, r) = 1 - \prod_{j=1}^{n} F_j(S_j; m, r), \quad m \in M, r \in Q. \tag{2}$$

In the following the $F_{X_j}(S_j; m, r)$ will be considered as functions of m and r. To this end let $g_{0j}(m, r) := \max g_j(t; m, r)$ and let F_{U_j} be the distribution

function of the random variable U_j introduced in (1); it is

$$F_{X_j}(S_j; m, r) = P(g_{0j}(m, r) + U_j \leqq S_j)$$
$$= P(U_j \leqq S_j - g_{0j}(m, r)),$$
$$\text{i.e. } F_{X_j}(S_j; m, r) = F_{U_j}(S_j - g_{0j}(m, r)).$$

The parameters m and r enter the distribution of X_j only via g_{0j} so that we may define

$$G_j(g_{0j}) := F_{U_j}(S_j - g_{0j}). \tag{3}$$

Numerically the $G_j(g_{0j})$ are identical to the $F_{X_j}(S_j; m, r)$ so that we may write

$$P(m, r) = 1 - \prod_{j=1}^{n} G_j(g_{0j}),$$
$$g_{0j} = g_{0j}(m, r),\ m \in M,\ r \in Q. \tag{4}$$

The psychometric function for detection by a single subsystem is given by (2) – or (4) – as in the special case $n = 1$.

Assumption 2 relates the transfer characteristics of the detecting subsystems to the distribution functions of the X_j. (1) in particular defines what will be called (temporal) *peak-detection* in the following; peak-detection may be taken as a particular approximation to the more general case where $g_j(t)$ is added to the trajectories $U_j(t)$ of some noise process and the stimulus is being detected if $X_j = \max[g_j(t) + U_j(t)] > S_j$; ("max" means, as before, the maximum on $(0, T)$)[2]. (1) may be shown to be valid if either the $U_j(t)$ vary only a little as compared to the variation of $g_j(t)$ within $J = (0, T)$, so that $U_j(t) \approx U_j$ a constant within J, or if the S_j are large compared to the $g_{0j} = \max g_j(t)$. Assumption 2 is often made when threshold curves are to be interpreted. Let, for instance, the threshold curves be plots of the temporal frequency w of a sinusoidally flickering stimulus versus the corresponding sensitivity to this frequency, and let the output of the single detecting system be given by $g(t) = m|H(w)|\sin(wt + \phi(w))$, where $H(w)$ is the Fourier transform of the impulse response on the detecting system. Then $\max g(t) = m|H(w)|$ so that $1/m = c_0|H(w)|$, c_0 some constant. An analogous interpretation holds if s is a sinusoidal grating. For another application of Assumption 2 see Roufs and Blommaert (1981).

Assumption 3 may appear to be somewhat artificial at first glance. However, assumptions of this sort are tacitly made in any static model of detection referring explicitly only to spatial characteristics of the detecting C_j, e.g. Abadi and Kulikowski (1973), Kulikowski and King-Smith (1973), Marr (1982), Shapley and Tolhurst (1973), Hauske et al. (1976), Graham (1977); the spatial pattern of activity generated by the stimulus is assumed not to depend upon time. This independence may be assumed for the case that the response has settled at the stationary state. Assumption 3 covers this situation as a special case. According to Assumption 3, *(i)* one has $g_j(t) = (f(r) + ra)h_{j1}(t)$ and $\mathrm{d}g_j(t)/\mathrm{d}t = (f(r) + ra_j)\mathrm{d}h_{j1}(t)/\mathrm{d}t = 0$ at $t = t_{0j}$, say. Obviously $\mathrm{d}g_j(t)/\mathrm{d}t|_{t=t_{0j}} = 0$ iff $\mathrm{d}h_{j1}(t)/\mathrm{d}t|_{t=t_{0j}} = 0$ for all $r \in Q$ so that t_{0j} and consequently $h_{j1}(t_{0j})$ does not vary with the value of r. For Assumption 3, *(ii)* the argument is analogous. If detection is by a single subsystem the conditions are thus sufficient (not necessary) to allow us to *predict* from $g_{0j} = f(r)h_{j1}(t_{0j}) + rh_{j2}(t_{0j}) = \text{constant}$ that $f(r)$ is a linear function of r. As will be seen below conditions of the sort postulated in Assumption 3 are also required if one wants to *infer* from a linear CIF that detection is by a single subsystem. Of course, neither the condition *(i)* nor *(ii)* in Assumption 3 are necessarily true. In general, t_{0j} may depend upon r; this implies that even if we know that detection is by a single, linear subsystem we cannot yet predict a linear CIF if not additional assumptions about the unit responses h_{ji}, $i = 1, 2$ are made.

To prepare the formulation of the main proposition let us recall the definition of log-concave (log-convex) functions. Let $K_j = g_{0j}(Q)$, i.e. let K_j be the range of the g_{0j}. Recall further that a function $G(x)$ is called strictly log-concave (log-convex) on some set Q if $\log G(x)$ is strictly concave (convex) on Q; then $\mathrm{d}^2 \log G(x)/\mathrm{d}x^2 < 0$ (> 0) on Q (Roberts and Varberg 1973). The following statement may now be proven:

(S1) Suppose

(a) Assumptions 1, 2, and 3 are true; if in particular *(i)* of Assumption 3 holds, the constants a_j are not all identical

(b) the $G_j(g_{0j})$ as defined in (3) are strictly log-concave on K_j (alternativly, the G_j are strictly log-convex on K_j).

(c) $f(r) = br + c$, $b \neq 0$.

Then detection is not by probability summation among independent subsystems.

