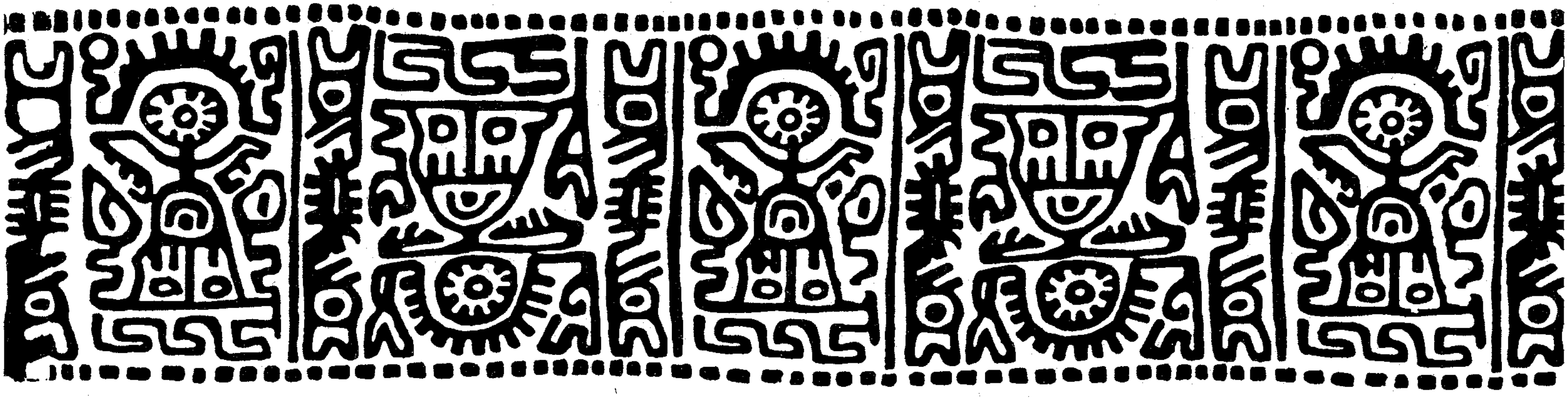
ALERTING SIGNALS AND AUDITORY DETECTION IN BRANCHING CHAINS

William J. McGill and Malvin C. Teich



CENTER FOR HUMAN INFORMATION PROCESSING

UNIVERSITY OF CALIFORNIA, SAN DIEGO
LA JOLLA, CALIFORNIA 92093-0109

CHIP 134
DECEMBER 1992

ALERTING SIGNALS AND AUDITORY DETECTION IN BRANCHING CHAINS

William J. McGill and Malvin C. Teich

Center for Human Information Processing
Department of Psychology, 0109
University of California, San Diego
La Jolla, CA 92093

CONTENTS

This paper is addressed to the information channel linking the ear and the brain. It is one of nature's most remarkable specialized networks. We formulate auditory detection as a stochastic process in which a sudden cascade of events in an alarm network signals the occurrence of a transient stimulus.

1. INTRODUCTION

Auditory detection is interpreted nearly always as a stimulus phenomenon. Transmission mechanisms are thought of primarily as refining information passed from the inner ear. Decision centers in the brain are assumed to operate with faithful copies of the stimulus itself. The latter is said to be recovered in the same way that image enhancement transforms noisy radio signals into elegant photographs. There is no question that a great deal of signal processing goes on in the auditory system. Convincing signs of sharpening are found nearly everywhere in the frequency domain, as well as in the reconstruction of auditory space.

Intensity is a different story. Although the ear exhibits extraordinary sensitivity near its threshold, increment detection is not very acute, even with phase-locked pure tones at constant frequency. The auditory system will sometimes have difficulty penetrating energy differences between closely related sine wave inputs. Why?

When weak inputs are amplified, the aim is to boost them up to audible level, not to analyze small differences. Differences might be rendered difficult to detect if the amplifier mechanism were somehow infiltrated by random noise. How in fact do random influences enter into auditory signal detection?

There are no easy answers to such questions. Internal noise is dutifully acknowledged as a factor in audition, but it is considered small compared with the impact of receptor mechanisms. Occasional reference is made to the idea that weak inputs might be amplified in transmission, generating a tell-tale internal noise, but few researchers deem it important to examine the consequences. We now raise the whole issue seriously.

Black-Box Representations:

Cognitive scientists favor black-box diagrams for representing the information flows characteristic of the major senses. These complex representations are arguably more realistic than pure receptor models. The latter, as we have noted, generally ignore system elements that process information. A black-box approach demands concern for all known stages of sensory communication. If serious attention is given to data transmission, controversial results are virtually guaranteed. Black-box analysis relies heavily on computer simulation to provide empirical validation of unorthodox outcomes, and more generally to convey a sense of how these systems work.

Despite much computer analysis, it is difficult to claim improved understanding of auditory detection on the basis of simulations. At key points, we are forced to rely on conjecture for envisioning obscure details of the system's organization. Because of these gaps, black-box diagrams remain vague and tentative as well as unusually complex compared with mathematical models.

Approximations must be invented for poorly-understood black-box components. For example, the branching chains developed in this paper are stochastic amplifiers. Electronic amplifiers might then be used as substitutes. The problem is that we seldom understand the machinery well enough to calculate how closely a conventional amplifier resembles a stochastic chain, or whether outcomes might be subject to change when minor details of the chain are modified. Only by comparing simulations can we tell whether diagrams are really different.

Despite these weaknesses, black-box approaches are useful because of the sheer complexity of the auditory system. Myriad connections form billions of alternative routes. It is futile to attempt signal tracing through systems of this size without some kind of roadmap of the organization. Simulations, combined with new knowledge of the underlying physiology, can suggest useful ideas about organization. Yet, because sensory networks are so poorly understood, the ideas remain vague. We need more than that.

Poisson Transmission Models:

An important alternative to black-box analysis invokes probability theory to deal with uncertainties posed by continuous branching in neural pathways. The chain of events between ear and brain is then expressible as a stochastic process. While this approach is more precise than a black-box diagram, it has its own problems. By far the most popular mechanism is the dead-time-modified Poisson point process (see Teich, 1992). It turns up regularly in auditory firing data. But detection as measured by listening experiments does not seem to display Poisson origins. Moreover, Poisson mechanisms lack memory, yet even at early transmission stages firing patterns show interspike correlations (Teich, 1992).

A clever way around these localized memory effects treats the auditory channel as an aggregate. Intensity information is then said to be conveyed by volume activity in the channel. Nonrandom bunching, even fixed spacing on any single line, diminishes in importance as new lines are included. A Poisson process (or a near miss) emerges whenever the aggregate is large, and pulses along the separate lines are sparse so that they appear to be coming at random. Memory loss is then a property of the collective, not the separate transmission lines.

This type of aggregate randomization is very common. It occurs, for example, whenever separate lanes of traffic are superposed. At a given measuring point, superposition makes an event in any traffic lane indistinguishable from events in other lanes. A monitoring system, aggregating data from independent pathways, will lose key details of the flow; but since the task is to detect small changes in overall traffic volume, jettisoning detail poses no real problem. In any event, when intensity information is represented via superposition, a Poisson-like aggregate flow results. This argument was set forth originally in the now-famous Cox-Smith (1954) superposition theorem, and was later proved rigorously in Çinlar's (1972) study of the mechanisms governing superposed point processes.

Despite its appealing simplicity, superposition fails to account for the major phenomena of auditory increment detection. Over the years, experimental data have been distilled into a limited number of key functional relations or laws. These same laws should emerge from theoretical treatments of Poisson detectability, but it does not appear to happen (McGill and Teich, 1991a).

At the other extreme, black-box representations are ordinarily too complex to offer significant insights on detection. They fail to explain why any of the empirical relations arise. We seem to require transmission mechanisms more complicated than a Poisson process and more explicit than a black-box diagram.

Markov processes offer such intermediate possibilities. When a Markov operator is configured to produce message growth, repetition of the operator at successive stages of transmission generates multiplicative increases in progeny. Both impulse count and the level of internal noise tend to grow with chain length. Moreover, a repetitive operator will propagate memory effects all through the transmission chain.

Markov Transmission:

If the flow of information between receptor and brain is taken to be a Markov process, what impact would that have on detectability? Passage along an auditory transmission channel with Markov properties generates messages that evolve as they move. If the chain is long, outcomes might be affected more by propagation noise than by acoustic noise in the original signal. Repeated branching will create random disturbances that grow as the chain lengthens. Internal noise might then overwhelm acoustic noise. A conjecture of this sort puts us at odds with most stimulus-oriented treatments of detection where acoustic noise is seen as the only random element, and where neural processing is thought to recover the stimulus virtually intact.

Increment Detection In Branching Chains:

A branching chain is a pulse-driven stochastic process. Information is passed along the chain as a constantly shifting impulse count reflecting the fates of individual message-events working their way up the chain to its final stages. Branching can occur in at least two ways. New message-events suddenly appear in the original chain, or the latter itself branches, creating new message-events in separate pathways.

In the simplest branching chains, transmission is initiated by a single start-up event. Hence, the number of start-up events is not necessarily linked to the size of an increment. Instead, stimuli modulate the growth of new message-events in the chain. There is a single start-up pulse and a characteristic output counting distribution corresponding to each stimulus increment.

