Alerting Signals and Detection in a Sensory Network*

WILLIAM J. McGILL[†] AND MALVIN C. TEICH[‡]

Columbia University

Detection theories are based on the proposition that nature designs our sense organs to recover signals from external noise. Not much credence is given to the idea that an essential part of the noise arises internally in conveying information to the brain or that the detected event may be an alert signal rather than a literal copy of the stimulus. Transmission is portrayed here as a birth-and-death process. New events are created and existing events are lost as messages pass from the sense organ to a central decision-making area. Messages are either alert signals or actual "images" of the stimulus formed by branching chains. In both cases, a telltale multiplicative noise appears at the output. We develop: (1) a branching-chain transmission mechanism, (2) its output counting distribution, (3) characteristic properties of branching-chain noise, and (4) examples of the effects of such noise on the discriminability of pure-tone signals. © 1995 Academic Press, Inc.

I. INTRODUCTION

Puzzles abound in modern sensory research.

The classical treatment of absolute visual thresholds devised by Hecht, Shlaer, & Pirenne (1942) continues to enjoy wide acceptance more than 50 years after its introduction. It depicts the visual system as essentially noiseless, arguing that threshold variability can be traced to the intrinsic variability of stimuli. The idea is elegantly simple but at odds with what we know of sensory processing and human judgment.

A huge network exists, more than 2 million neurons, connecting our eyes to the brain (Brown, 1965). If high-level decision centers can access every photon absorbed in each retina, the transmission system must be set up to count with perfect accuracy, permitting no variations of its own to intrude on threshold data. Hard-line advocates of the quantum theory of Hecht *et al.* say this is just what happens. (See Sakitt, 1972.)

Barlow's (1956) update of visual detection introduced an internal noise he labelled "dark light," but the interaction was seen as purely additive. Additive errors do not propagate into larger errors as they pass on up through the system. More recently, Barlow (1977) discussed evidence of central noise in visual detection.

Ideal Detection

Consider the early treatment by Tanner and Swets (1954) of ideal auditory detection. Their unique marriage of signal analysis and decision theory has retained its popularity in the front rank of auditory research for more than 40 years.

We have known from the beginning that ideal detection fails to account for either the slope or the location of auditory psychometric functions. Among other things, it does not explain the poor performance of the ear attempting to detect small intensity differences in pairs of phase-locked pure tones, or bursts of "frozen," i.e., repeatable, noise. Yet these difficulties have not diminished the appeal of ideal detection in the slightest. Here again, analysis is restricted to the stimulus domain. The scheme simply finesses sensory mechanisms.

Mixture distributions of signal and noise are partitioned into decision regions that minimize error, or maximize gain if errors are unequally weighted. The partition is made on a mathematical representation of the input, not on the sensory output. This reflects a popular belief that perceptual sharpening and image enhancement manage to reconstruct stimulus patterns accurately in higher brain centers. Where pure tones are concerned, this faith in the ear's precision is misplaced. Detection data suggest a sizable internal noise varying with signal intensity. Transmission mechanisms are, or should be, parts of the representation.

Classical Sensory Analysis

A bit less radical than by-passing the entire auditory system is the classical position of psychophysics. Carefully chosen stimuli are used to explore receptor mechanisms. Transmission phenomena are excluded. Decision centers in the brain are thought to be capable of looking back through the long neural chain to the periphery, isolating events as they occur in the sense organ. Hence, when we analyze a

critical band, the masking region associated with a pure tone, it is almost a given that we are dealing with an activation pattern in the inner ear. In fact, the evidence shows that critical bands are formed en route up the auditory neural pathway. (See Ehret & Merzenich, 1985.)

Margolis (1987) explains why it is so difficult to give up on these traditional paradigms despite persistent signs of trouble. We want a formula that will translate perceptions directly into receptor events or beyond the sense organ, into stimulus events. Perhaps that is not a good way to proceed.

In Search of Simplicity

Theories seldom fall from favor because they fail to work out satisfactorily. If a detection model offers a simple standard for comparison with data, it will tend to remain popular long after its advocates become convinced there is something wrong. This is an application of Kuhn's (1957) now-famous argument on the history of astronomy. A corollary principle is that most attempts to market complexity end up with no one listening, even when skeptics concede that the arguments are plausible.

Sometimes a theory will adopt ridiculous stances while attempting to cope with contrary evidence. This suggests less ignorance or obstinacy than a determined effort to preserve simplicity. The need to understand exactly what we are saying is so fundamental that arguments continue to be pushed despite clear indications that they are wrong. There is obvious benefit in assuming that brain mechanisms can reconstruct the main features of the external environment with high accuracy. It expresses a naive realism that ordinarily works well in accounting for our daily experience. What is more natural than to try to extend these ideas (wrongly, we think) into areas related to signal detection?

One important consequence is that auditory detection becomes centered on input signal analysis to the exclusion of virtually everything else.

New Paradigms Needed

Most contemporary researchers agree that time has expired on the simplest versions of auditory and visual detection. Sensory systems are not simple. Their work begins in the realm of stimulus energy and ends in decision centers of the brain. Detection problems are not limited to the stimulus domain or to receptor activity. The zones between reception and decision are a mostly uncharted region where messages move in ways that are only partly understood. Some are unaccountably lost. Some are overcome by noise. Noise increases as signals pass through the system. In recent years, with the benefit of much new work on summation, inhibition, and parallel–serial processing, our grasp of sensory mechanisms has improved greatly. This has spawned new versions of detection theory in both vision

and audition. (See Geisler, 1989; Green and Swets, 1988/1966).

The search for new paradigms led eventually to Laming's book, Sensory Analysis (1986). (Also see McGill & Teich, 1989.) Laming jettisoned much of the traditional apparatus of receptor analysis in favor of "black boxes" representing the network between receptor and brain.

The centerpiece is a linkage Laming calls differential coupling. In effect, he argues, a sensory system neutralizes incoming stimulation by dividing it into slightly asynchronous excitatory and inhibitory streams. Constant backgrounds thus become silent. Increments or decrements are detected as departures from this balanced state. The system behaves like an alert-signal generator. Laming (1986) does not offer detail on mechanisms. Instead he gives a battery of arguments drawn from the recent literature, all supporting the idea of differential coupling.

A closely related approach in perception is the concept of "adaptation level." (See Helson, 1947.) In neurophysiology it is spoken of as a "balance" of excitatory and inhibitory effects. We usually prefer these more traditional terms, but language should not divert us from the importance of Laming's idea. He puts a new face on increment detection, masking, and intensity discrimination. He uncovers regularities no one else had noticed. Our starting point then is Laming. We believe we have an elementary way to realize his differential coupler.

Stochastic Networks

Visual and auditory psychophysics deal with stimuli and receptor mechanisms. Visual and auditory detection theories concentrate on stimuli and decision rules. We now add (perhaps "insert" would be more accurate) a third level of analysis dealing with information transmission, as sketched in Fig. 1.

We are aware this brings us close to preaching complexity. How can it be kept simple and manageable? Nearly everyone agrees that transmission occurs prior to detection, but no one other than Laming (1986) or, a bit earlier, Marimont (1962) seems prepared to say what that might mean. Even Laming's differential coupler is a "black box"
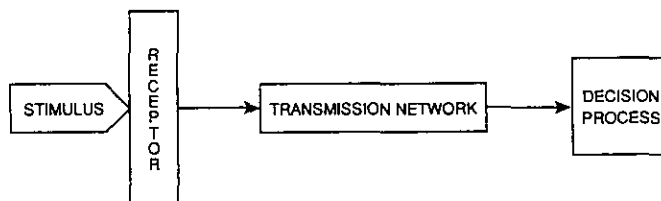

FIG. 1. Flow diagram of a sensory process. Stimulus and receptor mechanisms constitute the traditional realm of psychophysics. Stimulus and decision processes form the domain of signal detection theory. Ordinarily the transmission network linking receptor to decision is discounted. This network is the focus of the present paper.

with purely functional properties, whereas we ask for something quite specific.

Signal detection encountered consumer resistance at first because researchers did not see the point of a decision theory apart from studies of receptor function. What chance exists for a new analysis taking us right into the sensory labyrinth?

Conflicts between realism and simplicity are never-ending. One time-tested way to seek resolution is to invoke a stochastic process. A network can be kept simple by formulating it as a random walk, even when it is too complex for tracing signals.

In sensory communication, transmission is viewed as recurrent generation of a message at successive points or stages along a pathway. An input signal triggers the process. Transmission may fail, in which case the input event will be lost. In order to invoke a stochastic process, we say that information moves under the control of a propagation operator, i.e., a probability rule decreeing the birth and death of progeny at each successive stage. Repetition of the operator creates a cluster of impulses converging on the end stage and evolving as it moves. This leads to an output counting distribution rather than a detailed record of passage along the pathway. Counting variability combines with signal variability to limit detectability of incoming signals. Detection is said to occur after these noisy influences have come into play. Hence, decision data are predictable from, but not the same as, signal data.

II. SIMPLE OPERATORS—COMPLEX NETWORKS

The energy in a stimulus is transferred to receptor cells through some form of physical interaction. Activation leads to a discharge of primary neurons close to the receptor. Information marking the occurrence of a stimulus is contained in this discharge, but it takes place against a backdrop of unanalyzed, "spontaneous," noise events. These become incorporated into the discharge, leaving no way to distinguish signal from noise.

Decisions are made by counting output events. Counts that remain small reflect the effects of residual noise, whereas large counts tell us something new has happened. This formulation is classical detection theory, most clearly exemplified in Barlow's (1956) amendment of the quantum model of Hecht et al. (1942).

Suppose, as we have been suggesting, decisions are not made locally, but centrally. What happens when information is conveyed to a high-level decision region via less than perfect transmission lines?

Our first attempt to introduce such concepts inserted a Poisson link, obtained by superposing events occurring in separate neural pathways leading back from the periphery. (See McGill, 1967; Teich and McGill, 1976.) Momentary stimulus intensity was viewed as driving the Poisson rate.

Counting distributions were then worked out (McGill, 1967) and converted to detection laws for various inputs.