---

[2] Approximations for the distribution of X_j may be derived within the framework of extreme value theory (Leadbetter et al. 1983). A different approach which appears to be better suited for our purposes was suggested by Ditlivsen (1971): if for instance the noise is wide-sense stationary Gaussian with autocorrelation function $R(\tau)$, then $P(X_j \leqq S_j) \approx \Phi(S_j)\exp(J)$, where

$$J = \int_0^T [\Phi(g'(t)/a)g'(t) - a\phi(g'(t)/a)]\phi(S - g(t))/\Phi(S - g(t))\mathrm{d}t,$$

where Φ is the distribution function of a standardised Gauss-variable and ϕ is the corresponding density function, $g'(t) = \mathrm{d}g(t)/\mathrm{d}t$ and $a = \sqrt{R''(0)/2\Pi}$; the index j has been dropped for simplicity.

Equation (1) follows immediately if one puts $U_j(t) = U_j$ a constant during a trial

Proof. Suppose $P(f(r),r)$ is given by (4). Then $G := \prod_{j=1}^{n} G_j = q_0 = 1 - p_0$ a constant for all $r \in Q$; hence $dG/dr = 0$ on Q. Now $dG/dr = G \sum_{j=1}^{n} (dG_j/dr)/G_j$, and since $G = q_0$ a constant the sum has to equal 0 on Q. Its derivative with respect to r is then also identically 0 on Q, i.e.

$$\sum_{j=1}^{n} [(d^2G_j/dr^2)/G_j - ((dG_j/dr)/G_j)^2] = 0, \quad r \in Q. \quad (5)$$

From (a) we have $dG_j/dr = (dG_j/dg_{0j})dg_{0j}/dr$ and

$$d^2G_j/dr^2 = (d^2G_j/dg_{0j}^2)(dg_{0j}/dr)^2 + (dG_j/dg_{0j})d^2g_{0j}/dr^2.$$

According to Assumption 3, *(i)* the unit response $h_{j1}(t)$ and $h_{j2}(t)$ are independent of r. This, together with (c), yields $dg_{0j}/dr = (b + a_j)h_{j1}(t_{0j})$ and $d^2g_{0j}/dr^2 = 0$. If Assumption 3, *(ii)* holds we have $g_{0j} = f(r)h_{j1}(t_{0j}) + rh_{j2}$ with $h_{j2} = A_j$ a constant, so that $dg_{0j}/dr = bh_{j1}(t_{0j})$ and $d^2g_{0j}/dr^2 = 0$ on Q. Substituting the expressions for dG_j/dg_{0j}, d^2G_j/dg_{0j}^2 and dg_{0j}/dr into (5) we have

$$\sum_{j=1}^{n} D_j h_{j1}^2 (a_j + b)^2 = 0, \quad r \in Q \quad (6)$$

with $D_j := (d^2G_j/dg_{0j}^2)/G_j - ((dG_j/dg_{0j})/G_j)^2$; obviously, $D_j = d^2(\log G_j)/dr^2$, and according to (b) either $D_j < 0$ for all j or $D_j > 0$ for all j. Since the a_j are postulated to differ from each other it follows that $[h_{j1}(a_j + b)]^2 > 0$ for all j (excluding the trivial case $h_{j1} = 0$ for all j), and hence (6) and consequently (4), i.e. (2), cannot be true. □

Comments. According to *S1* the empirical finding of a linear CIF is incompatible with the hypothesis of detection by PSS, provided the assumptions (a) and (b) hold. Assumptions that are equivalent to Assumption 3 are, as already mentioned, commonly made in models concerned with the spatial properties of detecting systems; they will be further discussed in the Discussion section. In the following section it will be shown that (b) is indeed more the rule than an exception.

1.2 Distribution Functions Excluding Detection by PSS for Linear CIFs

Suppose that a linear CIF was observed and a decision has to be made as to whether detection was by a single subsystem or by PSS. Since the hypothesis of detection by PSS may be discarded if the G_j are all strictly log-concave (log-convex) for the values of the g_{0j} generated by the stimuli we have to discuss conditions under which strict log-concavity or log-convexity holds. It will first be shown that in particular log-concavity is a property of practically all $G_j(g_{0j}) = F_j(S_j - g_{0j})$ employed so far in psychophysics if only $q_0 = 1 - p_0 \geqq 0.5$. It will then be shown that the condition on q_0 may be dropped if n can assumed to be large.

To keep the notation simple the index j will be dropped in the following proposition:

(S2) Let F_U be the distribution function of the random variable U and let u^* be the corresponding mode of U; let further $g^* = S - u^*$. Suppose F_U is differentiable at least twice and the density function $f(u) = dF(u)/du$ is unimodal and strictly decreasing for $u > u^*$. Then $G(g_0) = F(S - g_0)$ is strictly log-concave for $g_0 < g^*$.