A remarkable property of chains branching to the zero-state (i.e., chains in which message-events disappear) is that continuous stimulation can render the chain nearly silent. This introduces a possibility that such chains might serve as change-of-state detectors. The auditory system is capable of creating not only an internal reconstruction of the stimulus environment, but added alerting signals triggered off only when changes occur. Increment detection could be mediated by such alerting signals, requiring no direct stimulus comparisons of any kind.

Difference discrimination, on the other hand, compares pairs of stimuli, one having an added increment and the other without it. Each member of the pair is separated from its counterpart by a blank interval. Insertion of the blank shifts adaptation level to zero so that single alerting signals can no longer be used to detect intensity changes. In effect, the format of difference discrimination forces paired comparisons, whereas increment detection does not. Laming (1986), who first looked at this problem, argued that increment data imply some kind of alerting signal. If branching chains are the transmission vehicle, amplified alarm-signals can be fashioned easily.

Multicomponent Branching Chains:

It is difficult to envision a single Markov operator covering the entire length of a branching chain from the inner ear to a terminus near the brain. More plausible would be a mechanism with two or more distinct structures: a feedback amplifier in the receptor area, and a longer secondary transmission network leading back to high-level decision centers. Amplifier circuitry would pick up a minute change in ambient intensity and convert it to a cascade of impulses at or near the receptor site. The cascade should vary in average size depending on the magnitude of the stimulus change. This initial alert-signal would then be injected into a chain with fixed parameters, conveying an amplified noisy record of the transient to a central registration point. Amplification parameters of the first chain would be set by the size of the stimulus increment, whereas the second operates with fixed properties, transmitting any stimulus received at its input.

If a way can be found to combine these two components into a single equivalent chain whose parameters are resultants of the separate components, we might then represent initial amplification and subsequent transmission as a single branching chain with the same parameters everywhere. Equivalent chains constructed via such averaging prove to be "birth-and-death" stochastic processes of the type described by Bharucha-Reid (1960, pp. 86-89).

Serial Processing:

Chain structures are serial processors. Auditory transmission is easily formulated as a type of serial processing, because nature's task is to move stimulus information from the inner ear to a central location. We learned long ago from studies of memory storage that serial and parallel mechanisms, though conceptually quite different, often produce indistinguishable outcomes (McGill, 1963, pp.344-347). Townsend (1976) and Vorberg & Ulrich (1987) have demonstrated outcome-equivalence among a variety of serial and parallel processors. No doubt parallel alternatives to branching chains will be invented. Our object here is to understand the mechanisms governing transmission noise, not to claim insights on currently hidden details of auditory networks.

Outline Of The Paper:

We examine the proposition that key properties of auditory signal detection are due to an internal transmission noise.

We construct a family of related chain-like processes to serve as models of transmission. Then we make arbitrary cuts in such branching chains, leading to two-step or multi-step mechanisms with rules that change at the cut. Multi-step chains are shown to be equivalent in output to simple birth-death processes.

We consider immigrant pulses from sources external to the transmitted message. They become incorporated into signals passing up the network, adding another kind of noise.

We study how a single branching chain grows silent under continuous stimulation, creating an adaptation level. What happens when such adapted chains encounter sudden intensity changes? Alerting signals are generated, notifying higher centers of a change in state. This important class of signals was overlooked prior to Laming's (1986) pioneering work on increment detection.

Finally we compare single-chain auditory alerting signals, with stimulus representations, i.e., full-scale copies or mappings of peripheral stimulus patterns.

2. BRANCHING CHAINS

Despite efforts to minimize contamination, information flowing from one place to another is subject to unexplained perturbations. These disturbances are random and can usually be traced to the transmission process itself. In designing a communication system, a critically important task is to minimize its internal noise.

Nature's systems pose a somewhat different problem. The design is given. Our job is to figure out how it works. We cannot speculate usefully about auditory internal noise unless we know what mechanism creates the noise. In the beginning at least, this information is just what we do not have. Data on anatomy and physiology are voluminous and growing rapidly, but transmission is so complex that no single organizing principle has yet suggested itself. Our chief guidance comes from discriminations made by the intact ear. Research has pinpointed stimuli that differ morphologically yet cannot be separated by the system. These invariances reveal what kind of information is getting through. Generally, the energy content of a brief acoustic stimulus defines its detectability.

Discrimination data show also that a sizable internal noise exists, although we do not understand its origins. The inferential leap to a processing mechanism is not easy because input and output are drawn from different measurement domains. Cross-correlations between stimulus events and the data on which a decision center acts, are a near-impossibility. Despite our growing knowledge of auditory topography, we are forced to proceed obliquely in conjecturing how noise might arise in transmission.

For instance, the internal noise accompanying a stochastic flow of information might act as a signature helping to identify the mechanism. A passive detection process in which message-events are simply aggregated and counted will typically obey a square-root detection law. In vision research this type of detectability is actually found. It is known as the DeVries-Rose law. (See Bouman, 1961.) Just-detectable increases are proportional to the square-root of the background level.

As noted earlier, auditory detection does not seem to behave this way. Detectable increments are either directly proportional to background level, or show fractional exponents between .85 and .95. Key auditory data are illustrated in Figure 1. Square-root-law detectability is simply not in evidence.

FIGURE 1 HERE -

One possible explanation of the discrepancy is examined in this paper: a multiplicative noise acting on neural messages from auditory receptors. This is typical of processes with repeated staging. Noise created in early stages becomes subject to further modification at later stages. Disturbances generated in this sequential way are obviously larger than noise that is simply added up.

The reasoning is indirect but still fairly easy to test. We know, for example, that auditory reception and transmission add unwanted noise to input signals. There is just too much difficulty detecting brief intensity increases in pure tones . The latter have fixed

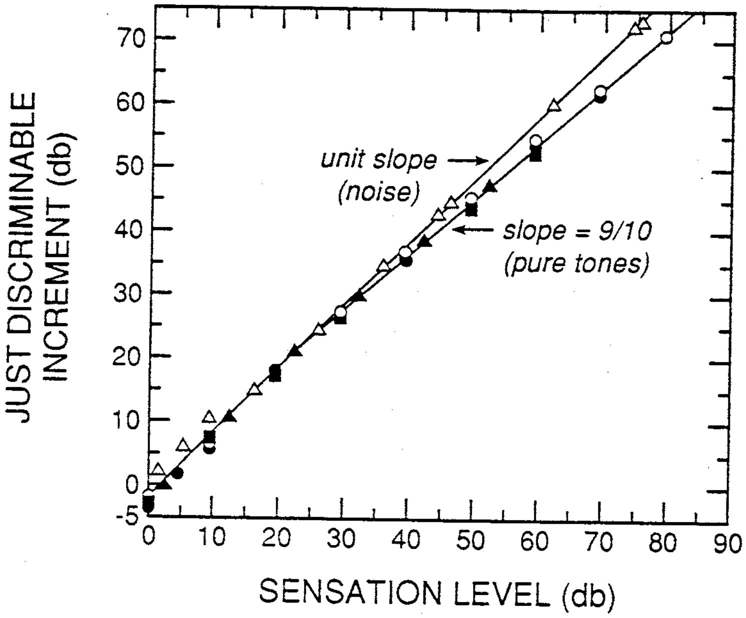


Figure 1: Experimental data on auditory increment detection. Abcissa is sensation level (db. above threshold) of background. Ordinate is sensation level of test signal just detectably different from background. Plotted data are mid-points of individual psychometric functions.

Unit slope depicts Weber's law. It characterizes detectability of noise increments over nearly entire range. Pure tone slope fitted by least squares is actually .905. Data on certain pure tones based on decrement detection: dark squares, dark circles, and dark triangles. These are from McGill and Goldberg (1968a, b). Open circles are based on 20 msc increments as determined by Campbell and Laskey (1967). Open triangles taken from Miller's (1947) classic study of wide-band noise increment-detection. Also see Green's (1960) study of noise increment-detection, confirming and extending Miller's findings.

Miller and Campbell-Lasky functions shifted horizontally for best match to McGill-Goldberg data. Below 20 db SL, wide band noise and phase-locked pure tone increment/decrement detection are reasonably well-fitted by Weber's Law. No indication of square-root detectability.

Viemeister (1971), and Moore & Rabb (1974) showed that "near-miss" to Weber's law (.905 slope) for pure tones above 20 db SL, is probably due to spread of excitation away from primary locus in receptor. When spread is controlled experimentally, Weber's law reappears.

energy. Fluctuations are negligible. Our inability to distinguish small differences suggests not only an internal noise, but one that varies with signal level (Green, 1967). We cannot be certain that the noise arises in transmission, but it is a clear possibility.

Listening experiments do not confirm Poisson models of auditory data-flow (McGill & Teich, 1991a, pp. 20-21). Pure tones and other fixed-energy signals generate Weber's law (Fig. 1) in controlled experiments on increment detection; not a square-root law. Hearing data, as we have already noted, are consistent with mechanisms whose variance is proportional to the mean-squared, not the mean. One such is a Markov transmission chain.