Acoustic sinewaves in wideband noise produced distributions identical to those reported for photodetection of laser energy. (See Peřina, 1967; Teich & McGill, 1976.) The approach also gave a good account of our acoustic signal/noise detection data. If sinewave signals are retained and input noise is reduced to zero, this mechanism will show just-noticeable differences proportional to the square root of intensity. The relation is a familiar one in visual research at low light levels where it is known as the deVries–Rose law (Graham, 1965). Actually, a Poisson network driven by a brief light flash should yield Neyman's Type-A distribution, but the Poisson and the Type-A are often difficult to tell apart. (See Teich, Prucnal, Vannucci, Breton, & McGill, 1982.)

Square-root detection is less common in audition. For pure tones it appears restricted to a narrow (10–15 dB) range above absolute threshold (Rabinowitz, Lim, Braida, & Durlach, 1976). Hence, although some of our predictions held up, our Poisson scheme failed with pure tones at middle intensities where the model was really aimed (Moore & Raab, 1974; Rabinowitz et al., 1976).

The main problem was that Poisson transmission did not alter detectability at these mid-level intensities. We could not find a way to put Weber's law in the output unless it was already there in the input. Eventually we grew skeptical of the approach.

Suppose, as an alternative, stimuli activate an amplifier–transmitter network. Following transmission, progeny of these inputs emerge at higher centers. We imagine the network to be a chain of discrete stages, each one more-or-less like the others. Transmission is idealized as an operator moving messages from stage to stage, creating and erasing events at each transition. All this is more than a Poisson network can do.

The number of stages is taken to be large. It puts us into an asymptotic domain produced by replicating a Markov operator, for example the one depicted in Fig. 2.

A transmission chain consists of r stages. Message-events are passed through a given stage only when they reproduce
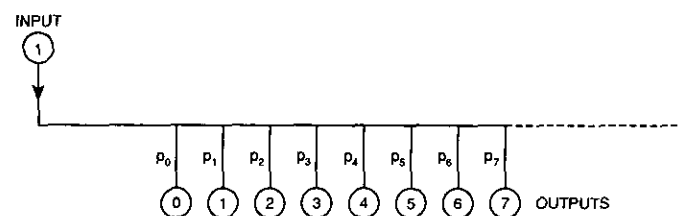

FIG. 2. A Markov-operator diagram of stochastic transmission for a single stage. Each input event is converted to one of several possible outputs with the probabilities p_0, p_1, p_2, etc. A possible outcome is that an input may completely disappear (probability p_0). Multiplication occurs whenever the output number equals or exceeds 2. A single input event starts the process. The network then generates a probability distribution of progeny at the output of the final stage.

themselves (or multiply, i.e., become two or more message-events) at that stage. The mechanism amplifies if progeny accumulate while passing from stage to stage. Moving an input through the entire chain requires at least one output at the end of the rth stage. Successful passage is never guaranteed. A sizable probability exists that all message-events will disappear, leaving no record of the input.

We characterize such a network as a stochastic branching chain. Not every chain displays all these properties at the same time. Some may not amplify. Others may be constructed so that the likelihood of erasures is small or nonexistent. When a chain performs both multiplication and erasure, the total event-count undergoes a random walk along the positive integers (and zero) as inputs move from stage to stage. It is then called a "birth–death" process. (See Bharucha-Reid, 1960, pp. 86–89.)

An important baseline condition reproduces the input message-event unchanged (on the average) at each stage. The chain transmits but does not amplify. This is done by balancing expected growth against expected losses. Messages might then seem to pass through such a network more-or-less intact, but when growth and decay are perfectly balanced in a chain that is sufficiently long, virtually all inputs are lost in transit. This result is an important key to differential coupling.

Long chains fail to transmit unless amplification, however small, occurs. This also means that a single input event, subjected to continuing small multiplications, must eventually give birth to an unbounded number of progeny.

At first thought, no room at all seems visible between these options: complete blockage or uncontrolled growth as the average birth rate passes from just under unity to just above unity. There are easy ways out of the dilemma, but the fact of its existence offers a sober warning about branching networks. They can easily get out of control.

Transmission noise is not "added" to messages. It is a built-in property of the network. Noise multiplies as messages move through successive stages, varying with the length and complexity of the branching chain. Hence, we contend with two basically different kinds of noise: (1) *additive noise* generated by events outside the transmitted message that somehow find a way into it and (2) *multiplicative noise* accompanying the passage of events from point to point in a network.

Barlow's dark light (1956) is an excellent example of additive noise. So are the "immigration" pulses of birth–death immigration processes. Immigration is an expected outcome of spontaneous activity in neural pathways. (See Bharucha-Reid, 1960, pp. 173–174.)

Weber's law is typical of multiplication noise. Disturbances early in the chain become subject to modification at later stages. Variance increases with each such alteration. It is evidently larger when noise sources are introduced

sequentially than when they are laid out side by side and added up.

Simplest Network

The device in Fig. 2 is a Markov propagation operator. It operates indiscriminately on every message event appearing at the input of each stage. A long chain brings on the asymptotic properties of the birth–death process. To characterize the chain we need to understand these asymptotic properties. Branching chains will amplify when propagation produces an average of slightly more than one output event for each input at each stage.

Such transmission is more complex than a Poisson link, but still oversimplified. Sensory mechanisms typically involve separate foci performing distinct operations. For example, a local amplifier might be followed by a common carrier. The first chain boosts signals, and the second conveys boosted data to a decision site. In combination, however, the output of the two chains is the same as a single branching chain with unvarying parameters (McGill & Teich, 1992). Hence, no great loss of generality is experienced by concentrating on single chains.

Consider first an extremely primitive network. Hecht *et al.* (1942) turned to it for an account of photons lost moving from a locus at the surface of the cornea to eventual absorption in the retina. Transmission is via a branching chain, but it is not regenerative. It produces random losses of photons passing through the ocular media. The retinal targets for these photons might then create a second transmission chain with new losses. Our argument is that the pair of sequences can be replaced by a single chain whose transition probabilities are the same from beginning to end. Both mechanisms will have exactly the same output. In this case no multiplications are permitted. An event either reproduces itself or disappears as it moves through either chain.

This Markov operator is diagrammed in Fig. 3.

Transition probability from state 1 at any given stage to state 1 at the next stage is fixed at p. Corresponding dropout

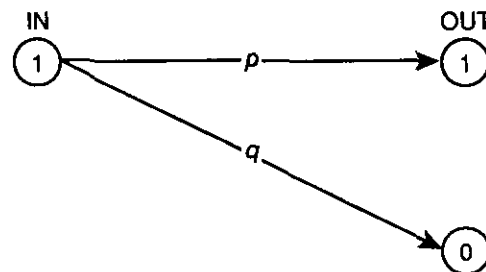**FIG. 3.** Markov diagram of a stage in a transmission network in which outputs are either transmitted or lost at random. A transition from state 1 to state 1 implies that an input event is transmitted to next stage. A transition from state 1 to state 0 means that the input is lost. No transmissions occur out of the zero-state. Any event entering the zero-state is lost, terminating its passage through network.

probability (transition from state 1 to state zero) is then also fixed and labeled q, where

$$p + q = 1.$$

Evidently these probabilities determine the statistics of passage through successive stages. The latter are independent, and the process is multiplicative. (Each stage triggers the next.) Accordingly, the probability-generating function at any point (see McGill & Teich, 1991b, Appendix) is obtained by applying a propagation operator to the generating function of the prior stage. (Also see Feller, 1957, pp. 248–264; Bharucha-Reid, 1960, pp. 19–27, for discussions of generating functions and their role in analyzing branching processes. Feller's book is a biblical text for students of stochastic processes. It is now in its 3rd edition, revised in 1970.)

Transmission through a sequence of these attenuating stages starts with a single input event. Our propagation operator is a discrete probability generating function (p.g.f.) moving the input from stage to stage. The p.g.f. is defined as follows:

$$G(s) \equiv \sum_{k=0}^{\infty} s^k p(k).$$

Allowed transitions in stage 1 limit the p.g.f. to two states: zero (message lost) or unity (message transmitted):

$$G_1(s) = q + ps. \tag{1}$$

At stage 2, a fraction q of the input messages is expected to drop out. Assuming homogenous operation, repeated applications produce for the rth stage

$$G_r(s) = q + pG_{r-1}(s),$$
$$= 1 + p(G_{r-1}(s) - 1), \tag{2}$$
$$\frac{G_r(s) - 1}{G_{r-1}(s) - 1} = p.$$

Multiplying this out over all r stages, when the initial condition is a single input (i.e., $G_0(s) = s$), we have

$$\frac{G_r(s) - 1}{G_{r-1}(s) - 1} \cdot \frac{G_{r-1}(s) - 1}{G_{r-2}(s) - 1} \cdots \frac{G_2(s) - 1}{G_1(s) - 1} \cdot \frac{G_1(s) - 1}{s - 1} = p^r. \tag{3}$$

Accordingly, the generating function for the output of the entire chain must be

$$G_r(s) = (1 - p^r) + p^r s, \tag{4}$$

where the iteration in Eq. (2) is now replaced by unconditional probabilities.

Probability of successful passage through the chain (state 1 output) is the transition probability p raised to a power set by the length of the chain. As the latter lengthens, probability of passage moves toward zero. Nearly all inputs are lost.

This primitive branching chain generates random deletions. Driven by a Poisson input, it eliminates input events at random but leaves the form of the distribution intact. (See Teich & Saleh, 1982.) Hecht et al. (1942) could then be assured that photons incident on the cornea retained a Poisson distribution at the retina despite losses in transit.