Proof. $G(g_0)$ is strictly log-concave on $I = [0, g^*)$ iff

$$(d^2G/dg_0^2)/G < [(dG/dg_0)/G]^2 \quad (*)$$

on I. Now $dG/dg_0 = -f(S - g_0)$ and with $u = S - g_0$, $d^2G/dg_0^2 = -(df/du)(du/dg) = df/du$, and since $df/du < 0$ for strictly decreasing f, (*) is true. □

Comments. (S2) implies that given a linear CIF we can reject the hypothesis of detection by PSS if *(i)* the densities f_j are strictly decreasing for $u_j > u_j^*$ and *(ii)* for all $j = 1, \ldots, n$, $G_j(g_{0j}) > q_0$ for $g_{0j} < g_j^*$. Condition *(i)* holds for all distributions commonly employed in psychophysics; the uniform distribution is a counter-example for *(i)*. However, this distribution is rarely used in the present context. The first condition of *(ii)*, $G_j(g_{0j}) > q_0$ for all j, is a trivial consequence of (4). The second condition, $g_j < g_j^*$ is seen to be necessary as follows. Note that if $g_{0j} \in I$ then $r \in Q$. Let in particular u_j^0 be the median of u_j, and let $g_{0j}^0 = S_j - u_j^0$, and let $q_0 = p_0 = 1/2$. Then (*) holds for all unimodal distributions with $u_j^0 \geqq u_j^*$, i.e. for all distributions that are either symmetrical or have a mode smaller than the median, provided dF/du is strictly decreasing for $u > u^*$. One sees that *(S2)* applies to "classical" distributions like the normal, the logistic, and the log-normal distribution, but also to other distributions less popular in psychophysics like the Cauchy-distribution, (representing, for instance, the distribution of the quotient of two independently and normally distributed random variables).

(S2) defines sufficient conditions for the log-concavity of the G_j that may imply certain restrictions with respect to the minimal value of q_0 that may be chosen. However, these restrictions do not have to hold necessarily as the following example shows:

Quick's Model. Suppose that for all j, $u_j = -y_j$ and the $y_j > 0$ are Weibull-distributed; then

$$\begin{aligned} G_j(g_{0j}) &= P(u_j \leqq S_j - g_{0j}) = P(-y_j \leqq S_j - g_{0j}) \\ &= P(g_{0j} - S_j \leqq y_j) = 1 - P(y_j < g_{0j} - S_j) \\ &= \exp[-(g_{0j} - S_j)^\beta], \beta > 0, g_{0j} \geqq S_j. \end{aligned}$$

If we want $g_{0j}=0$ to represent the state of no activation by a stimulus we have to put $S_j=0$. From (4) the probability of detection is then given

$$P(m,r)=1-\exp\left[-\sum_{j=1}^{n}(g_{0j}^{\beta})\right]. \tag{7}$$

This is the model of detection by PSS suggested by Quick (1974). Now $\log(G_j)=-(g_{0j}-S_j)^{\beta}$, $S_j \geqq 0$, is known to be strictly concave for all $g_{0j}>S_j$, if $\beta>1$, and to be strictly convex if $\beta<1$. Thus the G_j are all strictly log-concave for $\beta>1$ or strictly log-convex for $\beta<1$, so that if a linear CIF is observed the hypothesis of detection by PSS has to be rejected for all values of $q_0=1-p_0$ chosen by the experimenter; Quick's model is not compatible with linear CIFs, except for the case $\beta=1$; however, this case seems to have never been observed.

It will now be shown that under rather general conditions the assumption of a sufficiently large n [as in Wilson and Bergen's (1979) model] leads to further types of distribution functions that imply a rejection of the hypothesis of detection by PSS if a linear CIF is observed.

Extreme Value Distributions. For distributions like the normal or the log-normal log-concavity of the $G_j(g_{0j})$ can only be postulated if the g_{0j} do not exceed a certain level; this, however, may be guaranteed if only $q_0 \geqq 1/2$. If n is sufficiently large, this condition may be dropped; let again $\prod_{j=1}^{n} G_j=q_0$ and let M^* denote the geometric mean of the G_j. Obviously, $M^*=q_0^{1/n}$ and $M^*\to 1$ for $n\to\infty$ for all $q_0\in(0,1)$; the values of the G_j vary about a mean value M^* that is the closer to 1 the larger the value of n.

Let us begin with the special case that the u_j are all identically distributed with distribution function F_U, and $S_j-g_{0j}=S-g_0=u$ for all j. For greater simplicity we drop the subscript U in the following. Suppose for a moment that we know the value of n, and $F^n(u)=q_0$. For given probability of detection $p_0=1-q_0$ we know that $M^*=F(u)=q_0^{1/n}$ is the closer to 1 the greater the value of n. This may be expressed by saying that F assigns an appropriate mean and a suitable variance to u, or equivalently, defines implicitly an appropriate scale of the random variable u.

Let us now recall that two distributions F_1 and F_2 are said to be of the same type if they are related by $F_1(u)=F_2(au+b)$, with $a>0$ a scale parameter and b a location parameter (Feller 1966, p. 44). This definition enables us to express the distribution of u – called F_1 for the moment – in terms of an arbitrary distribution F_2 of the same type as F_1. Let, for instance, the distribution F_1 of the u_j be Gaussian; then we may write $F_1(u)=F_2(au+b)$, where we may take F to be the distribution of $N(0,1)$-variables. a and b are chosen such as to adjust the mean and the variance of the standardised variables to the proper values. Certainly, the parameters a and b depend upon the value of n, so that we may write a_n for a and b_n for b. The $a_n>0$, b_n are known as "norming constants"; they are characteristic for a given distribution and are unique up to a linear transformation (Leadbetter et al. 1983).