Markov Transmission:

Imagine a transmission system constructed in the form of a chain with r stages as illustrated in Figure 2. At each stage all events passing through are subject to the rule shown just below the chain. The diagram depicts a Markov operator applied whenever a message-event transits a given stage. We show the operator as having three distinct outcomes. With probability p_2 the message-event multiplies, becoming two separate events. It can also disappear (be wiped out) with probability p_0 . Or with probability p_1 the message-event passes through the stage unchanged, emerging exactly as it entered.

- FIGURE 2 HERE -

For each message-event at each stage:

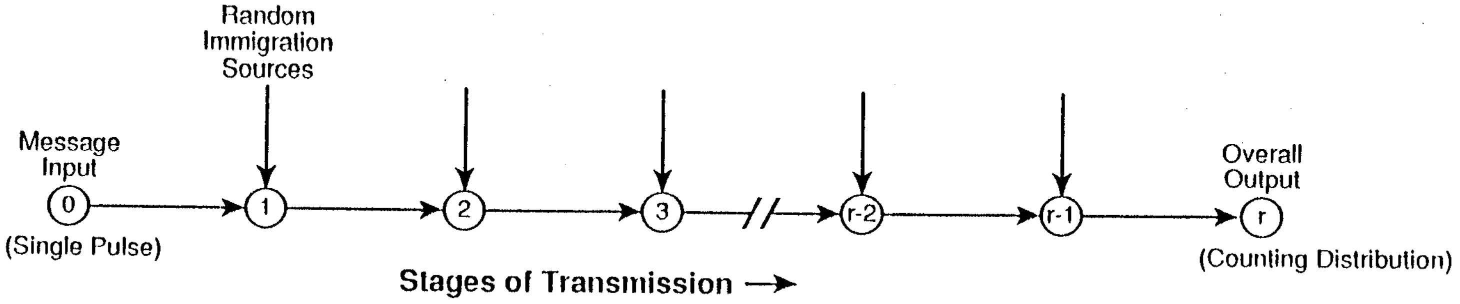
$$p_o+p_1+p_2=1.$$

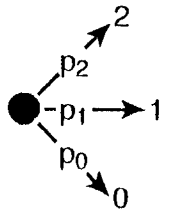
In other words one of the three outcomes occurs at each stage. There is no special magic associated with pair production. It is one possibility. Multiplication might result in a distribution of progeny. Our bread-and-butter question is whether we can solve for the output count given any particular assumption about the Markov operator. Later we consider chains with arbitrary progeny.

Our initial concern is limited to simple alerting signals such as might occur when the background intensity undergoes a sudden, brief transient. These alert-signals start with a single pulse or at most a small number of pulses. We want to know how a long chain modifies this initial count, and what type of noise is generated in transit. Consider the following explicit formulation of amplified transmission in a branching chain:

A single pulse appears at the first stage of a branching chain. It is operated on according to the rule in Figure 2. At each subsequent stage the same rule is applied to every message event passing through. Some of these events (pulses) will disappear. Others will multiply. Depending on the balance among the conditional probabilities, progeny of the start-up event will tend to increase in number with each application of the Markov operator. Thus a probability distribution of message events is generated at the final (output) stage. The mean of this output distribution

Branching Chain Mechanism





Propagation Operator

Figure 2: Diagram of branching chain. A single impulse appears at the input (stage zero) moving through chain from left to right. Chain consists of r stages each roughly identified as a tiny group of interconnected neurons. At stage 1, initial input pulse is transformed by propagation operator, becoming zero, one, or two pulses. Probabilities of transformation shown in operator diagram below the chain. At stage 2, pulse(s) transformed again. In the event transform moves pulse to zero-state either here or at any other stage, transmission ceases. Pulse is lost. Same type of propagation at all stages.

Total count behaves like a random walk, going up when pulses are multiplied, down when lost; unchanged when a given stage is passed intact. At each stage, intruder or "immigrant" pulses may also appear. These originate outside the message sequence and reflect external sources of noise. Once incorporated into chain, however, immigrants are processed as though they were message-events, becoming indistinguishable from them. Immigration is discussed in section on Distribution Family.

Objective of analysis is to calculate output counting distributions for a variety of propagation conditions, and different transmission mechanisms.

measures amplification produced by the chain, while its variance measures internal noise.

If input pulses are transmitted with perfect fidelity, the Markov operator must have the following form:

$$p_0 = 0,$$

$$p_1 = 1,$$

$$p_2 = 0.$$

A chain following this rule is completely transparent to its input. Delivery of the start-up pulse intact at the terminus of the chain is guaranteed. No output probability distribution is found.

Generating Functions:

Branching chains are multiplicative. The number of message-events operating at any given point is a random variable tending to grow as information passes up the chain.

Beginning with a single start-up event, the process leads to an output probability distribution. At every point in the chain, message-events come and go. We use a generating function to keep track of them. This function multiplies in an interesting way as events move from stage to stage, and it decomposes in an even more interesting way when a chain is cut. We can then alter propagation mechanisms for information passing through such a cut.

The generating function of any discrete probability distribution *p(k)*, where *k* is an arbitrary integer (including zero) is given by:

$$G(s)=p(0)s^{o}+p(1)s+p(2)s^{2}+\cdots+p(k)s^{k}+\cdots. \tag{1}$$

The variable *s* is introduced as an analytical device. The coefficients of s^k are probabilities constructed typically by applying a simple multiplication rule as *k* changes. Choosing *s* suitably small causes *G(s)* to converge, often on a closed expression. If we now operate on *G(s)* and either expand or differentiate the result with respect to *s*, we have new probabilities reflecting the outcome of the operation. For example, when the operation involves summing *n* samples drawn from the same parent distribution, the generating function of the sum is given by:

$$(G(s))^{n}.$$

It is then easy to find the distribution of a sum via the generating function of one of its components if the latter are all independent with the same probability distribution. In the case of a branching chain, the result shows how to move from a chain with a single start-up event to one with *n* start-up events. We treat each separate start-up as an independent chain and then consider the generating function of the sum.

Unraveling A Branching Chain:

Our branching process is said to consist of *r* identical stages with the same propagation operator applied at each stage. Because of this continued recycling, we are free to enter the chain at any intermediate point, say stage *h*, and direct our attention to whatever happens from then on.

The full *r-stage* chain begins with a single start-up event and undergoes branching at each successive stage according to the rules in Figure 2. Entering at stage *h* we expect to find a counting distribution of events due to prior branching. This distribution becomes an input for the subsequent *r-h* stages of transmission. Any event passing through stage *h* is now a start-up event for this shorter process. We label counting probabilities at stage *h* as $p_h(k)$, where *k* is the size of the count. It ranges from zero up to a maximum where all message-events multiply at all prior stages. The generating function of the entire chain, *r* stages long, can then be reconstructed as follows:

$$G_r(s) = \sum_k p_h(k)(G_{r-h}(s))^k. \tag{2}$$

In Eq.(2) $G_r(s)$ is the generating function of the full *r-stage* transmission process, whereas $G_{r-h}(s)$ is our analogous expression for the shorter chain *r-h* stages long created by an arbitrary cut at stage *h*. Summation extends over the possible counts recorded at stage *h*. We see that $G_r(s)$ is the expectation of the chains resulting when the starting point is taken to be stage *h* rather than stage zero, and when the single start-up event is replaced by a distribution of events.

In view of Eq.(1) defining a generating function:

$$G_r(s)=G_h(G_{r-h}(s)), \tag{3}$$

where the argument of the generating function on the right hand side is now $G_{r-h}(s)$ rather than s. Evidently we can unravel $G_r(s)$ step-by-step using only these simple rules. For example, let h=1. Then:

$$\begin{aligned} G_r(s) &= G_1(G_{r-1}(s)), \\ &= p_0+p_1G_{r-1}(s)+p_2(G_{r-1}(s))^2. \end{aligned} \tag{4}$$

This is the relation between generating functions at successive stages. It shows the transmission operator in Figure 2, one step beyond start-up, working to produce the final generating function at stage r. Evidently transitions between successive stages are sharply constrained by the rules in Figure 2. As a convenience when the subscript is unity, we usually omit it altogether. Hence:

$$G_r(s)=G(G_{r-1}(s))=G_{r-1}(G(s)).$$

Now unravel $G_{r-1}(s)$ using the same reduction scheme. We have:

$$G_r(s)=G(G(G_{r-2}(s))). \tag{5}$$

Recycling the same propagation operator over and over produces a multiplicative system of generating functions corresponding to the growth of multiplicative noise inside the branching chain.

One final point on the properties of stage-by-stage message evolution depicted as a nested sequence of generating functions: let $h=r$. The partition in Eq.(3) then becomes:

$$G_r(s)=G_r(G_o(s)),$$

from which it follows that:

$$G_o(s)=s. \tag{6}$$

The zero-stage generating function shows that all probability is concentrated at a count of unity. In other words, the process begins with a single start-up event as we have been claiming.

If the cut is made at $h=0$, Eq.3 partitions $G_r(s)$ from stage zero, i.e., from the very front end of the branching chain:

$$G_r(s)=G_0(G_r(s)).$$

We see that the zero-stage generating function is an identity operator transforming the argument of the generating function into itself.