An easy extrapolation exists to continuous transmissions resulting in the well-known "pure death" process. Suppose the transition at any stage requires a brief interval of time Δt. If the latter is very short, there is little likelihood of a change of state during Δt. (The only change possible would be a sudden transition to state zero.) Hence, as Δt decreases, q also decreases proportionally. This stable ratio is ordinarily expressed as a fixed loss parameter (death rate) μ, where

$$\lim_{\Delta t \to 0} \left(\frac{q}{\Delta t} \right) = \mu. \tag{5}$$

The network probability then becomes

$$p^r = (1 - q)^r = (1 - \mu \, \Delta t)^r = \left(1 - \frac{\mu r \, \Delta t}{r} \right)^r. \tag{6}$$

If we define $t = r \, \Delta t$ as the transit time for the full network, the limit of the passage probability is

$$\lim_{\substack{r \to \infty \\ \Delta t \to 0}} (p^r) = e^{-\mu t}. \tag{7}$$

Network length is measured by its transit time t, regarded here as fixed, while loss or dropout rate per increment of time is measured by μ. Evidently, μ is not a probability since it may take on arbitrary values, whereas the limit of the state 1 output in Eq. (7) is a probability. As in discrete cases, this probability moves toward zero when the network lengthens (t increases) or when the dropout tendency grows large (μ increases).

III. NEUROLOGICAL AMPLIFIER NETWORKS

Birth and Death of Message-Events

Iterative restrictions on probability-generating functions can also be written for branching chains with message-events that multiply, reproduce, or drop out at each stage.

Output is then a random mixture of these various outcomes. Consider, for example, the operator in Fig. 4.

Chains constructed from such stages typically amplify their input. Message-events tend to increase in number passing from stage to stage. A single input thus becomes a small, fluctuating cluster, slowly increasing in size as it moves up the chain.

There is no requirement forcing pairwise multiplication (two message-events replacing a single one). It is an option. We might well find a random distribution of progeny at each stage. Pairwise multiplication offers a neat mechanism and an easy place to start.

How the various possibilities play out will be determined by values assigned to the transition probabilities in Fig. 4. In pairwise multiplication, p_0, p_1, and p_2 represent conditional probabilities governing transitions for each message-event moving through each stage. Hence

$$p_0 + p_1 + p_2 = 1.$$

Whenever $p_2 > p_0$ the network will amplify. This means that average output consists of more than one event for each input event. Stage-to-stage net births are all greater than unity.

The probability generating function representing the operator in Fig. 4 is

$$G(s) = p_0 + p_1 s + p_2 s^2. \tag{8}$$

This expression depicts a single stage and the possibility of pairwise multiplication. We want the counting distribution for the full network (i.e., the output at stage r). The best way

MAIN
TRANSMISSION
CHANNEL

CONTROL IN

STIMULUS IN

−1

(no change) 0
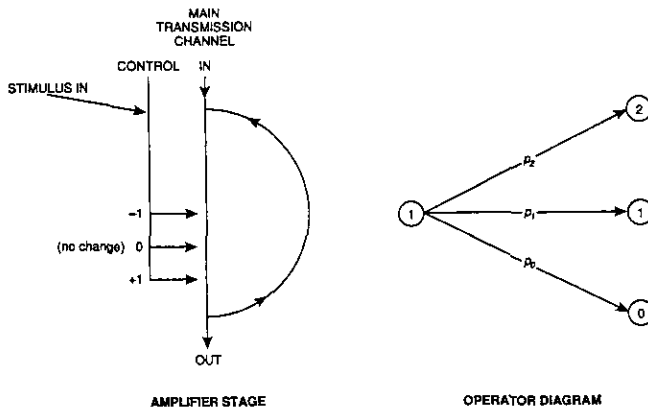
+1

OUT

AMPLIFIER STAGE

OPERATOR DIAGRAM

FIG. 4. A simple amplifier network. To the left is a functional diagram of a typical stage. The passage of message-events moving down the main transmission channel is modified by the control mechanism on the far left. Stimulus intensity sets the likelihood of multiplication. Feedback may occur repeatedly in a single site or a number of such amplifiers may be arrayed in sequence. The diagram to the right displays the Markov operator characterizing this mechanism. The operator shows allowable transitions for any event passing along the main channel and the probabilities of these transitions.

to proceed is to put an iterative restriction on the generating function at stage r:

$$G_r(s) = p_0 + p_1 G_{r-1}(s) + p_2 G_{r-1}^2(s). \tag{9}$$

This time, no easy way exists to multiply the process out. There is, however, a simple way to construct a time derivative. First, subtract $G_{r-1}(s)$ from both sides:

$$G_r(s) - G_{r-1}(s)$$
$$= p_0 - (p_0 + p_2) G_{r-1}(s) + p_2 G_{r-1}^2(s). \tag{10}$$

Next set $p_0 = \mu \, \Delta t$, as in Eq. (5); $p_2 = \lambda \, \Delta t$, in which λ is a fixed amplification parameter called the birth rate. Evidently, $t = r \, \Delta t$ must measure the length of the chain if successive stages differ by Δt. The time difference between the last two stages converges on a derivative:

$$\lim_{\Delta t \to 0} \left(\frac{G_t(s) - G_{t-\Delta t}(s)}{\Delta t} \right)$$
$$= \mu - (\mu + \lambda) G_{t-\Delta t}(s) + \lambda G_{t-\Delta t}^2(s). \tag{11}$$

In fact, identical restrictions hold for all adjacent stages. Accordingly,

$$G_\tau'(s) = \lambda(\mu/\lambda - G_\tau(s))(1 - G_\tau(s)). \tag{12}$$

The generating function in Eq. (12) depicts a chain evolving as a function of time τ. Our parameter t denotes a particular chain length; hence, a specific value of τ. We want the number of events in the final stage. As we see, a time derivative will produce the number, but the connection between stages and time disappears in the derivative. Equation (12) also describes a continuous chain in which all events are last-stage outputs emerging at varying times. (See Feller, 1957, p. 407; Bharucha-Reid, 1960, p. 87.) Hence, output counts are identical for the limit of a discrete chain of fixed length or a continuous chain observed for a fixed time period t. Both have the same counting distribution. In one case the outputs are bunched together; in the other, they form a discharge.

Equation (12) shows the quadratic factored to display its two roots. Since the coefficients of Eq. (11) sum to zero, a root will be found where $G_\tau(s) = 1$. As Δt decreases, the probability of a shift up or down becomes vanishingly small. A root then appears at $G_\tau(s) = 1$ in Eq. (12), reflecting this resistance to change.

A second root emerges in Eq. (11) when $G_t(s)$ reaches an asymptote as t increases without limit. In that case $G_t'(s) \to 0$. This second root is the asymptotic value of $G_t(s)$, namely

$$\lim_{t \to \infty} (G_t(s)) = \mu/\lambda.$$

This limit of the p.g.f. depicts a birth–death transmission process operating over infinite time. There is probability in state zero but nothing in any other finite state (since only the coefficient of s^0 survives at the limit). Hence, we have a significant probability of a zero count, a significant probability of an infinite count, and nothing between.

From the generating function at infinite time

$$\lim_{t \to \infty} (p_t(0)) = \mu/\lambda.$$

This is the limiting probability of a state-zero output (message lost). In branching processes it is called an *extinction probability*. When the loss parameter μ exceeds the gain parameter λ, extinction probability remains fixed at unity. Nearly all interesting cases arise when $\mu \leqslant \lambda$.

Evidently the operator in Eq. (9) has two different asymptotic limits. The derivative in Eq. (12) shows the number of stages growing large while transit time t remains finite. Now a further limit is obtained by letting t itself go to infinity. At this point the process breaks down. Every non-zero output is infinite. The initial limit, on the other hand, produces conventional distributions of output events.

Different limits explain many puzzling outcomes. For example, when $\lambda = \mu$ in a perfectly balanced network, extinction probability is found to be unity. This suggests that no messages can pass the chain. Yet when the differential equation, Eq. (12), is integrated in these circumstances, a conventional probability distribution results. Evidently we are at the intermediate limit (r infinite; t finite). Extinction is approached only as t becomes infinite. The probability of a zero count rises systematically with t up to the final extinction probability.

Adaptation Level

The chain's output is found by integrating Eq. (12) and then extracting a counting distribution from the probability generating function. A single input event triggers the first stage. Output spreads ultimately to many counts because multiplications and dropouts add variance continuously. This noise is a signature, stamping the propagation mechanism on all data passing through.

At adaptation level expected gains just balance losses, i.e., $\mu = \lambda$. Equation (12) becomes

$$\frac{G'_\tau(s)}{(1 - G_\tau(s))^2} = \lambda. \tag{12a}$$

Integration of both sides with respect to time produces

$$\frac{1}{1 - G_\tau(s)} \Bigg]_s^{G_t(s)} = \lambda t.$$

Since the chain is triggered by a single input pulse, the limits on the left-hand side must be $G_0(s) = s$ when $\tau = 0$ and $G_t(s)$ when $\tau = t$. We have

$$\frac{1}{1 - G_t(s)} = \lambda t + \frac{1}{1 - s},$$

$$G_t(s) = 1 + \frac{(s-1)}{1 - \lambda t(s-1)}. \tag{13}$$

This is the probability-generating function of a process with balanced gains and losses. The chain branches up and down continuously. To extract a counting distribution, we expand the right-hand side of Eq. (13) in powers of s. The coefficient of s^k will then be the probability, $p_t(k)$, that the chain's output consists of exactly k message-events. Expansion is easy. The general term of the probability distribution is

$$p_t(0) = \left(\frac{\lambda t}{1 + \lambda t}\right),$$

$$p_t(k) = \frac{(\lambda t)^{k-1}}{(1 + \lambda t)^{k+1}}, \quad k = 1, 2, 3, ..., \infty. \tag{14}$$

There is a detached "spike" of probability at zero count produced by message-events lost in transit. Non-zero counts have a geometric distribution, reflecting difficulty in achieving runs of multiplications as progeny pass from stage to stage in Fig. 4.

We are describing a normally quiet alerting-network in which any small change sets off an alarm. If the background level is even modestly high, near-silence prevails. A transient triggers the alarm. Hence, alerting-networks do not make comparisons. They react to changes. This was the idea that led Laming (1986) to differential coupling in increment detection.