The chosen value of q_0 determines the value of u, and vice versa. So we may conceive of q_0 as being a certain value of a function $H_n(u):=F^n(a_nu+b_n)$. If the sequences $\{a_n\}$, $a_n>0$, $\{b_n\}$ are the sequences of norming constants of F the function $H_n(u)$ changes only little with increasing n if only n has already become sufficiently large; then $H_n(u)$ converges (weakly) towards a distribution function $H(u)$, called the limiting distribution. If F^n converges towards H, F is said to be in the *domain of attraction* of H. Because of the slow changes of $H_n(u)$ with n for sufficiently large n we have the approximation

$$F^n(u)\approx H((u-b_n)/a_n). \tag{8}$$

A detailed evaluation of conditions of weak convergence may be found in Leadbetter et al. (1983). Here it is sufficient to state that most distributions employed in visual psychophysics converge weakly towards a limiting distribution As Gnedenko (1943) has shown the Poisson distribution is an example of a distribution not belonging to any domain of attraction. Fisher and Tippett (1928) and Gnedenko (1943) have shown that for independent and identically distributed random variables there exist only three limiting distributions. They are of the form $H_i(u)=\exp(\Gamma_i(u))$ $i=1,2,3$, with

$$\begin{aligned}
&\Gamma_1(u)=-u^{-\beta}, \beta>0, u>0, \Gamma_1(u)=0, u\leqq 0\\
&\Gamma_2(u)=-(-u)^{\beta}, \beta>0, u\leqq 0, \Gamma_2(u)=1, u>0\\
&\Gamma_3(u)=-\exp(u), \qquad -\infty<u<\infty.
\end{aligned} \tag{9}$$

The following result for nonidentically distributed random variables is now easily derived:

(S3). Let $G_j(g_{0j})=F_j(S_j-g_{0j})$ and suppose that *(i)* for each j, F_j belongs to a domain of attraction, and *(ii)* $F_j(a_n(S_j-g_{0j})+b_n)\to 1$ as $n\to\infty$. Then

$$P_r(m)=P(m,r)\approx 1-\exp\left[-c_0\sum_{j=1}^{n}\Gamma^{(j)}(A_{jn}(S_j-g_{0j})+B_{jn})\right],\ c_0=1/n \tag{10}$$

with $\Gamma^{(j)}$ defining the domain of attraction to which F_j belongs; $\Gamma^{(j)}$ is one of the three functions Γ_i given in (9), and $A_{jn}=1/a_{nj}$, $B_{jn}=-b_{nj}/a_{nj}$.

Remark. In general one will assume that the F_j all belong to the same domain of attraction (provided they

belong to a domain of attraction at all), simply because it is plausible that the noise characteristics of different subsystems are similar. However, (10) allows for the more general interpretation that the F_j may also belong to different domains of attraction; if this should be the case we will have to assume that for each of the occurring domains of attractions there is more than 1 subsystem having a distribution function belonging to this domain.

Proof. Let $v_{nj}=a_{nj}(S_j-g_{0j})+b_{nj}=a_{nj}v_j+b_{nj}$, i.e. $v_j:=S_j-g_{0j}$. Then $[F_j(v_{nj})]^n=\exp[n\log F_j(v_{nj})]$. Because of $F_j(v_{nj})\to 1$ with increasing n one has $\log F_j\approx -(1-F_j)$ and

$$\lim_{n\to\infty} n\log F_j(v_{nj})=-\lim_{n\to\infty} n(1-F_j(v_{nj})$$
$$=\log H^{(j)}(v_j)=\Gamma^{(j)}(v_j), \tag{11}$$

with $H^{(j)}$ the limiting distribution of F_j. Further,

$$\lim_{n\to\infty}\prod_{j=1}^{n} F_j(v_j)=\lim_{n\to\infty}\exp\left[\sum_{j=1}^{n}\log F_j(v_j)\right]$$
$$=\lim_{n\to\infty}\exp\left[-c_0\sum_{j=1}^{n} n(1-F_j)\right]$$
$$=\exp\left[-c_0\sum_{j=1}^{n}\lim_{n\to\infty} n(1-F_j(v_j)\right].$$

This yields, together with (8) and (11), expression (10) for $P(m,r)$. □

Comments. Assumption *(i)* implies a restriction of modelling detection processes to distributions being members of a domain of attraction; fortunately, this class encompasses most distributions employed in visual psychophysics except for the Poisson distribution which, however, is not of interest in the present discussion. The meaning of assumption *(ii)* may be seen as follows: suppose that for some F_k *(ii)* does not hold. Then the remaining F_j have to be even closer to 1, implying that the stimulus is detected predominantly by the subsystem C_k, i.e. we have detection by a single subsystem contrary to our assumption. Thus *(ii)* is not a condition restricting the generality of the result (10).

If the distribution F is already of the type of a limiting distribution the "$\approx$"-sign in (10) may be replaced by an "$=$"-sign.

While (10) is certainly far too general for practical computations, further restrictions on (10) are not necessary if we want to decide whether the hypothesis of detection by PSS may be kept provided a linear CIF was observed. If we have reason to assume that n is sufficiently large to justify (8) and if *(i)* and *(ii)* of *(S1)* can be adopted, then $G_j(g_{0j})=\exp(\Gamma^{(j)}(A_j(S_j-g_{0j})+B_j))$ and from (9) one sees that the G_j are either all strictly log-concave or log-convex: Γ_1 is strictly concave, Γ_2 is strictly convex for $\beta<1$ and strictly concave for $\beta>1$ and Γ_3 is strictly concave. The G_j will thus be all log-concave if the F_j are in the domain of attraction of $H(u)=\exp(\Gamma_3(u))$, and they will all be either log-concave or log-convex if only the exponents β are the same for all F_j, or at least satisfy the same inequality, provided the limiting distribution is defined in terms of Γ_1 or Γ_2. A linear CIF does not allow then for detection by PSS and supports the hypothesis of detection by PSS and supports the hypothesis of detection by a single subsystem instead.

Suppose that the U_j are independent and identically distributed so that $A_j=A$ and $B_j=B$, with A and B constants for all j, and the F_j belong to the domain of attraction of Γ_2. Suppose further that $S_j=S$ for all j, and $AS=-B$. Then (10) reduces to Quick's (1974) model (7).