<u>Differential Equation of the Chain</u>:

The analysis in the previous section shows that cutting a branching chain at the first stage, one step beyond start-up, partitions the generating function of the full chain:

$$G_r(s) = G(G_{r-1}(s)).$$

If we then expand the right hand side as in Eq. (4), we find:

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

This linkage between successive generating functions is a reflection of the Markov transmission operator depicted earlier in Figure 2. Suppose further that each stage is Δt long and let:

$$\begin{aligned} t &= r\Delta t, \\ p_0 &= \mu\Delta t, \\ p_2 &= \lambda\Delta t. \end{aligned}$$

Then Eq. (4) becomes:

$$G_t(s) = \mu\Delta t + (1 - \mu\Delta t - \lambda\Delta t)\, G_{t-\Delta t}(s) + \lambda\Delta t (G_{t-\Delta t}(s))^2.$$

Evidently the generating function is converging on a differential equation if terms can be suitably defined:

$$\frac{G_t(s) - G_{t-\Delta t}(s)}{\Delta t} = \mu - (\mu + \lambda)\, G_{t-\Delta t}(s) + \lambda (G_{t-\Delta t}(s))^2. \tag{7}$$

We want the limit of this expression as r approaches infinity and Δt approaches zero while the product $r\Delta t$ remains finite. The gain and loss parameters, λ and μ, are taken to be fixed and finite. Hence the products $\lambda\Delta t$ and $\mu\Delta t$ must approach zero as Δt approaches zero. This means that as the process converges from discrete stages to a smooth flow in continuous time, the instantaneous probability of any change (up or down) becomes vanishingly small.

An overwhelming tendency exists for message-events to remain unaltered during the very brief time consumed by passage through any given transmission stage. Accordingly:

$$\lim_{\Delta t \to 0}\left(\frac{G_t(s) - G_{t-\Delta t}(s)}{\Delta t}\right) \Rightarrow G_t'(s) .$$

The expression on the right is the derivative of the generating function at a point close to the input end of the branching chain. Our parameter t measures the length of the chain since it corresponds precisely to stage r, the last step of the discrete transmission process.

Restrictions identical to Eq. (7) hold also for every intermediate point along the chain following start-up. Accordingly, t can be treated either as a fixed constant associated with the length of the chain, or as a time variable expressing the current state of the message. When both ideas are required in the same expression, we use t for the (fixed) length of the process, and τ for the time variable. Ordinarily t will designate the interval during which a counting process in a branching chain comes under observation.

In view of Eq. (7) and because the process behaves in the same way at each point along its entire length, we are led to a characteristic differential equation for the generating function governing the counting process at the output of the chain:

$$\frac{G_\tau'(s)}{(\mu/\lambda - G_\tau(s))(1-G_\tau(s))} = \lambda . \tag{8}$$

The generating function corresponding to an observation time t is found by integrating Eq.(8) between limits $G_\tau(s) = s$ at start-up and $G_\tau(s) = G_t(s)$ at the chain's output. Equation (8) shows the original quadratic expression in Eq. (7) factored now into its two roots.

One root is unity, an outcome mandated by the restriction that the coefficients of Eq. (7) sum to zero. This in turn implies an even more fundamental restriction, namely that message-events tend to remain unchanged during the very brief time of passage through a single stage. The second root, μ/λ, is an upper limit for $G_\tau(s)$ as τ approaches infinity. This root conveys important information on the asymptotic behavior of the process. More on that shortly.

Generating Function of the Process:

The solution of Eq. (8) can be found in most tables of integrals. Our notation produces:

$$G_t(s) = \frac{\mu/\lambda - \left(\frac{\mu/\lambda - s}{1-s}\right)e^{-(\lambda-\mu)t}}{1 - \left(\frac{\mu/\lambda - s}{1-s}\right)e^{-(\lambda-\mu)t}}. \qquad (9)$$

This is the generating function for a branching chain similar in character to Figure 2 but operating in continuous time rather than discrete stages. If a way can be found to expand Eq. (9) as a power series, the coefficients of successive powers of s will be probabilities of all possible output counts produced by the continuous process approximating our branching chain.

The development thus far parallels the generating-function analysis of a linear "birth-and-death process". The latter is a stochastic mechanism resembling an upward-drifting random walk on the positive integers (and zero). It is described in many easily accessible sources, notably Bharucha-Reid (1960, pp 9-56 and 86-89); and Feller (1957, pp 407-411).

The limiting scheme we have used, moving from discrete stages to continuous time, is a popular elementary argument employed, for example, in going from a (discrete) geometric distribution to its (continuous) exponential counterpart. It is, however, quite inelegant and ordinarily shunned by analysts. Elegant or not, we persist in viewing the linear birth-and-death process chiefly as a limiting form for multi-stage branching chains. Such chains appear to be nature's way of moving sensory data, but they are discrete and hard to analyze. A stochastic process is not only easy to study, it is also more informative on the types of noise created in transmission than discrete analyses tend to be.

Expansion of Eq.(9) in powers of *s* proves to be no great problem. First subtract unity from both sides. This leads directly to:

$$\frac{G_t(s)-1}{s-1} = \frac{1-\mu/\lambda}{(1-s)-(\mu/\lambda-s)\,e^{-(\lambda-\mu)t}},$$

$$\frac{G_t(s)-1}{s-1} = \frac{1-p_t(0)}{1-\left(\frac{1-e^{-(\lambda-\mu)t}}{1-\mu/\lambda e^{-(\lambda-\mu)t}}\right)s}\,. \tag{10}$$

The numerator on the right hand side of Eq. (10) is the probability of a non-zero count. We establish this point by calculating the exact probability of a zero count (i.e., silence at the output), using the generating function. The significance of zero-count is that our mechanism constructs a detached spike of probability at the zero-state. This is why the survival probability appears in the numerator of Eq. (10). Zero-count probability is found by setting s = 0 in Eq. (9):

$$p_t(o) = \frac{\mu/\lambda\,(1-e^{-(\lambda-\mu)t})}{1-\mu/\lambda e^{-(\lambda-\mu)t}}\,. \tag{11}$$

Additional simplification occurs if we also define:

$$\beta = \frac{(1-\mu/\lambda)\,e^{-(\lambda-\mu)t}}{1-\mu/\lambda e^{-(\lambda-\mu)t}},$$

(12)

and:

$$F_t(s) = \frac{\beta s}{1-(1-\beta)s}.$$

(13)

The adjusted generating function in Eq. (10) then reduces to:

$$\frac{G_t(s)-1}{s-1} = \frac{1-p_t(o)}{1-(1-\beta)s}.$$

(10a)

All elements of ß are fixed when chain parameters are established. Hence ß will be an important constant of any branching chain.

With these steps we have reconfigured the original generating function, $G_t(s)$, into greatly simplified form:

$$G_t(s) = p_t(o) + (1-p_t(o))F_t(s).$$

(9a)

Evidently $F_t(s)$ in Eqs.(9a) and (13) represents the survivors probability distribution, i.e., message-events destined never to jump to the zero-state. Its generating function tells us that this survivors distribution is geometric and shifted one step to the right of the zero-state. Further, the overall generating function, $G_t(s)$, is shown to be the compound of a binomial (extinction or survival), with a second generating function constructed only from message-events in the survivor category. The latter prove to have a geometric distribution because new message events are produced via a fixed multiplication rule. As the process runs, each additional survivor increases the output count by unity while its probability is multiplied against that of the prior count.

Branching Chain Counting Distribution:

With Eq. (9a) it is now easy to extract the counting distribution of the full branching chain based on the mechanism in Figure 2, and assuming a single start-up event. We invoke Eqs. (11), (13), and (9a) in order to expand $F_t(s)$ as a geometric series in powers of s. The probability of an output count consisting of exactly k message events is identified as $p_t(k)$. This probability turns out to be the coefficient of s^k in the expansion of Eq.(9a):

$$p_t(o) = \frac{\mu/\lambda(1-e^{-(\lambda-\mu)t})}{1-\mu/\lambda e^{-(\lambda-\mu)t}}, \qquad k=0,$$

$$p_t(k) = (1-p_t(0))\beta(1-\beta)^{k-1}, \qquad k=1,2,3,4,... \tag{14}$$

Equation (14) is our sought-after counting distribution at the output of a branching chain. Notice the detached spike of probability at the zero count.

In stochastic-process literature this distribution is called a linear birth-and-death process. Our analysis characterizes it as a noisy stochastic amplifier governing transmission of a single input event from start-up to destination via repeated identical stages. If stimulus magnitudes act on the chain's growth parameters instead of triggering independent inputs, the expected number of output events will reflect any desired amplification of the input pulse. Of course, this amplified output will also contain an added noise associated with repeated branching.

Accordingly, geometric branching has a substantially larger variance than its Poisson counterpart. This leads to Weber's-law increment detection as contrasted with square-root-law detectability in a Poisson point process.

Mean and Variance of Counting Distribution:

1)Mean: Average output produced by continuous branching in the counting distribution, Eq.(14), can be calculated in several different ways. Note that only the non-zero states contribute to the count. Hence:

$$mean=(1-p_t(0))\cdot\frac{1}{\beta},$$

where $1/\beta$ is the expectation of the shifted geometric distribution corresponding to $F_t(s)$. In view of Eqs.(11) and (12) we then have:

$$mean=e^{(\lambda-\mu)t}. \tag{15}$$

Branching causes the average output to grow exponentially with the difference between gain and loss parameters of the chain.