A steady background sets up opposed processes, one attempting to amplify the input, the other trying to knock it out. At adaptation level, each cancels the other, creating a balance that persists until a transient occurs. The stochastic process defined by such a mechanism is called a "birth-and-death" process. (See Bharucha-Reid, 1960, pp. 86–89; Feller, 1957, pp. 407–411.) Outputs show a detached spike of probability at zero response and a shifted geometric distribution above zero; i.e., the distribution starts out at a count of unity. Variance increases with intensity, making transients of fixed size harder to detect as the background level increases. Thus, Weber's law is built into an alerting network if it is constructed in the form of a branching chain.

Output Counting Distribution

A sharp transient triggers the alarm. The balance between λ and μ is suddenly altered at the input. The birth rate

parameter escalates while the death rate remains unaffected. A secondary process rapidly comes into play aiming to restore equilibrium. As the parameters diverge, the chain shifts into a new counting mode in which mean output increases exponentially with the size of the momentary gap between birth and death rates.

What is this new distribution? The differential equation for its probability-generating function as the latter evolves with the length of the chain (i.e., with time) is given by

$$\frac{G_\tau'(s)}{(\mu/\lambda - G_\tau(s))(1 - G_\tau(s))} = \lambda, \qquad (12b)$$

where τ is the (time) variable of integration. Integration runs between limits $\tau = 0$ and $\tau = t$. When $\tau = 0$, $G_0(s) = s$, i.e., the process begins with a single event at the input. The solution of Eq. (12b) can be found in any standard table of integrals; for example, see Hodgman, Selby, & Weast (1954, p. 250). Our notation produces

$$G_t(s) = \frac{\mu/\lambda - \left(\dfrac{\mu/\lambda - s}{1 - s}\right) e^{-(\lambda - \mu)t}}{1 - \left(\dfrac{\mu/\lambda - s}{1 - s}\right) e^{-(\lambda - \mu)t}}. \qquad (15)$$

This generating function describes the transmission mechanism depicted in Fig. 4 when birth and death parameters λ, μ, and t are arbitrary. It reduces to Eq. (13) as $\lambda \to \mu$.

The counting distribution corresponding to Eq. (15) is well-known in stochastic processes. (See, for example, Bharucha-Reid, 1960, p. 88; McGill & Teich, 1992, p. 26.) We do not reproduce it here because our goal is a relation between signal intensity and detectability. Parameters such as λ, μ, and t are not observable in psychophysics. Neither are impulse counts for that matter, but detections are closely tied to them. Birth and death rates must be estimated to get at counting distributions. Thus, we need a way to relate λ, μ, and t to intensity.

Log Transform

When the difference $(\lambda - \mu)$ is linear with intensity, output increases exponentially. This is important for detecting weak signals, but it can be a serious problem as intensity increases. Large bursts of activity flood the system, paralyzing it. Ideally, outputs ought to grow the way a loudness function grows—as a fractional power of intensity. Adaptation in an alerting network makes even linear growth a feasible option for weak inputs. Alert signals need never be large to reach threshold if the baseline is virtual silence. Whatever form a transform takes, evidently its chief role is to curb output at high intensities.

The average output of the branching chain characterized by Eq. (15) can be calculated without elaborate ritual. The mean gain through any stage is

$$m = (1 + p_2 - p_0).$$

Each input rides through, is increased to two with probability p_2, or is wiped out with probability p_0. After r stages, the gain through the entire network will be

$$m_r = (1 + p_2 - p_0)^r.$$

Now apply the limiting process as $r \to \infty$ and $\Delta t \to 0$:

$$m_t = \lim_{\substack{r \to \infty \\ \Delta t \to 0}} (m_r) = \left(1 + \frac{(\lambda - \mu)r\,\Delta t}{r}\right)^r$$
$$= e^{(\lambda - \mu)t}. \qquad (16)$$

This is the exponential growth we find worrisome. A propagation operator enforces such growth whenever λ exceeds μ, unless intensity sets parameters via a transform rather than directly.

Earlier we spoke of energy exchanges in the sense organ as if they were the primal events of the stimulus message. This is the prevailing view in sensory psychology and neurophysiology. Among vision researchers, for example, the principle that light energy carries critical stimulus information is widely accepted. The same is true in audition although disagreements must be recorded. For instance, Jeffress (1964, 1968) showed that his electronic model of an amplitude detector yielded results very close to human performance. Green and Swets (1988/1966), following a similar line, proved that when a sinewave masked by Gaussian noise is known exactly, an amplitude ratio will measure the signal's detectability in the noise. Laming (1986) too asserted that amplitudes provide the crucial intensity information in auditory signal processing.

Readers encountering these arguments for the first time may be forgiven expressions of bafflement. Energy and amplitude are functions of each other. If we know how information flows in one, it is easy to calculate how it flows in the other. What can the argument be about?

Generally it is about an underlying mechanism. For example, if negative masking is presumed to occur just above absolute threshold, amplitude measurements become appropriate. They produce the desired outcome whereas energy measurements typically do not. Arguments such as this can never be settled. Each measurement translates readily into any of its counterparts. The measurement we use depends on the mechanism assumed to be operating.

These issues are brought into sharp focus by the need to relate branching-chain parameters λ, μ, and t to intensity and to estimate them from intensity measurements. Which measurements?

We are guided by a proposal of Rushton (1961) in a study of visual coding. He worked with the optic nerve of *Limulus*, the horseshoe crab, using data provided by Fuortes (1959). Rushton showed that neural firing rates are proportional to the logarithm of stimulus intensity over a wide range of intensities and backgrounds. The relation he settled on was $\log(1 + ax)$, in which x is the intensity variable and a is constant. Rushton speculated that this transform would bring linearity to other modalities as well. Indeed, we now know that primary auditory neurons in mammals behave similarly (Sachs & Abbas, 1974; Teich & Khanna, 1985).

Sense organs are required "at the lower end of their intensity range to detect signals ... approach(ing) the prevailing noise level," whereas in the upper ranges the "task is not merely to detect" but to "make fine discriminations with regard to the time course, spatial location, and quality" (Rushton, 1961, p. 177). To achieve these objectives in the present application, Rushton's prescription is applied as

$$\beta \ln(1 + ax) = (\lambda - \mu) t, \tag{17}$$

where β is a proportionality constant expressing the arbitrary base of the logarithm. In Eq. (17), a and β are constants chosen to match intensity units on the left-hand side with the branching parameters on the right. The difference $(\lambda - \mu)$ is the net birth rate of neural events stage-by-stage, and t measures the length of the chain.

When this log transform is introduced into Eq. (16), it creates a power-law relation between stimulus intensity and the chain's output:

$$m_t = e^{(\lambda - \mu) t}$$
$$= e^{\beta \ln(1 + ax)}$$
$$= (1 + ax)^{\beta}. \tag{18}$$

Mean output of the branching chain increases as a power function of intensity. The process has two stages: (1) a log transform from intensity to the net birth-rate of the chain as set forth in Eq. (17), and (2) exponential growth of message-events within the chain. The combination produces Eq. (18). It behaves more-or-less as a loudness function for pure tones or other narrowband stimuli. Given the low-level intensities characterizing barely detectable increments, we can generally take $\beta = 1$ over the narrow range of a psychometric function. Full impact of the exponent is encountered only when measurements are attempted on a wide range of intensities. (See Teich & Lachs, 1983.)

Our argument leading to Eqs. (17) and (18) is based chiefly on Rushton's work, but it is similar in content to proposals by Marimont (1962) and MacKay (1963) seeking to explain the discrepancy between nerve impulse data and magnitude estimates when both are viewed as functions of intensity.

Marimont's idea was that a log transform at receptor level can be counterbalanced by an "exponentiating" device higher up, yielding a power law in the output. MacKay approved this approach but noted that no solid basis was known in neurophysiology for the required second-stage exponential transform. He suggested it might be part of an evaluative mechanism underlying magnitude estimates. Here, we claim that a branching process provides an ideal rationale for the missing exponential link.

Ryan, Braverman, Woolf, & Axelsson (1989) used an imaging technique, 2-deoxyglucose (2-DG) autoradiography, to show that cellular activity spreads as pure-tone stimuli ascend through the auditory tract. The data offer strong indications of branching. Moreover, a recent study by Zeng & Shannon (1994) analyzed loudness matches made by hearing-impaired patients with implants at different locations in the auditory tract. The data show intensity coding to be a two-stage process similar to the one envisioned for the visual system by Marimont in 1962. Loudness balances reported by Zeng & Shannon suggest logarithmic compression in the periphery followed by an exponential expansion higher up. The two stages combine to produce a power function. This is just what Eqs. (17) and (18) say.

In Eq. (18), the product $(\lambda - \mu) t$ is the log of the mean count t seconds after stimulation. This transform may well be propagated along the entire length of a neural transmission chain, preserving its form at every stage, as suggested by Rushton. A configuration such as that would be designed to maintain a constant balance between excitation and inhibition through λ and μ at every point. At the output, progeny would not all emerge at the same instant. An alerting signal embodies a brief discharge varying in size with stimulus magnitude. If an energy exchange occurs in the sense organ, the counting record will preserve it. So at least for the restricted range of the psychometric function, the mean output count in Eq. (18) is proportional to stimulus energy.

Power Function Output

Apply the transform in Eq. (17) to the branching-chain generating function in Eq. (15) so that mean output is a power function of stimulus energy. We find

$$G_x(s) = \frac{\pi(x) - \left(\frac{\pi(x) - s}{1 - s}\right)\left(\frac{1}{1 + ax}\right)^{\beta}}{1 - \left(\frac{\pi(x) - s}{1 - s}\right)\left(\frac{1}{1 + ax}\right)^{\beta}}. \tag{19}$$

Here $\pi(x) = \mu/\lambda$ is the extinction probability corresponding to an incremental intensity x above adaptation level. Extinction probability is at or very near unity in the vicinity

of adaptation level, dropping off to zero as incremental intensity increases. Without difficulty we obtain

$$G_x(s) = 1 + \frac{(1+ax)^\beta (s-1)}{1 - \gamma(x)(s-1)}. \tag{20}$$

Introduced into Eq. (20) is a new intensity variable

$$\gamma(x) = \frac{(1+ax)^\beta - 1}{1 - \pi(x)},$$

combining the power function and extinction. This new measure simplifies the argument materially. We analyze its significance in the next section on *ultra-linear output*.