So far we have shown that either commonly employed distribution functions or simply the assumption of a sufficiently large n together with certain mild restrictions on the F_j imply that linear CIFs exclude the possibility of detection by PSS. However, for a given experiment we usually do not know anything about the distribution functions; the U_j and therefore the G_j are not directly observable. In the following section a test concerning a type of the G_j that is of interest in the present context is suggested.

2 Parallelism of Psychometric Function on log*m*-Scales

In this section we consider sets Σ^* of stimulus patterns s that are not superpositions but single patterns differing with respect to a certain property like spatial or temporal frequency, duration, size etc. It will be shown that if Assumptions 1 and 2 hold and detection is by PSS, and if further the psychometric functions from the family $H^*(\Sigma^*)$ are parallel on the log$m$-scale, $m\in M$, then the $P_r(m)\in H^*$ are given by Quick's model. The converse also holds, i.e. if the $P_r(m)$ are given by Quick's model, then they are parallel on the log$m$-scale. In combination with linear CIFs the assumption of detection by PSS leads then to a contradiction: the $G(g_{0j})=1-P_r(m)$ are then log-concave (or log-convex) and we have to conclude that detection is by a single subsystem.

We have to begin with a formal definition of parallelism on log$m$-scales.

Definition 4. Let $H^*(\Sigma^*)$ be a given psychometric family. Let $x:=\log m$, and for each $P_r(m)\in H^*(\Sigma^*)$, let $A(x,r):=P_r(e^x)=P(e^x,r)$. If for any $r\in Q$, $r'\in Q$, $r\neq r'$ there exists a constant $\alpha(r,r')\in\mathbb{R}$, which is independent of $m\in M$, such that $A(x,r)=A(x+\alpha(r,r'),r')$, then the elements of $H^*(\Sigma^*)$ will be called log-*parallel*, and $H^*(\Sigma^*)$ will be called a log-*parallel psychometric family*.

Let $H^*(\Sigma^*)$ be a log-parallel family. The elements of this family may then be written in the form $P_r(m) = P(ma(r))$, with $a: Q \to \mathbb{R}$ some function of $r \in Q$, not of $m \in M$, since if $P(m(r), r) = P(m(r'), r') = p_0 \in (0, 1)$ some constant, then log-parallelity implies that there exists a constant $\alpha(r, r')$ which is independent of the value of p_0 such that $\log m(r') = \log m(r) + \alpha(r, r')$ for all $m(r)$ and corresponding $m(r')$ from M. For each pair (r, r') there exists then a real number $a(r, r') > 0$ such that $\alpha(r, r') = \log a(r, r')$; therefore we have $m(r)a(r, r') = m(r')$. Consequently one finds for fixed $m(r')$, that for all $r \in Q$, $P_r(m(r)) = \text{const}$ if and only if $g = m(r)a(r, r') = \text{const}$. So $P_r(m)$ depends only upon the value of g, and suppressing the fixed r' we arrive at the above representation. Conversely, the representation $P_r(m) = P(ma(r))$ also implies log-parallelism.

The following proposition may now be proved:

(S4). Suppose detection is by PSS and Assumptions 1 and 2 hold; let the number n of subsystems involved in the detection task be the same for all $r \in Q$. Let the h_{0j} denote again the maxima of the unit responses of the C_j; we assume that the set

$$\{x + \log h_{01}, \ldots, x + \log h_{0n};\ x \in R,\ r \in Q\} \subset R^n$$

contains inner points. Then the $P_r(m) \in H^*(\Sigma^*)$ are given by Quick's model (7) if and only if $H^*(\Sigma^*)$ is a log-parallel psychometric family.

Proof. To show that log-parallelism is implied by Quick's model is the trivial part of the proof: let $h_{0j} = \max h_j(t)$ so that $g_{0j} = m h_{0j}$. Then $P_r(m) = 1 - \exp\left(-\sum_{j=1}^{n} g_{0j}^{\beta}\right)$ and $\sum_{j=1}^{n} g_{0j}^{\beta} = m^{\beta} \sum_{j=1}^{n} h_{0j}^{\beta}$ so that $P_r(m) = P(ma(r))$ with $a(r) = \left(\sum_{j=1}^{n} h_{0j}^{\beta}\right)^{1/\beta}$ for all $m \in M$, $r \in Q$; consequently $H^*(\Sigma^*)$ is a log-parallel psychometric family.

To show that log-parallelism implies – under the conditions of *(S4)* – Quick's model is not trivial. This part of the proof is due to Bolthausen (1987).

Let again $F_j := P(X_j \leqq S_j)$, and let us define $H_j = H(m b_{0j}(r)) = 1 - F_j$. From (4), $P(m) = 1 - \prod_{j=1}^{n} (1 - H_j)$, and $P(m) \to 0$ for $m \to 0$ implies then $H_j(m h_{0j}) \to 0$ for all j, and $P(m) \to 1$ for $m \to \infty$ implies $H_j(m h_{0j}) \to 1$ for all j. Let $x = \log m$, $t_j = \log h_{0j}$, and let $\phi(x) := \log(1 - H_j)$. From Definition 4 one sees that log-parallelism implies that there exist functions Φ and Γ such that

$$\sum_j \phi(x + t_j) = \Phi[x + \Gamma(t_1, \ldots, t_n)]. \tag{12}$$

It will be assumed that ϕ, Φ and Γ are differentiable. If such a relation holds for a set of points $\{t_1, \ldots, t_n\}$ it will also hold for sets $\{x + t_1, \ldots, x + t_n\}$ so that (12) will hold for an open set $\{t_1, \ldots, t_n\}$ of points. Differentiation of (12) with respect to t_1 yields

$$\phi'(x + t_i) = \Phi'(x + \Gamma)\Gamma_i, \qquad \Gamma_i = \partial\Gamma/\partial t_i. \tag{13}$$