A second path to the same result uses our adjusted generating function, Eq.(10a). The limit of this expression as s approaches unity generates the average count. This property of Eq.(10a) is based on results obtained with so-called "tails generating functions". (See Feller, 1957, pp 249 - 250, and our own analysis a bit later.) Accordingly, set $s = 1$ in Eq.(10a) to find the mean count. It is the value just calculated.

2)<u>Variance:</u> The variance of Eq.(14) can be calculated directly or from its generating function. One of the best approaches uses our adjustment in Eq.(10a) based, as we show later, on the tails distribution of Eq.(14). Distributions with a detached spike of probability at zero are actually somewhat easier to analyze in tails form. In any event, let:

$$\frac{G_t(s)-1}{s-1}=Q_t(s)\ . \tag{16}$$

The variance is found (see below in this paper) from:

$$variance=2Q_t^{\prime}(1)+Q_t(1)-Q_t^2(1)\ , \tag{17}$$

where $Q_t(1)$ is the counting distribution's mean value as we have just calculated it, $Q_t'(1)$ is the first derivative of $Q_t(s)$ taken with respect to s, and s is set equal to unity. For the variance of Eq.(14) we have then:

$$\begin{aligned} variance &= \frac{2}{1-\mu/\lambda}(1-e^{-(\lambda-\mu)t})\,(e^{2(\lambda-\mu)t})+e^{(\lambda-\mu)t}-e^{2(\lambda-\mu)t}, \\ &= \left(\frac{2}{1-\mu/\lambda}-1\right).\,e^{(\lambda-\mu)t}(e^{(\lambda-\mu)t}-1), \\ &= \left(\frac{\lambda+\mu}{\lambda-\mu}\right).\,e^{(\lambda-\mu)t}(e^{(\lambda-\mu)t}-1)\,. \end{aligned} \tag{17a}$$

If the multiplier (gain) parameter λ is appreciably greater than the dropout (loss) parameter μ, the variance of the output will be proportional to the mean squared. As average output increases in a branching chain, counting variance grows more rapidly than its counterpart Poisson variance. The enhanced growth rate implies that a branching chain will exhibit Weber's-law detectability rather than square-root-law detectability. Just-discriminable increases in output should then be proportional to the background level itself, not to its square root. This variance growth is large. It is due chiefly to multiplier effects on individual message-events passing through the branching chain. Whenever any of these message-events happens to drop out or to increase in number, additional variance is created.

If the parameters of the chain approach each other in magnitude, the branching process becomes balanced, and the above arguments no longer apply. We then have a baseline condition resembling an adaptation level. The mean and variance in such circumstances are obtained by letting λ approach μ in Eqs. (15) and (17a). Hence, as $\lambda - \mu \to 0$ so that the squared difference is negligibly small:

$$\text{mean} \to 1, \qquad \text{variance} \to 2\lambda t. \tag{18}$$

The mean output count is just what we would expect from a symmetric Markov process. A single start-up event begins the transmission. Gain and loss parameters are perfectly balanced (see Fig. 2). A single message-event should then be expected at the output. The variance, on the other hand, is something of a surprise. It is large, increasing with both the multiplication parameter and the length of chain. When the output of a branching chain distribution (Eq. 14) is examined under these circumstances, we find a dense concentration of probability at the count of zero, together with a long, very low geometric tail governing the non-zero states. Hence, a balanced transmission operator will deliver a replica of its start-up pulse on the average, but the output displays substantial variance. Later, when we speak of a transmission network as "virtually silent" under continuous stimulation, it is this kind of silence we have in mind, a mostly silent but occasionally noisy background rather than utter transparency. We return to this baseline distribution shortly. It is one of the simplest and therefore most important examples of a branching chain.

Tails Generating Function:

In the previous section an adjusted generating function was employed to calculate the mean and variance of our branching-chain counting distribution, Eq.(14). The significance of the adjustment was passed over lightly, except for an observation that its basis is the tails distribution of Eq.(14).

We now show how a tails generating function produces the adjustment described in Eq.(10a). Our argument roughly parallels that given in Feller (1957, pp.248-250). The matter is taken up in detail here because of the peculiar character of the branching chain distribution with its isolated spike of probability at count zero, and its long geometric tail corresponding to non-zero counts. In such distributions the tails generating function has an unusually convenient form, making for easy recognition if it should turn up as the result of some analytic operation.

The classical probability generating function defined in Eq.(1) is only one of a variety of possibilities. Consider, for instance, the following closely related expansion:

$$Q_t(s)=q_t(0)+q_t(1)s+q_t(2)s^2+q_t(3)s^3+\cdots, \tag{19}$$

where:

$$q_t(k) = \sum_{i=k+1}^{i=\infty} p_t(i) \,. \tag{20}$$

The terms $p_t(i)$ in Eq.(20) are evidently counting probabilities given by Eq.(14), and $q_t(k)$ is defined as the tail beyond $p_t(k)$. We label $Q_t(s)$ a "tails generating function" because it is constructed from the tails associated with each counting probability $p_t(k)$.

Multiply both sides of Eq.(19) by s-1. We have:

$$(s-1)\cdot Q_t(s) = -q_t(0) + \sum_{k=1}^{k=\infty} p_t(k)\, s^k,$$

$$G_t(s) = (s-1)\cdot Q_t(s) + 1 \,. \tag{21}$$

Accordingly, the tails generating function, $Q_t(s)$, defined in Eq. (19) may be obtained directly from $G_t(s)$,the standard generating function, without actually forming and summing the tails series:

$$Q_t(s) = \frac{G_t(s) - 1}{s-1} \,.$$

This is exactly the form of our adjustment in Eq. (10a).

If both sides of Eq. (21) are differentiated with respect to s, we find:

$$G_t'(s) = (s-1)Q_t'(s) + Q_t(s)\,. \tag{22}$$

Set $s = 1$ and observe that $G_t'(1)$ is the mean of the branching chain distribution. Hence:

$$mean = Q_t(1)\,,$$

as claimed in the preceding section.

A second differentiation of Eq. (22) produces:

$$\begin{aligned} G_t''(s) &= \big(sQ_t''(s) + Q_t'(s)\big) - Q_t''(s) + Q_t'(s)\,, \\ &= (s-1)Q_t''(s) + 2Q_t'(s)\,. \end{aligned} \tag{23}$$

Now set s=1 in Eq. (23), producing the second factorial moment on the left hand side and $2Q_t'(1)$ on the right side. Thus:

$$2nd\ factorial\ moment = 2Q_t'(1)$$

This leads immediately to Eq. (17) for the variance of the branching chain distribution in view of the relation between factorial moments and central moments. (See, for example, Bharucha-Reid, 1960, p. 20, Equation 1.36.)

A tails generating function is evidently valuable since it offers an easy route to the moments of branching chain counting-distributions, but as we have already suggested, convenience is not our primary interest. Look back to Eq. (10a) and see that it has an unusually simple form compared with Eq. (9):

$$Q_t(s) = \frac{A}{1-Bs}, \tag{10b}$$

where now $A = 1-p_t(0)$, and $B = 1-\beta$. These two constants yield the zero count probability, and the key parameter of the survivors' distribution. If several different branching chains are constructed in sequence, and the tails generating function of the output proves to have the form prescribed in Eq. (10b), we know at once that a single equivalent chain must exist, and that its distribution parameters are given by A and B. We can then infer the rule by which components combine to produce the parameters of the single chain.

Auditory transmission is a branching system so complex that research has only begun to analyze its properties. An argument such as the one in the preceding paragraph permits us to consider the entire system as a single entity when the prescribed conditions exist. We are not forced to identify and analyze each separate component before a system can be pieced together. Branching chains connected in sequence, are equivalent to a single chain. This implies what we have always believed, that auditory psychophysics must play a significant role in determining system structure as well as in measuring what gets through. The description is vague at first, in the way that a single chain is vague, achieving greater specificity and detail as knowledge progresses.

Extinction Probability:

The roots of the quadratic underlying the birth/death process differential equation, Eq.(8), turn out to be unity and μ/λ respectively. The first root emerges because the coefficients of Eq.(7) sum to zero. Then if $G_t(s) = 1$, the derivative in Eq.(7) becomes zero, and this defines a root.

We now show that the second root, μ/λ, is an asymptote for the zero-count probability when the chain becomes infinite in extent. Notice that to form a derivative from Eq.(7) the number of stages, r, is allowed to approach infinity while the transit time per stage, Δt, goes to zero so that the product $r\Delta t$ remains fixed and finite. The asymptote under consideration here is the one that develops as t itself goes to infinity. In order to get to this asymptote, we require two separate limiting steps. In the first, the number of stages in the chain goes to infinity but the time t required to pass through them remains finite. In the second, t itself approaches infinity.

Now consider the limit of the generating function for a branching chain, $G_t(s)$ in Eq.(9), following the second step. Evidently, as $t \to \infty$ with λ greater than μ:

$$G_t(s) \Rightarrow \mu/\lambda.$$

The only probability left in the generating function is located at the zero count. Equations (11) and (12) reveal that as $t \to \infty$:

$$\begin{gathered}\beta \to 0,\\ p_t(0) \to \mu/\lambda.\end{gathered}$$

(24)

Zero-count probability grows with time, eventually reaching an asymptote defined by the intensity parameters of the transmission system. We label this asymptote:

$$\begin{aligned}\pi &= \lim\ p_t(0),\\ &= \mu/\lambda.\end{aligned}$$

(25)

The limit is approached as $t \to \infty$.