The p.g.f. in Eq. (20) resembles an earlier expression for a balanced network given in Eq. (13). Here, the generating function extends to increments above adaptation level and embodies the log transform in Eq. (17). A subscript x in the p.g.f. shows that it is now intensity-based. As with Eq. (14) for a balanced line, we see a spike of probability covering lost inputs in the zero state, as well as a geometric distribution of output counts reflecting the consequences of multiplication noise in transmission.

These points are readily verified in the counting distribution extracted from Eq. (20),

$$p_x(0) = \frac{\pi(x)\,\gamma(x)}{1 + \gamma(x)},$$

$$p_x(k) = [1 - p_x(0)] \left(\frac{1}{1+\gamma(x)}\right)\left(\frac{\gamma(x)}{1+\gamma(x)}\right)^{k-1}, \tag{21}$$

where $k = 1, 2, 3, ..., \infty$; $\pi(x) = \mu_x/\lambda_x$; $\gamma(x) = [(1+ax)^\beta - 1]/[1 - \pi(x)]$.

Output probability $p_x(k)$ is geometric. Zero output indicates failure to survive. The non-zero portion of Eq. (21) is a *Bose-Einstein* (BE) distribution shifted up by one count. A separate spike of probability is added at the zero count. The BE parameter [in this case $\gamma(x)$] measures an expected number of events.

Branching generates larger variances than we encounter in transmission devices passing information without multiplication. These effects are not subtle. They modify the output away from Poisson form and toward Bose-Einstein (geometric) form. Signal detectability is thus altered in the direction of Weber's law.

If branching chains mediate messages passing between sense organ and brain, transmission can be expected to behave in the manner just described. It is not clear in any given case whether the message passed is an alert-signal or a portion of a literal copy of the input. The distinction is important because the two types of transmission follow different detection laws. This point is expanded in our section on *alarm signals vs representations*.

Ultra-linear Output

Rushton's log transform leads to Eq. (20), an intensity-based generating function for a branching chain. It happens that ax and β are usually small in increment detection. Accordingly, a linear approximation of $(1+ax)^\beta$ is generally suitable for threshold work. Psychometric functions are handled by letting $\beta = 1$ and absorbing the exponent into the units constant a:

$$G_x(s) = 1 + \frac{(1+ax)(s-1)}{1 - \gamma(x)(s-1)}.$$

Now let x, the incremental intensity, go to zero. This moves the generating function to adaptation level. We need

$$\lim_{x \to 0} \gamma(x) = \lim_{x \to 0} \left(\frac{ax}{1 - \pi(x)}\right) = \gamma_0.$$

With suitable restrictions on $1 - \pi(x)$, a limit develops because extinction probability approaches unity as x approaches zero. We let γ_0 symbolize this limit. Accordingly, at $x = 0$ we obtain

$$G_x(s) = 1 + \frac{(s-1)}{1 - \gamma_0(s-1)},$$

which, except for notation, is Eq. (13).

Intensities are expressed as increments above adaptation level. With $\beta = 1$ we have

$$\gamma(x) = ax/[1 - \pi(x)],$$

$$= \gamma_0 + ax, \tag{22}$$

where γ_0 represents an equivalent background intensity, and x is an increment above it. Extinction, incremental intensity, and adaptation level are then governed by a simple relation

$$\pi(x) = \frac{\gamma_0}{\gamma_0 + ax}. \tag{23}$$

These expressions lead to an intensity-based solution for the branching-chain differential equation, Eq. (12),

$$G_x(s) = 1 + \frac{(1+ax)(s-1)}{1 - (\gamma_0 + ax)(s-1)}. \tag{24}$$

Extraction of the output counting distribution is simple:

$$p_x(0) = \frac{\gamma_0}{1 + \gamma_0 + ax},$$

$$p_x(k) = [1 - p_x(0)] \left(\frac{1}{1 + \gamma_0 + ax}\right)\left(\frac{\gamma_0 + ax}{1 + \gamma_0 + ax}\right)^{k-1};$$

(25)

here $k = 1, 2, 3, ..., \infty$; a is a units constant; x is an incremental intensity above adaptation level; γ_0 is an intensity parameter corresponding to adaptation level. Equation (25) is Eq. (21) with $\gamma(x) = \gamma_0 + ax$ and $\beta = 1$.

The counting distribution in Eq. (25) represents the response of a branching chain (see Fig. 4) when a transient stimulus adds an intensity increment to a steady background and when intensities are subject to the log transform in Eq. (17).

At adaptation level ($x = 0$), this chain will emit a variable discharge averaging one output event for each input. The most probable output at adaptation level is silence, since $p_x(0) \gg p_x$ ($k > 0$) unless γ_0 is very small. Any transient generates a burst of activity whose mean count varies with increment intensity. We argued earlier that over the narrow range of a psychometric function, mean output count grows in rough linear agreement with increment energy. This conclusion is built into Eq. (25), where the mean is $1 + ax$.

Our expression in Eq. (25) centers on an adaptation level γ_0, fixed by the background level. Evidently the magnitude of γ_0 cannot be linear with intensity if Eq. (18) is correct, but the near-silence of a balanced line means that we need not specify an exact relation. We may take γ_0 as having whatever value emerges from experiment. Since increments are generally quite small in detection experiments, output will remain linear with increment energy no matter what value is chosen for γ_0. Thus, our ultra-linear restriction seems fairly realistic. In effect, adaptation levels create conditions that resemble absolute threshold wherever adaptation occurs.

Alarm Signals vs Representations

The branching chains described thus far can be characterized as "alert" or "alarm" signals activated whenever a sudden transient perturbs a continuous background. In increment detection the alert is set off by any marked change in level. Comparisons are not involved. Alert signals are thus not representations, à la Hecht et al. (1942) in which a random number of photons are detected at the retina. An alarm is triggered by the arrival of any number of photons. Alarms are defined by their warning role. They indicate only that something new has happened.

Branching chains may also operate in a representational mode to boost a weak signal. If the input is a cluster of events, a mechanism of this type will produce an amplified version of the process envisioned by Hecht et al. (1942). The number of photons at the retina obeys a Poisson distribution. These peripheral events serve as input to a birth-and-death transmission chain with $\lambda \gg \mu$. Unlike the alerting mode, birth and death rates remain fixed. They are system constants reflecting losses in transmission and the amplification needed to detect very weak signals despite such losses. These opposed influences combine to produce an output fluctuation equivalent to a low-level masking noise, plus a boost in the average size of the output cluster. Square-root-law detectability persists because output means and variances remain proportional. Both are governed by the number of input events.

Closely related results are found when the restriction to a Poisson number of inputs is lifted and a Neyman Type-A distribution of inputs is substituted. The latter characterizes nerve impulses propagating in peripheral optic nerve fibers (Saleh & Teich, 1985). Representational transmission in this instance is similar to the visual threshold mechanism suggested by McGill (1967) and studied by Teich et al. (1982). Here, however, we have a cascade of amplification stages rather than a single stage.

Since birth and death rates remain fixed in representational transmission, auditory representations also code intensity into a number of input events. Each will then give birth to a separate branching process, even when all are propagating within a single physiological chain. Hence, output distributions are equivalent to a sum: the convolution of separate branching processes making up the representation. Each is defined by an expression resembling Eq. (20). Intensity alters the number of components.

These convolution distributions are not simple. Although each zero-state is reduced in size compared with chains in which birth and death rates are equal, zero-states are still detached from the rest of the distribution. However, if the number of chains is large, the central limit theorem will take over, driving aggregate output toward a compound Poisson distribution (McGill & Teich, 1992) or toward a noncentral negative binomial when transmission is a birth–death–immigration process (Li & Teich 1991, 1992).

The important distinction when a representation is detected rather than an alert signal is that separate chains connect each input event to its respective central counterpart. Intensity is coded into the number of input events. The count monitored at the decision area is then the aggregate output of all individual chains making up the representation. A square-root detection law results.

Both mechanisms, alarm signals and representations, probably operate side by side. The one actually at work in detection would depend on conditions. Increment detection of pure tones at high or moderate intensity levels is readily mediated by an alert signal, whereas at absolute threshold stimuli appear to be detected as representations. The function governing increment detection would then shift from

deVries–Rose form to Weber's law at some mid-level intensity. The data indicate that this is indeed what happens (Rabinowitz *et al.* 1976).

IV. DISCUSSION

Probable Structure of a Sensory Network

Branching chains transmit either (1) alarm signals or (2) representations of the initial neural response developed in the sense organ. A threshold pure tone or a small spot of light would normally activate only one or a very few chains in its representation. If this number stays small as the background increases, output will still appear to come from a single chain. The branching structure in Fig. 5 illustrates how this single-chain impression develops if the signal spreads over only a few chains. Hence, there ought to be an intensity region just above zero where thresholds remain nearly fixed. Negative masking does not occur.

In the alerting mode, outputs are dominated by the spike of probability in the zero-state. When more than one branching chain becomes activated, aggregate output is biased away from the negative binomial distribution. There is a bulge at the low end of the distribution attributable to the zero-state of each component. If the number of chains proves to be large, as when intense signals or wideband backgrounds are studied, the central limit theorem will eventually drive aggregate output toward Gaussian form, ironing out such anomalies. Besides, a detection criterion set in the upper tail of the convolution distribution will generally show little alteration traceable to the zero-state. Means and variances of aggregates are easy to calculate, but the distributions themselves are difficult because of the detached zero-states.
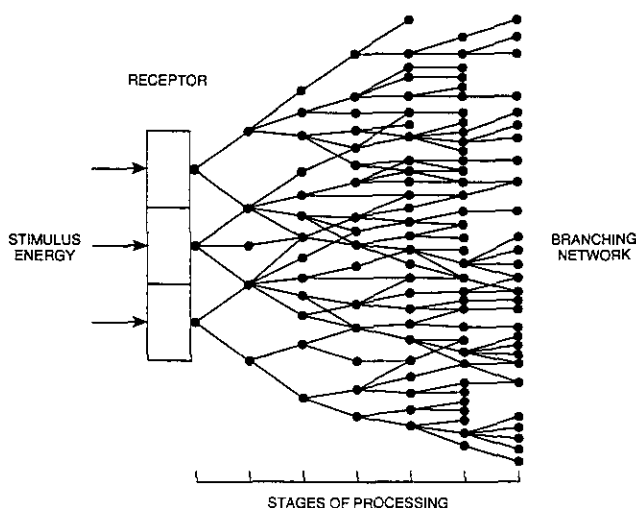

RECEPTOR

STIMULUS ENERGY

BRANCHING NETWORK

STAGES OF PROCESSING

FIG. 5. Interactions among nearby chains leading back from a sensory organ and forming a branching network. If there is room to spread out and the number of separate chains is not too large, they may be treated as a single chain with altered parameters. At successive stages the volume of activity increases but spreads out over an increasingly larger region.