Differentiation of (13) with respect to t_j leads to

$$0 = \Phi''(x + \Gamma)\Gamma_i\Gamma_j + \Phi'(x + \Gamma)\Gamma_{ij}, \qquad \Gamma_{ij} = \partial^2\Gamma/\partial t_i \partial t_j.$$

For fixed $t_1, \ldots, t_n$ this is a linear, homogeneous differential equation with constant coefficients for $f(x) = \Phi(x + \Gamma)$; the corresponding characteristic polynomial has solutions $\mu_1 = -\Gamma_{ij}/\Gamma_i\Gamma_j$, $\mu_2 = 0$, so that

$$f(x) = c\,e^{\mu x} + b, \tag{14}$$

where c, b, and μ may depend upon t_i and t_j. If we now integrate (12) with respect to x we have

$$\phi(x + t_i) = \Phi(x + \Gamma(t_1, \ldots, t_n))\Gamma_i, \qquad i = 1, \ldots, n.$$

Substituting (14) for Φ shows that Φ is not specific for t_i and t_j, implying in particular that μ in (14) is not specific for t_i and t_j, i.e. $\mu_{ij} = \mu$. One has

$$\phi(x + t_j) = c_1\,e^{\mu x} + b_i,\ c_i = c\Gamma_i,\ b_i = b\Gamma_i.$$

Then $1 - H_j = \exp(c\Gamma_j m^{\mu} + b_j)$, and since $H_j(m h_{0j}) \to 0$ for $m \to 0$ and $H_j(m h_{0j}) \to 1$ for $m \to \infty$ it follows that $b_j = 0$, $c_j < 0$ and $\mu > 0$, and from the definition of the t_j we may conclude that $c\Gamma_j = h_{0j}$. □

Comment. A related result was presented by Green and Luce (1975) who made the additional assumption that the number n of subsystems involved depends upon the value of the parameter r representing the attribute with regard to which the elements of Σ^* differ. Certainly, n does not *necessarily* depend upon r. It does not appear to be possible to generalise the proof presented by Green and Luce to cover the case considered here, namely $n = \text{const}$.

(S4) implies that if we assume detection by PSS and postulate that the probability of detection is given in terms of Quick's model, we should observe log-parallel psychometric functions. If log-parallel psychometric functions are not observed, Quick's model is inadequate. Simultaneously, peak-detection by a single linear subsystem also implies log-parallel psychometric functions, because then $P_r(m) = P(m h_{0j}(r) + U_j > S_j) = P(m h_{0j}(r))$; this prediction of log-parallelism holds independent of the distribution defining $P_r(m)$. So a lack of log-parallelism indicates also that the peak-detection assumption is not justified.

3 Discussion

It was demonstrated that it is possible, given the validity of Assumption 1 to 3, to deduce from the form of threshold curves and general properties of psycho-

metric functions that detection is by a single subsystem. In particular, under the conditions of *(S1)*, linear CIFs and Quick's model for detection by PSS are incompatible with each other, implying that it is not possible to reinterprete for instance the data of Kulikowski and King-Smith (1973) in terms of the model of Graham (1977), unless one is prepared to accept the "steepness parameter" β in Quick's model to be equal to 1. Given again that the conditions of *(S1)* may be assumed, linear CIFs are indicative of detection by a single subsystem unless one is prepared to assume distribution functions with (a) multimodal densities, or (b) densities not strictly decreasing for $u > u^*$, u^* the mode of u, or (c) densities with mode greater than the median. Note that *S2* only states some sufficient conditions. In any case, the combination of the finding of a linear CIF with the assumption of detection by PSS requires additional assumptions about the stochastic properties of the "noise" variables that deviate from those usually made. However, if detection by PSS is postulated and if the number n of subsystems involved in the detection task is assumed to be large as in the Wilson and Bergen (1979) model, and possibly in Graham's (1977) model, one is led to distribution functions that always exclude the possibility of detection by PSS, regardless of the value of q_0. It should be noted that values of $n=20$ or $n=30$ may already be "large"; however, if the u_j are Gaussian n has to be considerably larger for (13) to hold (cf. Galambos 1978).

Let us now consider the conditions (a)–(c) of *(S1)*. According to (a) the discussion of linear CIFs with respect to detection by PSS was restricted to stochastically independent subsystems and temporal peak-detection. Proposition *S1* can be generalised to the case of dependent subsystems at least for particular forms of dependencies. It is not clear yet which form of dependency among the subsystems is the most adequate to assume. However, the following hypothesis seems to be plausible: neighbouring subsystems may be stochastically dependent, and the dependency turns into independence with increasing distance (local and/or functional) among subsystems. If further n is "large" it is known that under rather general side conditions the same limiting distributions as for independent u_j hold, i.e. $P(m,r)$ is again given by (13), and we are again led to expressions for $P(m,r)$ of the form (13). For a detailed discussion of this property of extreme value statistics cf. Leadbetter et al. (1983) and Galambos (1978).

The restriction to temporal peak-detection is not critical either; in many situations the assumption of peak-detection appears to be justified, e.g. Roufs and Blommaert (1981). *(S1)* may be generalised to the case of detection by TPS in individual subsystems; approximations for $P(m,r)$ in case of detection by TPS will be discussed elsewhere.