In stochastic processes π is called an *extinction probability*. Once message-events are lost, they cannot be recaptured. As the transmission process runs, zero-count probability increases with time (i.e., length of the chain) eventually reaching its maximum value π when $t \to \infty$. Simultaneously the survivors distribution, although still geometric in form, becomes vanishingly small (see β in Eq.(24)) over all finite counts. When $t \to \infty$ and the multiplier λ is greater than the dropout parameter μ, counts that are not zero become infinite, indicating that the mean and variance of the counting distribution, Eq.(14), must be infinite as well.

Certain applications of birth/death processes are centered on extinction calculations - - studies of species survival, for example. In our case no special role is played by extinction since neural transmission times are relatively short. Extinction probability is best thought of as a handy constant of a transmission system, serving to characterize its changing resistance to input data at different intensities. Later, when we attempt to relate system parameters λ and μ to measurements of stimulus intensity, the extinction constant will be quite useful.

What if the loss parameter should prevail, i.e., $\mu > \lambda$? Equation (9) tells us that as $t \to \infty$:

$$G_t(s) \Rightarrow 1,$$

from which it follows that extinction probability $\pi = 1$. Ultimate extinction is virtually certain; hardly surprising since the dropout parameter dominates transmission.

Finally, if parameters λ and μ are equal and the system is perfectly balanced, we must first consider the limit of Eq. (9) as $\mu \to \lambda$ before letting t go to infinity. The simplest way to do this is to turn to Eqs. (11) and (12). As $\mu \to \lambda$:

$$p_t(0) \to \frac{\lambda t}{1+\lambda t},$$
$$\beta \to \frac{1}{1+\lambda t}.$$

Using these limits it is now easy to construct a generating function and a probability distribution via Eq.(9a) for the special case in which $\mu = \lambda$. The extinction probability is evidently:

$$\pi = \lim_{t \to \infty} p_t(0) = 1.$$

Despite identical parameters and a balanced system equally likely to multiply and decay with no net gain or loss, ultimate extinction is still practically certain.

Gathering these results together, we have for the extinction properties of a branching chain:

$$\begin{aligned} \pi &= \mu/\lambda, \qquad \textit{for} \;\; \mu<\lambda, \\ &= \;\; 1, \qquad\qquad\quad \mu\geq\lambda. \end{aligned} \tag{27}$$

The same generating function, Eq. (9), holds whether transmission is headed for amplification or extinction.

Our experience here with λ and μ equal suggests that a family of distributions will emerge as varying restrictions are placed on transmission parameters. Pure multiplication occurs when $\mu \to 0$; pure erosion, whenever $\lambda \to 0$. Moreover, multiplication need not be twofold. We might have any number of progeny, even a distribution in which the number of offspring changes at random from instant to instant. These differences are important but not fundamental. The counting distributions are altered in detail but have a common ancestry. All work in essentially the same way, and their generating functions are close relatives.

Log Transform:

As message-events move through a branching chain, a rapidly increasing flood of activity develops, ending with the chain's output discharge. We know from Eq.(15) that the average size of this discharge grows exponentially with the difference between the chain's multiplication and dropout parameters. Powerful amplifier effects are an important aid for detecting weak auditory signals near ambient noise level, but when inputs are of even moderate intensity, exponential growth also carries the risk of overload and possible breakdown.

A transformation of the input signal can act as a counterbalance to runaway amplification in a branching chain. In this section we consider the effect of a particular log-transform relating stimulus intensity to the chain parameters λ and μ. It exactly counters the exponential growth of message-events created by the branching-chain parameters, yielding an average output proportional to input intensity. The form of the output counting distribution is unaffected by this transformation. Everything remains as it was, save for a smaller mean output count and greatly diminished higher moments. These are attractive linear properties.

The transform we consider is the following:

$$\ln (1+ax) = (\lambda_x - \mu_x)\, t, \tag{28}$$

where x denotes a particular stimulus intensity.

An expression similar to Eq.(28) was proposed by Rushton (1961) as a basis for visual coding. Working with the optic nerve of the horseshoe crab (Fuortes, 1959), Rushton found that Eq.(28) produced linear growth of neural responses over a wide range. He speculated that log transforms should improve the coding of other modalities as well. 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 discrimination with regard to time course, spatial location, and quality" (Rushton, 1961, p. 177).

A log transform operating on net growth of message-events, offers advantages at both ends of the intensity continuum. At the low end where detections occur, the scale is spread out. It is also foreshortened at the high end where overload threatens. Branching chains become virtually silent under the stimulus of a continuous background. Hence, increments are always detected as new signals. In effect such detection implies log scaling since there is only one process at all background levels.

In Eq. (28) subscripts appear on the multiplication and dropout parameters because they are now taken to be functions of stimulus intensity. The latter is treated as a variable labelled x. A constant a matches units of intensity to the parameters of transmission. As before, t measures the apparent length of the branching chain. Chain length is assumed constant as intensity changes. Hence, the full effect of stimuli will be directed at the pair of transmission parameters, or rather at the difference between them. A momentary auditory signal may increase λ_x, or depress μ_x , or may change both simultaneously.

Equation (28) indicates that the log can be negative when $\mu_x > \lambda_x$. The argument of the logarithm itself must remain nonnegative. This leaves a narrow range in which ax might go negative, but as a practical matter, we deal here exclusively with positive (or zero) intensities for which $x \geq 0$. No barrier exists prohibiting symmetrical treatment of increments and decrements in a steady background. Each poses essentially the same detection problem, but for the moment we are limited to increments. Later, when adaptation levels are developed, departures up and down can be handled easily within the same branching-chain framework.

When stimulus intensity is zero, multiplication and dropout parameters are equal. This creates a "baseline" noise about which we have spoken previously. We first consider our log transform for nonzero stimuli. Then by allowing intensity to go to zero, we look at the baseline noise.

Let:

$$\pi(x) = \frac{\mu_x}{\lambda_x}, \tag{29}$$

and substitute Eq. (28) into the generating function for a branching chain, Eq. (9). We then have:

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

(30)

Now define:

$$\gamma(x) = \frac{ax}{1-\pi(x)} .$$

(31)

Next, add and subtract unity from Eq. (30). It brings us to:

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

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

(32)

This is the tails generating function corresponding to the intensity-based branching chain produced from Eq. (30). We need only isolate the coefficient of s in the denominator, and divide out the term $1 + \gamma(x)$ in order to put the tails generating function into the form decreed by Eq. (10b). Accordingly:

$$Q_x(s) = \frac{(1+ax) \ / \ (1+\gamma(x))}{1 - \left(\frac{\gamma(x)}{1+\gamma(x)}\right)s} .$$

(32a)

Equation (32a) is now in proper tails format. Using Eqs. (10b), (11), and (12), we next write down the zero-count probability, as well as the parameter ß of the survivors distribution:

$$p_x(o) = \frac{\gamma(x) - ax}{1+\gamma(x)},$$
$$\beta = \frac{1}{1+\gamma(x)}.$$

(33)

As we have seen, these are the basic ingredients of every branching chain. Evidently the geometric form of the survivors' distribution is not altered by a log transform mapping intensities into λ and μ. Our output counting distribution, transformed by Eq. (28), is:

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

$$p_x(k) = (1-p_x(0))\beta(1-\beta)^{k-1}, \qquad k = 1,2,3,4\cdots.$$

(34)

The shape of the distribution is identical with Eq. (14). Only parameters change. The new mean and variance can be found easily from $Q_x(s)$ in Eq. (32).

As before, the mean is given by $Q_x(1)$:

$$\text{mean} = 1 + ax.$$

(35)

Similarly, the variance is constructed from $Q_x(s)$ and its first derivative in s, $Q'_x(s)$:

$$\begin{aligned} variance &= 2Q_x'(1) + Q_x(1) - Q_x^2(1)\,, \\ &= 2\gamma(x)\,(1+ax) + (1+ax) - (1+ax)^2, \\ &= \left(\frac{1+\pi(x)}{1-\pi(x)}\right)(ax)\,(1+ax)\,. \end{aligned}$$

(36)

When stimulus intensity *x* is substantially above zero, variance is again proportional to the mean squared; but now the mean is linear in intensity rather than growing exponentially.

Suppose intensity is zero. We then confront the baseline noise condition of this branching-chain transmission mechanism. In that event $\mu_x = \lambda_x$, and $\pi(x) = 1$. Adjusted intensity $\gamma(x)$, defined in Eq. (31), approaches a limit because both numerator and denominator go toward zero simultaneously:

$$\lim_{x \to o} \gamma(x) = \frac{ax}{1-\pi(x)} = \gamma\,.$$

(37)

The limit here is taken to be a background intensity establishing the rate at which $\pi(x)$ approaches unity as *ax* decreases. This implies that our representation of a steady background signal will be absorbed into the constant γ. Intensities should then be interpreted as increments above γ rather than zero. The system we are uncovering describes incremental changes in steady backgrounds instead of absolute signal levels referenced to zero.