When inputs are stochastic, the statistics of the stimulus are overlaid on the variability of the branching chain (or chains) involved in transmission. Output variation is then a compound of the signal intensity distribution and the chain's multiplicative noise (Diament & Teich, 1992). In earlier papers we attempted to study such compounding, assuming that the stimulus created a Poisson-like message flow. Results were unsatisfactory for pure tone stimuli at moderate and high intensities (McGill, 1967). In order to correct the deficiencies of the Poisson approach, we turned next to branching chains in which each start-up event is said to produce a variable cluster of output pulses. As we see, a branching chain adapts to continuous pure-tone backgrounds and displays modified Bose–Einstein (i.e., geometric, not Poisson) outputs in responding to brief increments. The robustness of this mechanism reflects the simple structure of Markov-process noise. Despite our skepticism of Laming's (1986) arguments in McGill & Teich (1989), an intimate relation between differential coupling and these branching chains is now evident.

Importance of Pure Tones

A psychophysical transform converting intensity changes in input signals to just-noticeable differences in perception, remains as elusive today as when first conjectured in the 19th century. An adequate formulation requires much more than the stimulus. We need mechanisms for altering the input into what is in fact perceived and at least several transmission stages appear to be involved.

If we were certain that transmission is mediated by branching chains, the next steps would be clear. We would work out counting probabilities for the mixture of acoustic-stimulus fluctuations and internal noise generated in a branching chain. These calculations are not child's play but they can be done. Unfortunately, we are insecure about virtually all our premises. Consequently, we try to keep stimuli and experimental conditions as simple as possible in order to follow what is going on.

Visual data present special challenges for detection because classical light is stochastic. (See Teich & Saleh, 1988, 1990.) There is a real problem about separating Poisson photon-number fluctuations from multiplicative noise generated in transmission. This inherent confounding has helped to assure the longevity of the approach of Hecht *et al.* (1942). Pure stimulus models of detection are always simpler and more attractive to argue than those with transmission noise.

Auditory research offers better control and somewhat simpler tests because an analytically simple stimulus, the so-called "pure tone," is relatively easy to deliver to the ear intact if signal durations are not too short. Admittedly we do not know what the ear considers analytically simple, but would be well-advised to avoid getting too cute. Pure tones

can be fixed in frequency, phase-locked, and precisely timed. Variability is thus confined to the tone generator's error level which, with today's equipment, can be made negligible. Fluctuations in perception must then be attributable to noise contributed by the auditory system itself.

Essentially for these reasons, in recent years we have witnessed a revived interest in one of the traditional problems of auditory psychophysics: pure-tone intensity discrimination.

Narrowband Intensity Discrimination and Increment Detection

The experimental relation between intensity of a pure-tone masker and the size of a just-detectable increment (or decrement) is shown in Fig. 6. Data are from McGill & Goldberg (1968b). They were obtained with pure tones at 1000 Hz and signal durations of 15 or 20 ms. Also plotted are closely related results obtained by Campbell & Lasky (1967). A masking function is developed over eight log units of background intensity (80 dB). For the first 20–30 dB above absolute threshold, the slope is unity: Weber's law. Figure 6 shows little evidence of square-root detection near absolute threshold as in the paper of Rabinowitz et al. (1976), which compiled results from as many as 14 different studies. If square-root detection is in the McGill–Goldberg data, the range must be very small. Above 30 dB, slope is 0.9 in log coordinates: a "near miss" to Weber's law. Short bursts of wideband Gaussian noise do not show any of these departures. Slope is unity over virtually the entire range (Miller, 1947).

What accounts for the difference? McGill & Goldberg (1968b) sought an explanation in a power function, i.e., a "loudness" function relating masker intensity to the Poisson

FIG. 6. Pure-tone intensity discrimination. Abscissa is sensation level (dB above threshold) of a 1000-Hz test tone. Ordinate is sensation level of a tone just detectably weaker than the test tone. (McGill & Goldberg, 1968a, b). Unit slope depicts Weber's law. Slope fitted by least-squares is 0.905. Data are for three listeners: (closed squares) 15-ms tones, (closed circles) 15-ms tones, and (closed triangles) 20-ms tones. Open circles represent data on 20-ms pure tones obtained by Campbell & Lasky (1967). The Campbell–Lasky function is moved horizontally for the best match to the McGill–Goldberg data.

process they assumed to intervene between receptor and higher centers. They argued that such indirect loudness measurements supported their Poisson transmission hypothesis. Others were skeptical. The idea attracted few adherents, and a conviction grew that something else was involved in the near miss. Teich & Lachs (1979) attributed Weber's law to neural refractoriness and the near miss to spread of excitation in the cochlea. Subsequent neurophysiological measurements in primary auditory-nerve fibers (Teich & Khanna, 1985) revealed that refractoriness was too weak to convert the ascending spike train to one in which Weber's law might be found. Furthermore, aggregation of point processes in a collection of ascending fibers tends to wash out the effects of refractoriness in individual units. We now believe that the likely origin of Weber's law is multiplicative noise generated within branching chains. Spread of excitation across these chains appears to be responsible for the near miss.

Viemeister (1972) discovered that when pure tones are centered in a slot or notch of low-level, wideband masking noise, Weber's law is restored in increment detection. Masking slope remains at unity over nearly the full range of background intensities. Moore & Raab (1974) repeated Viemeister's experiment and concluded that the notch effectively blocks listening off center frequency as a way to improve detection. If it could be shown that a given network displays Weber's law in individual chains stretching back from the ear, and also the near-miss phenomenon when excitation spreads uncontrolled as intensity increases, to adjacent chains serving nearby frequencies, it might explain Viemeister's result.

In a recent chapter, McGill & Teich (1991a) made such a demonstration using branching chains operating in the alerting mode and an output counting distribution equivalent to Eq. (25). Pure-tone detection follows Weber's law in individual branching chains because of the proportionality between means and standard deviations in geometric distributions. If increasing intensity caused alert signals to spread out, activating more and more remote transmission chains, a near miss should develop. Pure tones embedded in notched noise would encounter the edges of the notch as excitation spreads, preventing activation of more remote chains and limiting bandwidth to the region of the notch. This should effectively fix the degrees of freedom in the aggregate alert signal, restoring Weber's law.

McGill & Teich (1991a) did not incorporate an adaptation level into their treatment. Their demonstration was carried out for $\gamma_0 = 1$, since in that special case the zero-state is not detached. Hence, the argument was appropriate for pairs of comparison tones with a brief (nearly) blank interval between each member of the pair, i.e., classical intensity discrimination.

A superior format for making the same experimental comparisons would be a run of just-detectable increments in
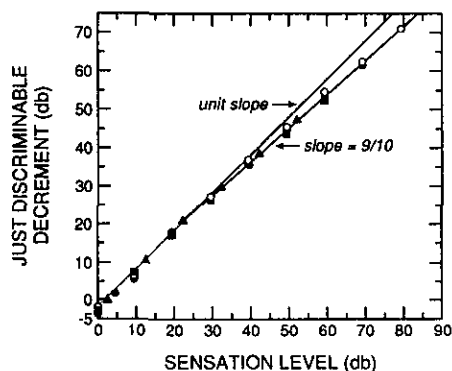
a steady tonal background. The adaptation level is then set by the steady background. Increment detection would be mediated directly by an alert signal. No explicit stimulus comparisons would ever be required. If excitation spread, more than one branching chain would be set up, and if spread were uncontrolled, we should encounter the near-miss phenomenon.

What happens when the two formats are compared in a narrowband case involving single chains? Is there a measurable difference between intensity discrimination and increment detection when the same pairs of pure tones are used to generate comparisons? Laming (1986) posed the issue originally, formulating it as a general question in sensory analysis.

The output counting distribution of a single branching chain operating in alert-signal mode is given in Eq. (25). An observer making paired comparisons on the basis of such counts will generate correct choices in accord with the rule

$$P(c) = P(j > k) + \tfrac{1}{2} \cdot P(j = k), \tag{26}$$

where $P(c)$ is the probability of a correct response in two-alternative forced choice, j is the count generated by the incrementally stronger pure tone, and k is the count generated by the weaker pure tone.

The probability of a correct response in two-alternative forced-choice (2AFC) is the probability that the strong signal count exceeds the weak count added to the probability of guessing correctly when they coincide.

If Eq. (25) is inserted (see McGill & Teich, 1991b), we find

$$
P(c) = \left(\frac{\gamma_0}{1 + \gamma_0 + aE_0}\right)\left(\frac{2 + \gamma_0 + 2aE_s}{2 + 2\gamma_0 + 2aE_s}\right)
$$
$$
+ \left(\frac{1 + aE_0}{1 + \gamma_0 + aE_0}\right)\left(\frac{1 + aE_s}{1 + \gamma_0 + aE_s}\right)
$$
$$
\times \left(\frac{1 + 2\gamma_0 + 2aE_s}{2 + 4\gamma_0 + 2aE_0 + 2aE_s}\right). \tag{27}
$$

The choice probability is developed from a chain adapted to a pure-tone background whose intensity fixes the size of γ_0 and incremented by two different pure-tone signals. Here the incremental energies are set at $x = E_0$, or alternatively $x = E_s$, slightly larger.