A necessary condition for *(i)* of Assumption 3 to hold is that σ_1 and σ_2 have identical time courses. Then a sufficient condition for *(i)*, i.e. $h_{j2}(t)=a_j h_{j1}(t)$ for all $t \in (0,T)$ and a_j a constant is the spatio-temporal separability of the impulse responses of the C_j. However, this need not be true. Alternatively, we may assume that there exists a subinterval I_j of $(0,T)$ such that for all $t \in I_j$, $h_{j1}(t)=\text{const}$ and $h_{j2}(t)=\text{const}$ and $g_j(t) < g_j(t')$ for $t \notin I_j$, $t' \in I_j$. Then we have again $h_{j2}(t) = a_j h_{j1}(t)$ for $t \in I_j$, $a_j = \text{const}$ on I_j. Here not only the time courses of σ_1 and σ_2 have to be identical but they are chosen that the deterministic responses of the C_j are stationary within some subinterval of $(0,T)$.

If *(ii)* of Assumption 3 can be adopted then $g_j(t) = f(r)h_{j1}(t) + rh_{j2}$ and $h_{j2} = K_j$ a constant within $(0,T)$, $dg_j(t)/dt = f(r)dh_{j1}(t)/dt$ and $dg_j(t)/dt = 0$ for $t = t_{0j}$ and t_{0j} does not depend upon r. Here the time courses of σ_1 and σ_2 need not be identical. For instance, the pattern $r\sigma_2$ may already be presented to the subject when the trial begins and remains stationary during $(0,T)$; we have then to assume that the response of the subsystems belonging to $C(\Sigma^*)$ to $r\sigma_2$ is stationary throughout $(0,T)$. However, this procedure may lead to adaptational processes not allowing for *(iii)* to hold. Another procedure would be to present the pattern $r\sigma_2$ such that transients are avoided and the response to this pattern is stationary within some subinterval of $(0,T)$; $f(r)\sigma_1$ is presented such that the response to this pattern is nonzero only if the whole response to $r\sigma_2$ is stationary.

On the other hand, if stimulus patterns are to be presented as in Wilson and Bergen (1979) and Bergen et al.'s (1979) experiments – the temporal course of the stimuli is defined by a Gaussian function – then the $g_j(t)$ cannot be assumed to be stationary. If, under these conditions, one does not want the temporal variables to appear explicitly in the modelling of the detection process one is automatically led to Assumption 3, *(i)*; clearly the value of t_{0j} is independent of the value of r in this case. But then the model is no longer compatible with the finding of a linear CIF, in particular not if $P(m,r)$ is defined in terms of Quick's function. This casts some doubts on interpretations of data within the framework of Graham's (1977) or Wilson and Bergen (1979) model. For a critique of static models of spatial stimulus processing see Fleet et al. (1985).

Suppose now that Assumption 3, *(i)* is not justified so that a_j is not a constant, but varies in some way during $(0,T)$. Then (5) together with $dg_j(t)/dt = 0$ at $t = t_{0j}$ implies $f(r)/r = -h'_{j2}(t_{0j})/h'_{j1}(t_{0j})$ and t_{0j} will no longer be independent of r if for instance $f(r) = br + c$ with $b \neq 0$ and $c \neq 0$. As a consequence, we can no longer postulate that $d^2 g_{0j}/dr^2 = 0$ everywhere in $(0,T)$

and we do not arrive at (9), i.e. a linear CIF is no longer counterindicative of detection by PSS. Unfortunately, we have now lost the possibility of a direct interpretation of detection data with respect to the spatial properties of the C_j.

Let us assume again that Assumption 3 holds. If, however, $a_j = \hat{a}$ a constant for all j it is easy to show that then $\hat{a} = -b$, b the slope of the CIF. Then we would have $dg_{0j}/dr = 0$ everywhere in $(0, T)$ and the log-concavity/log-convexity of the G_j would no longer be counterindicative of detection by PSS. However, if the the patterns s_1 and s_2 are spatially different the postulate that the a_j differ is natural: let, for instance, s_1 be a thin line such that its spatial amplitude spectrum is flat over the range of frequencies to which the human visual system is sensitive, and let s_2 be a sinusoidal grating with spatial frequency w_0. Let the C_j be spatial frequency channels. Suppose C_j and C_k are channels with center frequencies (i.e. frequencies to which they are most sensitive) w_j and w_k differing by at least the larger of the bandwidths of C_j and C_k. Let $w_0 \leqq w_j < w_k$. Then there will be be virtually no activation of C_k by w_0 and the relation $h_{j2} = -bh_{j1}$ *and* $h_{k2} = -bh_{k1}$ cannot hold simultaneously. For subsystems of the sort considered by Wilson and Bergen (1979) analogous arguments hold.

Let us consider the condition (b) of *(S1)*. So far the interpretation of linear CIFs as supporting the hypothesis of detection by a single, linear subsystem were based on more or less plausible *assumptions* about the underlying probability distributions. The question whether these assumptions can be tested empirically cannot be answered satisfactorily yet. For instance testing for the log-parallelity of psychometric functions means to test only a *sufficient* condition for log-concavity or log-convexity of the G_j and some future work concerning empirical tests is required here. It should be recalled that log-parallelism is a necessary condition for peak-detection by a single, linear system, regardless of the distribution underlying the psychometric function. So the *lack* of log-parallelism indicates that the hypotheses of *peak-detection* by a single subsystem on the one hand and, on the other hand, of detection by PSS as specified by Quick's model are invalid. Still, log-parallelism seems to have been rarely discussed with respect to visual detection data, with the exception of Nachmias (1967) whose data do support the hypothesis of log-parallelism.