To develop the baseline counting distribution, insert Eq. (37) into Eq. (32a). Then as x → 0:

$$Q_o(s) = \frac{1/1+\gamma}{1-(\frac{\gamma}{1+\gamma})s},$$

(38)

and in view of Eq. (10b):

$$p_o(o) = \frac{\gamma}{1+\gamma},$$
$$\beta = \frac{1}{1+\gamma}.$$

(39)

These expressions were worked out earlier for the general case. They are cast now into their simplest form: a perfectly balanced branching chain grown almost silent at its adaptation level. The "almost" is reflected in a long, low tail across non-zero counts.

The output distribution is produced from the constants in Eq. (39):

$$p_o(o) = \frac{\gamma}{1+\gamma}, \qquad k=o,$$

$$p_o(k) = \frac{\gamma^{k-1}}{(1+\gamma)^{k+1}}, \qquad k=1,2,3,4\cdots.$$

(40)

Mean and variance are easily obtained from $Q_o(s)$ with s = 1:

mean = 1,

variance = 2γ.

(41)

So the branching chain is not completely silent when no increment is applied -- just virtually silent. There is variance. In fact, the variance is found to be proportional to the level of the background signal. It is easy to see how Weber's law would emerge from such a mechanism. Its resolving power declines systematically as background intensity γ increases. Any incremental change must be large at high intensities in order to achieve discriminability equivalent to much smaller changes near threshold level. This is the heart of Weber's law.

Our discussion sets the stage for the obvious next step in simplification. It is easy to conjecture an "ultralinear" branching chain in which:

$$\gamma(x) = \gamma + ax. \tag{42}$$

Internal representations of background intensity and incremental intensities are assumed perfectly additive. In view of Eq. (31), this means that the extinction constant can now be defined explicitly in terms of stimulus measurements:

$$\pi(x) = \frac{\gamma}{\gamma + ax}. \tag{43}$$

Apart from the operational simplicity of an ultralinear branching chain, it possesses the great virtue that its parameters are given by intensities used to stimulate the system rather than by unknown internal constants. Our constant γ becomes an adaptation level while the term *ax* denotes an increment above the adaptation level. A detection theory cast in this form can easily be wrong, whereas in the format established by Eq. (9), there is always a sense in which the theory remains unchallengeable. It can never be quite pinned down.

3. DISTRIBUTION FAMILY

The Markov process depicted in Figure 2 is an extremely simple mechanism for amplifying weak auditory messages. Forces acting on pulse-like message-events as they pass along a transmission line are deemed to act in exactly the same way on all message-events. This includes the original start-up pulse and all its progeny, right up to the end of the line. Stable operators of this type cause neural messages to evolve in very predictable ways during transmission, creating a noisy amplified version of the input at the output site.

Our argument has been that such transmission phenomena should also affect the discriminability of output messages. In fact the changing detectability of intensity increments above a steady background signal may depend more on the changing statistics of transmission than on the characteristic responses of auditory receptors. We do not know this to be the case, nor do we assume it. We simply deal with it as an important conjecture linking internal noise with the transmission process. Something is obviously missing from the traditional stimulus-based detection theory. Transmission may prove to be the missing link.

As a family of related counting distributions is identified, we seek also to preserve the simplicity encountered thusfar. Our object is to illustrate the scope and power of very elementary ideas about sensory transmission, rather than strive for the most general formulation. This means that genuinely abstract family members will be omitted from our compilation. We concentrate on elementary forms as we develop our picture of the branching-chain family in Table 1.

Most of the results in this section can be obtained by playing changes on the parameters λ and μ in Eq.(9). This is in fact how we operated in our analysis of the log transform just presented. Here, however, we return repeatedly to the basic differential equation in order to display each different mechanism as an independent solution worthy in its own right.

Pure Erosion:

Consider a transmission chain unable to multiply. Its propagation diagram is illustrated in Figure 3a.

No progeny are created other than a reproduction of the initial start-up pulse. At each successive stage, this single pulse is either passed on to the next stage, or it is lost, terminating transmission. Evidently there can be only these two output conditions: a count of zero when the start-up event is lost, or a count of unity indicating successful passage through the chain. The longer the latter, the greater the likelihood that a given input fails to survive.

\- FIGURE 3 HERE -

To analyze output in such circumstances, start with our basic differential equation in Eq. (7) and set λ equal to zero. This step eliminates any possibility of amplification. The chain is now purely erosive. A given input is represented by at most one pulse somewhere in the chain, and if that pulse should be lost, transmission ends. Our process differential equation modified from Eq. (7), sets up as follows:

$$G'_\tau(s) = \mu(1-G_\tau(s)). \tag{44}$$

We let:

$$\begin{aligned} &y=G_\tau(s), \\ &dy/d\tau=G'_\tau(s). \end{aligned} \tag{45}$$

Integrating the differential equation produces:

$$\int_s^{G_t(s)} \frac{dy}{1-y} = \mu\int_o^t d\tau,$$

$$\ln(1-s)-\ln(1-G_t(s))=\mu t,$$

$$\ln\left(\frac{1-G_t(s)}{1-s}\right)=-\mu t,$$

$$G_t(s)=(1-e^{-\mu t})+e^{-\mu t}s. \tag{46}$$

The generating function governing erosive transmission, Eq. (46), has only two terms in its expansion. Coefficients of s^0 and s correspond to the respective probabilities of the two possible outputs (zero and unity). As the chain increases in length (measured by t), survival probability (coefficient of s) drops off exponentially, approaching zero when t grows large.

Suppose there are many start-ups instead of the single chain depicted in Eq.(44). The summed output produced by multiple chains can be obtained easily via Eq.(46). With n inputs, the generating function of the sum will be given by :

$$G_t^n(s)= (1+e^{-\mu t}(s-1))^n. \tag{47}$$

Equation (47) signals a binomial distribution. If the chain is sufficiently long, survival probability will be small. With n suitably large, a limit can be constructed in which the product:

$$\theta=ne^{-\mu t}$$

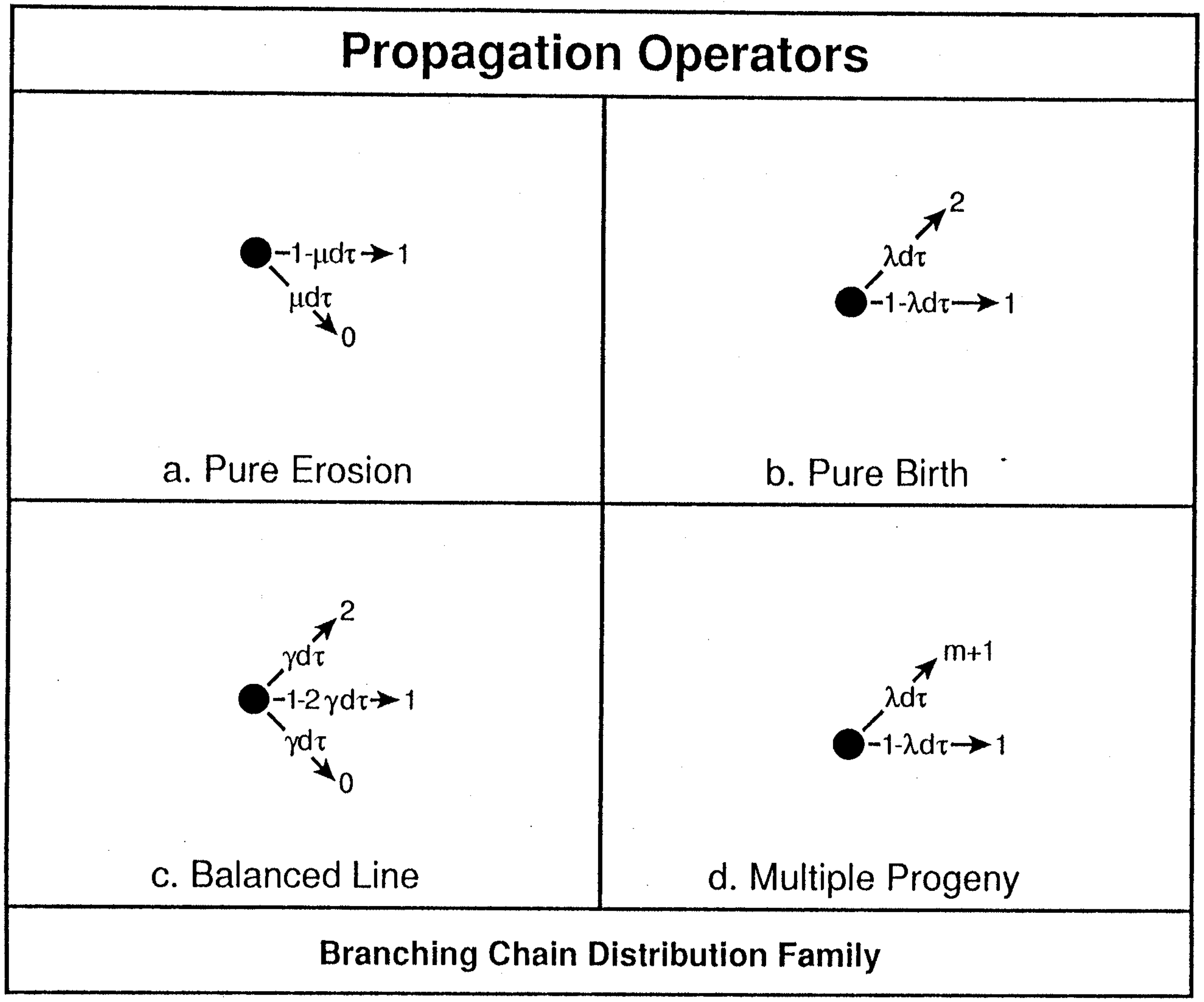


Figure 3: Propagation operators reflecting typical restrictions on branching-chain transmission. All four operators belong to same distribution family.

remains finite as its constituent terms go toward infinity and zero respectively. Then:

$$G_t^n(s) \Rightarrow e^{\theta(s-1)} . \tag{48}$$

When output is summed across many branching chains, the limiting distribution is Poisson. Hence, Eq. (48) is a neat little example of Cox & Smith's (1954) superposition theorem, cited earlier.