The detection format is a steady background or pedestal tone with a pair of superimposed increments differing slightly in intensity. All components are at the same frequency. When $P(c)$ is plotted against E_0, we have a psychometric function describing the detectability of E_s against E_0, where each lies above an adaptation level, γ_0. Counting distributions are based on Eq. (25).

With everything simple, why should the psychometric function be so formidable? There are two reasons. First,

there is a problem with the spike of probability corresponding to a count of zero. The first term of Eq. (27) is generated when the count k attributable to the weaker tone is zero. The second term appears when $k > 0$. A second reason is that we are computing a three-parameter (E_0, E_s, γ_0) expression in order to encompass several different experimental comparisons. For example, when adaptation level is set for background silence ($\gamma_0 = 0$), the psychometric function in Eq. (27) reduces to

$$P(c) = \frac{1 + 2aE_s}{2 + 2aE_0 + 2aE_s} \approx \frac{E_s}{E_0 + E_s}. \tag{28}$$

This is essentially the result reported earlier by McGill & Teich (1991a). A slight difference arises because McGill & Teich set $\gamma_0 = 1$, whereas here we take $\gamma_0 = 0$. If E_0 is of modest size, the difference is negligible. Equation (28) is the performance of an energy observer detecting two different levels of narrowband noise by making comparisons between them (see Green & McGill, 1970, Eq. (9a); McGill & Teich, 1991a, Eq. 27). We find noise here because an ultra-linear chain is a proportional counter, and the stochastic process defining the chain is a narrowband noise generator. At the terminus, a listener would not be able to decide whether the noise came from the auditory pathway or from a narrowband noise input. Branching converts deterministic inputs into stochastic outputs by infusing them with multiplicative noise generated in transmission.

If a pure tone spreads out over a frequency region encompassing a number of independent alarm generators, the narrowband noise in Eq. (28) will not always appear to widen, although its average level will change. This reflects the influence of the zero-state on the aggregate response of a small group of chains. Weber's law continues to characterize auditory alarm signals until inputs are intense enough to spread excitation over a wider region. Then the near miss appears. We should not assume that these same things happen in all experimental formats. Different formats activate different mechanisms.

Let $E_0 = 0$ in Eq. (27), eliminating one comparison tone. The format is now a steady-state background and a transient of fixed intensity detected against it: pure-tone increment detection. Both are at the same frequency. In that event,

$$
P(c) = \left(\frac{\gamma_0}{1 + \gamma_0}\right)\left(\frac{2 + \gamma_0 + 2aE_s}{2 + 2\gamma_0 + 2aE_s}\right)
$$
$$
+ \left(\frac{1}{1 + \gamma_0}\right)\left(\frac{1 + aE_s}{1 + \gamma_0 + aE_s}\right)\left(\frac{1 + 2\gamma_0 + 2aE_s}{2 + 4\gamma_0 + 2aE_s}\right),
$$
$$
P(c) \approx \frac{1 + \gamma_0/2 + aE_s}{1 + \gamma_0 + aE_s},
$$
$$
P(c) \approx \frac{\gamma_0/2 + aE_s}{\gamma_0 + aE_s}. \tag{29}
$$

The second term of Eq. (29) should be negligible compared to the first whenever the background level γ_0 is modestly high. Since the second term develops when background contributions are greater than zero, it follows that a branching chain must be virtually silent at adaptation level, a fact now established in several different ways.

Equation (29) again portrays an observer discriminating two levels of narrowband noise, one of them created by an alarm signal. To compare this result for increment detection against Eq. (28) involving intensity discrimination, we need to plot $P(c)$ against E_0 in both cases, but E_0 does not signify a background intensity in Eq. (29). Accordingly, let

$$\gamma_0 = aE_0.$$

This shifts the origin in Eq. (29) from adaptation level down to zero intensity. The same shift applied to incremental intensity produces

$$E_s = E_0 + \Delta E.$$

The new background level E_0 and increment ΔE can now be substituted back into Eqs. (28) and (29). When background intensity is modestly greater than zero (i.e., silence), these notional changes alter the appearance of the choice probability in Eq. (29) to a form that is much easier to recognize:

$$P(c) \approx 1 - \frac{1}{2} \cdot \frac{E_0}{E_s}. \tag{30}$$

Both intensities are measured against a zero origin, and Eq. (30) is comparable to its counterpart in Eq. (28). Both exhibit Weber's law: detectability is determined by the intensity ratio. When the two detection curves are plotted against $\log E_0$, we find that increment detection (Eq. 30) is about 3 dB more sensitive than intensity discrimination (Eq. 28). A narrowband comparison such as this would be appropriate near threshold or at low background levels.

Since branching chains are nearly silent at adaptation level, an unconventional strategy for increment detection emerges from Eq. (30). The observer listens at the output. If anything comes through, anything at all, even a single count, an increment is claimed for the interval in which it appears. If nothing comes through, the observer guesses. Counts rarely occur in both intervals. If it should ever happen, the observer guesses again or discards such trials from the data. This strategy is essentially the one described by Sakitt (1972) for absolute visual threshold. It yields results quite close to Eq. (30).

We see that detection strategies involving an adaptation level are likely to be different from pure intensity discrimination where explicit comparisons are required.

Experimenters generally prefer increment detection. They regard it as more sensitive than other threshold methods. Our analysis suggests that a basis for this preference exists in the information at the output end of a branching chain. [Compare Eqs. (28) and (30).]

Multiple Chains

Suppose we consider the same calculation assuming a number of independent chains, as might happen when a moderately intense pure tone serves as a background in the alerting mode. Exact solutions for the psychometric function are difficult in such cases, but a normal approximation can be constructed provided the number of chains is sufficiently large. Earlier, McGill & Teich (1991a) showed that in pure-tone intensity discrimination, i.e., $\gamma_0 = 1$, detectability is fixed whenever the signal-to-background level and v, the number of chains, remain constant.

What happens in increment detection? To learn the answer we need a wideband generalization of Eq. (30). Begin with the mean and variance of Eq. (25), now modified so that all intensities are referred to the same zero origin:

$$\text{mean count} = 1 + a\,\Delta E,$$
$$\text{variance} = 2aE_0 + a\,\Delta E \cdot (1 + 2aE_0 + a\,\Delta E). \tag{31}$$

These parameters describe a single-chain alerting distribution when an increment ΔE is added to the tonal background. Unmodified backgrounds are given by Eq. (31) with $\Delta E = 0$.

If the same energies are spread out uniformly across v independent transmission channels, the counting distribution is driven toward Gaussian form and the parameters become

$$\text{mean count} = v + a\,\Delta E,$$
$$\text{variance} = v[2aE_0/v + (a\Delta E/v) \tag{31a}$$
$$\cdot (1 + 2aE_0/v + a\,\Delta E/v)].$$

The standard normal deviate describing a paired comparison is then

$$\frac{a\,\Delta E}{[4aE_0 + (a\,\Delta E)(1 + 2aE_0/v + a\,\Delta E/v)]^{1/2}} \simeq d.$$

Constant terms divided by the background intensity become small and can be neglected. Hence,

$$\frac{v \cdot \Delta E/E_0}{2 + \Delta E/E_0} \simeq d^2.$$

Thus for v, the number of independent sensory channels, sufficiently large, and $\Delta E/E_0$ sufficiently small, the standard normal deviate characterizing increment detection in an ultra-linear network may be approximated by

$$(v\, \Delta E/2E_0)^{1/2} \simeq d. \tag{32}$$

Recall that the process was initiated by an energy exchange in the receptor. When this increment spreads out over a large number of channels responding to a steady-state background or pedestal, the aggregate action resembles amplitude detection. It would not be correct to define the stimulus in such cases as an amplitude. We say instead that amplitude-like detection occurs when an energy increment activates many independent branching chains. A near miss to Weber's law will also develop if v increases in an uncontrolled way as energy spreads to nearby receptor sites.

The sensitivity advantage of increment detection is preserved with wideband alarm signals.

V. SUMMARY

We have shown that stochastic models of sensory transmission are fairly easy to construct. The systems installed by nature behind the major senses almost certainly include mechanisms to amplify weak signals. As the principle governing amplification we offer a repetitive Markov operator that multiplies message-events according to a fixed probability rule as data pass from point to point in transmission. Continuous repetition of the operator moves information through successive "stages" of transmission, increasing the average message-size and adding an internal noise to the amplified record. We attribute Weber's law in large part to this internal noise.

There are, of course, many alternative views (cf. Treisman, 1964). Researchers have identified Weber's law with neural refractoriness (Teich & Lachs, 1979) and with receptor saturation (Lachs, Al-Shaikh, Bi, Saia, & Teich, 1984). Once sensory systems are addressed in their full complexity, explanations of Weber's law can be constructed from a variety of principles. All these need to be sorted out.

As noted, the auditory system probably amplifies the data it moves to higher centers. The machinery should then involve at least two processes: amplification and transmission. Here we formulate branching chains as unitary mechanisms with the same properties everywhere. Despite the apparent conflict, this single-chain approach is not a serious problem. Multi-step processes, in which one chain amplifies and a second transmits, are also equivalent to a *single* branching chain. The two systems differ in structure, but their outputs are identical. See McGill & Teich (1992).

Introducing a log transform early in transmission, as nature seems to do, counteracts the effects of amplification. The transform first compresses intensity, and then the network promptly undoes it. Such counterbalancing leads to a power function in the output.

The log transform is not introduced as a mathematical convenience to make calculations easy. A solid body of evidence establishes that log conversions similar to Eq. (17) actually occur early in transmission. Rushton (1961, pp. 171–181) and Zeng & Shannon (1994) found the transform in unrelated studies of sensory coding. Rushton based his conclusions on impulse data recorded from the visual receptor of the horseshoe crab, whereas Zeng & Shannon analyzed loudness judgments in hearing-impaired humans implanted with electronic transducers in the cochlea or brain stem. The formulation by Zeng & Shannon of intensity coding at "high" frequency is very close to the one presented in this paper. In the cat, primary eighth-nerve fibers synapsing directly with hair-cell receptors in the inner ear show similar effects. A log transform is found across the junction. (See Teich & Khanna, 1985.)