Condition (c) of *(S1)*, namely that the CIF $f(r)$ is linear, is given empirically. The question as to when a straight line is the best curve to be fitted through a set of points need not be discussed here. In any case one could argue that a linear relationship of the sort presented by Kulikowski and King-Smith (1973), p. 1460 is only an approximation to the true nonlinear relationship. Such an argument excludes the hypothesis of detection by a single subsystem for nonempirical reasons. If for instance Assumption 3 *does* hold and detection *is indeed* by a single system, then a linear CIF is *predicted* as the "true" relationship. The situation is different if one does not allow for Assumption 3; since then the $h_{j1}(t_{0j})$ and $h_{j2}(t_{0j})$ depend in general upon r a linear CIF can only be predicted if additional assumptions are made about the spatio-temporal impulse responses of the detecting single system. The consequences of the decision to drop Assumption 3 have, however, been commented upon above.

Our results suggest that it may be too early to assume that detection is necessarily by PSS. The data of Kulikowski et al. and Hauske et al. may well be taken to support the hypothesis of detection by a single subsystem. However, they are not yet conclusive: log-parallelism of psychometric functions would allow a unique interpretation in terms of single subsystem detection, but unfortunately they have not been reported by the authors. To the extent the hypothesis of detection by a single subsystem is plausible for these types of experiment the hypothesis of detection by PSS is plausible for other experimental situations; examples are the experiments by Mostafavi and Sakrison (1976) and Sakitt (1971). Sakitt argued that for her data detection by PSS could be excluded; however, she considered only detection by PS among identically activated subsystems. A re-analysis of her data allowing for not identically activated subsystems shows that detection by PSS is the most plausible hypothesis for her data. Summarising we may say that the experimental situation or task determines which mode of detection is chosen by the subject.

References

Abadi RV, Kulikowski JJ (1973) Linear summation of spatial harmonics in human vision. Vision Res 13:1625–1628

Bergen JR, Wilson HR, Cowan JD (1979) Further evidence for four mechanisms mediating vision at threshold: sensitivities to complex gratings and aperiodic stimuli. J Opt Soc Am 69:1580–1587

Bolthausen E (1987) Personal communication

Ditlivsen O (1971) Extremes and first passage times with applications in civil engineering. Dissertation. Technical University of Denmark, Copenhagen

Falmagne JC (1982) Psychometric functions theory. J Math Psychol 25:1–50

Feller W (1966) An introduction to probability theory and its applications, vol II. Wiley, New York London Sydney

Fisher RA, Tippett LH (1928) Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc Cambridge Phil Soc 24:180–190

Fleet DJ, Hallett PE, Jepson AD (1985) Spatiotemporal inseparability in early visual processing. Biol Cybern 52:153–164

Galambos J (1978) The asymptotic theory of extreme order statistics. Wiley, New York
Gnedenko MV (1943) Sur la distribution limite du terme maximum d'une serie aléatoire. Ann Math 44:423–453
Graham N (1977) Visual detection of aperiodic spatial stimuli by probability summation among narrow band channels. Vision Res 17:637–653
Green DM, Luce RD (1975) Parallel psychometric functions from a set of independent detectors. Psychol Rev 82:483–486
Hauske G, Wolf W, Lupp U (1976) Matched filters in human vision. Biol Cybern 22:181–188
Kulikowski JJ, King-Smith PE (1973) Spatial arrangement of line, edge and grating detectors revealed by sub-threshold summation. Vision Res 13:1455–1478
Leadbetter MR, Lindgren G, Rootzén H (1983) Extremes and related properties of random sequences and processes. Springer, New York Berlin Heidelberg
Marr D (1982) Vision. Freeman, San Francisco
Mostafavi H, Sakrison DJ (1976) Structure and properties of a single channel in the human visual system. Vision Res 16:957–968
Nachmias J (1967) Effect of exposure duration on visual contrast sensitivity with square wave gratings. J Opt Soc Am 57:421–427
Quick Jr, RF (1974) A vector-magnitude model for contrast detection. Kybernetik 16:65–67
Roberts AW, Varberg DE (1973) Convex functions. Academic Press, New York
Roufs JAJ, Blommaert FJJ (1981) Temporal impulse and step responses of the visual system obtained psychophysically by means of a drift-correcting perturbation technique. Vision Res XX:1203–1221
Sachs MB, Nachmias J, Robson JG (1971) Spatial-frequency channels in human vision. J Opt Soc Am 61:1176–1186
Sakitt B (1971) Configuration dependence of scotopic spatial vision. J Physiol 216:513–529
Shapley RM, Tolhurst DJ (1973) Edge detectors in human vision. J Physiol Lond 229:165–183
Wilson HR, Bergen JR (1979) A four mechanism model for threshold spatial vision. Vision Res 19:19–32

Received: January 4, 1988

Professor Dr. U. Mortensen
Westfälische Wilhelms-Universität
FB 8, Institut III
Fliednerstrasse 21
D-4400 Münster
Federal Republic of Germany

Verantwortlich für den Textteil: Prof. Dr. W. Reichardt, Max-Planck-Institut für biologische Kybernetik, Spemannstr. 38, D-7400 Tübingen. Verantwortlich für den Anzeigenteil: E. Lückermann, Springer-Verlag, Heidelberger Platz 3, D-1000 Berlin 33, Fernsprecher: (030)8207-0, Telex: 185411. Druck der Brühlschen Universitätsdruckerei, Gießen. Printed in Germany. – Springer-Verlag GmbH & Co KG, 1000 Berlin 33