A unique property of birth-and-death transmission is its adaptation level. Branching chains with perfectly balanced birth and death rates quickly drift into a state of adaptation to steady-state backgrounds. Output is largely (but not entirely) suppressed in the adapted condition. Hence, changes are detected in nearly the same way at all background levels. A log relation between increment intensity and the net birth rate of a branching chain will rarely generate large data counts in detection experiments. Within the rough 10 dB range of a psychometric function, there is little difference between log and linear scales. The log transform has its principal impact on the location of the adaptation level.

The perspective contributed by transmission is spectacular. Intensity is not necessarily coded into the neural image of an input signal. Instead, it can regulate the buildup of events inside an alarm mechanism. Steady-state backgrounds cause the latter to come into equilibrium so that multiplications and losses are nearly equal. Responding then virtually ceases until a change of some sort occurs. Detection involves an alert signal instead of a copy or image of the input.

Our results suggest that if background level is sufficiently high, the onset of an increment will trigger the alert signal. Listeners seem to be reporting on the properties of this alarm rather than the increment itself.

REFERENCES

Barlow, H. B. (1956). Retinal noise and absolute threshold. *Journal of the Optical Society of America*, 46, 634–639.

Barlow, H. B. (1977). Retinal and central factors in human vision limited by noise. In H. B. Barlow & P. Fall (Eds.), *Photoreception in vertebrates*. London: Academic Press.

Bharucha-Reid, A. (1960). *Elements of the theory of Markov processes and their applications.* New York: McGraw-Hill.

Brown, J. (1965). The structure of the visual system. In C. H. Graham (Ed.), *Vision and visual perception.* New York: Wiley.

Campbell, R., & Lasky, E. (1967). Masker level and sinusoidal-signal detection. *Journal of the Acoustical Society of America, 42,* 972–976.

Diament, P., & Teich, M. C. (1992). Evolution of the statistical properties of photons passed through a traveling-wave laser amplifier. *IEEE Journal of Quantum Electronics, 28,* 1325–1334.

Ehret, G., & Merzenich, M. (1985). Auditory midbrain responses parallel spectral integration phenomena. *Science, 229,* 1245–1247.

Feller, W. (1957/1970). *An introduction to probability theory and its applications:* Vol. 1. (2nd ed., 1957/ 3rd ed. rev., 1970) New York: Wiley.

Fuortes, M. (1959). Initiation of impulses in the visual cells of *Limulus. Journal of Physiology, 148,* 14–28.

Geisler, W. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review, 96,* 267–314.

Graham, C. (1965). Color mixture and color systems. In C. H. Graham (Ed.), *Vision and visual perception.* New York: Wiley.

Green, D., & McGill, W. J. (1970). On the equivalence of detection probabilities and well-known statistical quantities. *Psychological Review, 77,* 294–301.

Green, D., & Swets, J. (1988/1966). *Signal detection theory and psychophysics.* Originally published, New York: Wiley (1966). Revised and reprinted, Los Altos, CA: Peninsula Publishing (1988).

Hecht, S., Shlaer, S., & Pirenne, M. (1942). Energy, quanta and vision. *Journal of General Physiology, 25,* 819–840.

Helson, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *American Journal of Psychology, 60,* 1–29.

Hodgman, C., Selby, S., & Weast, R. (Eds.). (1954). *Mathematical Tables from Handbook of Chemistry and Physics.* Cleveland, OH: Chemical Rubber Publishing Co.

Jeffress, L. (1964). Stimulus-oriented approach to detection. *Journal of the Acoustical Society of America, 36,* 766–774.

Jeffress, L. (1968). Mathematical and electrical models of auditory detection. *Journal of the Acoustical Society of America, 44,* 187–203.

Kuhn, T. (1957). *The Copernican revolution.* Cambridge, MA: Harvard Univ. Press.

Lachs, G., Al-Shaikh, R., Bi, Q., Saia, R. A., & Teich, M. C. (1984). A neural-counting model based on physiological characteristics of the peripheral auditory system. V. Application to loudness estimation and intensity discrimination. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-14,* 819–836.

Laming, D. (1986). *Sensory analysis.* London: Academic Press.

Li, T., & Teich, M. C. (1991). Bit-error rate for a lightwave communication system incorporating an erbium-doped fibre amplifier. *Electronics Letters, 27,* 598–600.

Li, T., & Teich, M. C. (1992). Performance of a lightwave system incorporating a cascade of erbium-doped fiber amplifiers. *Optics Communications, 91,* 41–45.

MacKay, D. M. (1963). Psychophysics of perceived intensity: A theoretical basis for Fechner's and Stevens' laws. *Science, 139,* 1213–1216.

Margolis, H. (1987). *Patterns, thinking, and cognition.* Chicago: Univ. of Chicago Press.

Marimont, R. B. (1962). Model for visual response to contrast. *Journal of the Optical Society of America, 52,* 800–806.

McGill, W. J. (1967). Neural counting mechanisms and energy detection in audition. *Journal of Mathematical Psychology, 4,* 351–376.

McGill, W. J., & Goldberg, J. (1968a). Pure tone intensity discrimination and energy detection. *Journal of the Acoustical Society of America, 44,* 576–581.

McGill, W. J., & Goldberg, J. (1968b). A study of the near-miss involving Weber's law and pure-tone intensity discrimination. *Perception and Psychophysics, 4,* 105–109.

McGill, W. J., & Teich, M. C. (1989). A unique approach to stimulus detection theory in psychophysics based upon the properties of zero-mean Gaussian noise. (Review of Laming's *Sensory Analysis*). *Journal of Mathematical Psychology, 33,* 99–108.

McGill, W. J., & Teich, M. C. (1991a). Auditory signal detection and amplification in a neural transmission network. In M. L. Commons, J. A. Nevin, & M. C. Davison (Eds.), *Signal detection* (Chap. 1, pp. 1–37). Hillsdale, NJ: Erlbaum.

McGill, W. J., & Teich, M. C. (1991b). Simple Models of Sensory Transmission. (CHIP Report No. 132). University of California, San Diego: Center for Human Information Processing.

McGill, W. J., & Teich, M. C. (1992). Alerting Signals and Auditory Detection in Branching Chains. (CHIP Report No. 134). University of California, San Diego: Center for Human Information Processing.

Miller, G. A. (1947). Sensitivity to changes in the intensity of white noise and its relation to masking and loudness. *Journal of the Acoustical Society of America, 19,* 609–619.

Moore, B., & Raab, D. (1974). Pure tone intensity discrimination: Some experiments relating to the "near miss" to Weber's law. *Journal of the Acoustical Society of America, 55,* 1049–1054.

Peřina, J. (1967). Superposition of coherent and incoherent fields, *Physics Letters, 24A,* 333–334.

Rabinowitz, W. M., Lim, J. S., Braida, L. D., & Durlach, N. I. (1976). Intensity perception. VI. Summary of recent data on deviations from Weber's law for 1000-Hz tone pulses. *Journal of the Acoustical Society of America, 59,* 1506–1509.

Rushton, W. A. H. (1961). Peripheral coding in the nervous system. In W. Rosenblith (Ed.), *Sensory Communication* (pp. 169–181) Cambridge, MA: MIT Press.

Ryan, A. F., Braverman, S., Woolf, N. K., & Axelsson, G. A. (1989). Auditory neural activity evoked by pure-tone stimulation as a function of intensity. *Brain Research, 483,* 283–293.

Sachs, M. B., & Abbas, P. J. (1974). Rate versus level functions for auditory nerve fibers in cats: Tone burst stimuli. *Journal of the Acoustical Society of America, 56,* 1835–1846.

Sakitt, B. (1972). Counting every quantum. *Journal of Physiology, 223,* 131–150.

Saleh, B. E. A., & Teich, M. C. (1985). Multiplication and refractoriness in the cat's retinal ganglion-cell discharge at low light levels. *Biological Cybernetics, 52,* 101–107.

Tanner, W., & Swets, J. (1954). The human use of information: I. Signal detection for the case of the signal known exactly. *Transactions of the IRE Professional Group on Information Theory, PGIT-4,* 213–221.

Teich, M. C., & Khanna, S. M. (1985). Pulse-number distribution for the neural spike train in the cat's auditory nerve. *Journal of the Acoustical Society of America, 77,* 1110–1128.

Teich, M. C., & Lachs, G. (1979). A neural-counting model incorporating refractoriness and spread of excitation: I. Application to intensity discrimination. *Journal of the Acoustical Society of America, 66,* 1738–1749.

Teich, M. C., & Lachs, G. (1983). A neural-counting model incorporating refractoriness and spread of excitation: III. Application to intensity discrimination and loudness estimation for variable-bandwidth noise stimuli. *Acustica, 53,* 225–236.

Teich, M. C., & McGill, W. J. (1976). Neural counting and photon counting in the presence of dead time. *Physical Review Letters, 36,* 754–758, 1473 (erratum).

Teich, M. C., Prucnal, P. R., Vannucci, G., Breton, M. E., & McGill, W. J. (1982). Multiplication noise in the human visual system at threshold: 1. Quantum fluctuations and minimum detectable energy. *Journal of the Optical Society of America, 72,* 419–431.

Teich, M. C., & Saleh, B. E. A. (1982). Effects of random deletion and additive noise on bunched and antibunched photon-counting statistics. *Optics Letters, 7,* 365–367.

Teich, M. C., & Saleh, B. E. A. (1988). Photon bunching and antibunching. In E. Wolf (Ed.), *Progress in Optics XXVI* (pp. 1–104) Amsterdam: Elsevier.

Teich, M. C., & Saleh, B. E. A. (1990). Squeezed and antibunched light. *Physics Today*, **43**(6), 26–34.

Treisman, M. (1964). Noise and Weber's law: The discrimination of brightness and other dimensions. *Psychological Review*, **71**, 314–330.

Viemeister, N. (1972). Intensity discrimination of pulsed sinusoids. The effect of filtered noise. *Journal of the Acoustical Society of America*, **51**, 1265–1269.

Zeng, F.-G., & Shannon, R. V. (1994). Loudness-coding mechanisms inferred from electrical stimulation of the human auditory system. *Science*, **264**, 564–566.