Hecht, Schlaer, and Pirenne (1942) employed a very similar transmission scheme to account for photon losses as light flashes make their way through the cornea, the ocular media, and surface obstructions on the retina, to reach rod-receptors. This famous paper draws attention to the Poisson character of visual detection at absolute threshold. Threshold flashes entering a dark-adapted eye are found to contain about 100 photons. Psychometric functions measured under the same conditions suggest a count of 5-10 events for determining whether or not a flash will be seen. Hecht et al. estimate actual losses in the passage of light energy through the eye, finding that 90-95% of the incident light is lost. They conclude that the visual psychometric function probably displays statistics of the light stimulus rather than the variable sensitivity of visual receptors. The latter are viewed essentially as a noise-free detection system. (See Sackitt, 1972.)

The argument centers on receptors but seems also to assume that the entire visual detection chain is noise-free. Any number of psychometric functions can appear identical if signals happen to be contaminated by an unmeasured internal noise. This point was demonstrated by Barlow (1956). So a count of 5-10 events cannot be established unless we also know the internal noise level. Moreover, we have shown that a passive transmission system, in conformity with the Cox-Smith theorem, produces a Poisson-like process at its output. Output may thus be shaped as much by the statistics of transmission, as by variations in input.

Since Poisson fluctuations arise so easily in transmission, it becomes virtually impossible to determine the actual source of threshold phenomena. Sensory analysis abounds in such paradoxes. Experience teaches great caution in reasoning from the shape of psychometric functions to the structure of sensory processes.

- TABLE 1 HERE -

Pure Birth:

This member of the branching-chain family provides an ideal amplifier for weak input signals. Transmission in such cases is protected from loss or dropout. A single start-up pulse builds rapidly into a cascade of message-events. The propagation operator for achieving this buildup is depicted in Figure 3b.

Initial stages of auditory detection probably involve some form of feedback amplification. The extraordinary sensitivity of the auditory system suggests that it must have a good way to dig very weak signals out of the ambient noise. Amplifier mechanisms should

TABLE 1

BRANCHING CHAIN DISTRIBUTION FAMILY

Description	*Generating Function*	*Tails Function*	*Mean*	*Variance*	*Zero-State Probability*	*Restrictions*
Noiseless Transmission	s	1	1	0	0	$\lambda=0, \mu=0$
Pure Erosion	$1 + e^{-\mu t}(s-1)$	$e^{-\mu t}$	$e^{-\mu t}$	$(e^{-\mu t})(1-e^{-\mu t})$	$1-e^{-\mu t}$	$\lambda=0$
Pure Birth	$\frac{e^{-\lambda t}s}{1-(1-e^{-\lambda t})s}$	$\frac{1}{1-(1-e^{-\lambda t})s}$	$e^{\lambda t}$	$e^{\lambda t}(e^{\lambda t}-1)$	0	$\mu=0$
Balanced-Line	$1 + \frac{s-1}{1-\gamma(s-1)}$	$\frac{1}{1-\gamma(s-1)}$	1	2γ	$\frac{\gamma}{1+\gamma}$	γ is noise factor of a balanced-line
Multiple Progeny	$\frac{e^{-\lambda t}s}{(1-(1-e^{-m\lambda t})s^m)^{\frac{1}{m}}}$	*	$e^{m\lambda t}$	$me^{m\lambda t}(e^{m\lambda t}-1)$	0	$\mu=0$ $m=1,2,3,4,\cdots$
Branching Chain	$\frac{\mu/\lambda-\left(\frac{\mu/\lambda-s}{1-s}\right)e^{-(\lambda-\mu)t}}{1-\left(\frac{\mu/\lambda-s}{1-s}\right)e^{-(\lambda-\mu)t}}$	$\frac{1-p_t(0)}{1-(1-\beta)s}$	$e^{(\lambda-\mu)t}$	$\left(\frac{\lambda+\mu}{\lambda-\mu}\right)e^{(\lambda-\mu)t}(e^{(\lambda-\mu)t}-1)$	$\frac{(\mu/\lambda)(1-e^{-(\lambda-\mu)t})}{1-(\mu/\lambda)e^{-(\lambda-\mu)t}}$	$m = 1$
Log Transform	$1+\frac{(1+ax)(s-1)}{1-\gamma(x)(s-1)}$	$\frac{1+ax}{1-\gamma(x)(s-1)}$	$1 + ax$	$\left(\frac{1+\pi(x)}{1-\pi(x)}\right)(ax)(1+ax)$	$\frac{\pi(x)\gamma(x)}{1+\gamma(x)}$	$\ln(1+ax)=(\lambda_x-\mu_x)t$ $m = 1$
Immigration into Balanced-Line	$\left(1+\frac{s-1}{1-\gamma(s-1)}\right)\cdot\left(\frac{1}{(1-\gamma(s-1))}\right)^{\nu t/\gamma}$	*	$1 + \nu t$	$2\gamma+\nu t+\gamma\nu t$	$\left(\frac{\gamma}{1+\gamma}\right)\cdot\left(\frac{1}{1+\gamma}\right)^{\nu t/\gamma}$	ν is immigration rate γ is noise factor of a balanced-line $m = 1$

* No Simple Expression

be located close to the receptor organ, and would probably deliver their output into a conventional transmission network conveying message-events to a decision center at much higher level. Our ideal for this first amplifier stage of detection would be a pure birth process. The problem of analyzing two different types of transmission chain arrayed in sequence will be considered shortly.

In order to construct a chain with pure birth properties, we start with our basic differential equation, Eq.(7) but now set $\mu = 0$. This step eliminates all possibility of loss. A chain governed by such propagation will retain every message-event it creates, and the number will grow with successive stages. Our stochastic process restriction then becomes:

$$G_\tau'(s) = -\lambda G_\tau(s)(1-G_\tau(s)). \tag{49}$$

Again we invoke the simplified notation of Eq.(45) to integrate the differential equation:

$$\int_s^{G_t(s)} \frac{dy}{y(y-1)} = \lambda \int_0^t d\tau. \tag{50}$$

An easy solution suggests itself if we note:

$$\frac{1}{y(y-1)} = \frac{1}{y-1} - \frac{1}{y},$$

$$\int_{s}^{G_t(s)} \frac{dy}{y-1} - \int_{s}^{G_t(s)} \frac{dy}{y} = \lambda t,$$

$$\ln\left(\frac{G_t(s)-1}{s-1}\right) - \ln\left(\frac{G_t(s)}{s}\right) = \lambda t,$$

$$\frac{G_t(s)-1}{G_t(s)} = e^{\lambda t}\left(\frac{s-1}{s}\right).$$

(51)

In this form, Eq. (51) can be solved for $G_t(s)$ as a function of s. We find:

$$\begin{aligned} G_t(s) &= \frac{se^{-\lambda t}}{1-(1-e^{-\lambda t})s}, \\ &= \sum_{k=o}^{k=\infty} e^{-\lambda t}(1-e^{-\lambda t})^k s^{k+1}. \end{aligned}$$

(52)

Evidently the pure birth output distribution is geometric. It is also shifted one step to the right so that it spans only the non-zero counts. The process begins with a single start-up event. Since μ is set equal to zero, no losses are encountered in transmission. Zero counts are excluded by these requirements.

In physics, Eq. (52) is known as a Yule-Furry process. It is an idealized model for electron-photon cascades. (See Bharucha-Reid, 1960, p. 247.) Sensory analysis uses birth processes such as Eq.(52) to model the early stages of transmission where focus is on detecting a signal in ambient noise, or extracting it from a continuing background.

Prior to any neural activity, we imagine that auditory receptors, acting as transducers, set up excitation in the form of stimulating events proportional to the energy of an auditory signal. The first such event starts up the neural amplifier while the others control its amplification parameter. With a local amplifier situated immediately behind the receptor, even weak transients can produce an impulse cascade. These impulses are thought to function as input for a second branching chain leading away from the receptor area. Additional boosting is probably required along this second chain in order to bring stimulus information securely through to a higher center where it can be evaluated. All these amplifier and boosting operations should generate a tell-tale internal noise. The latter will then affect the discriminability of auditory signals differing slightly in intensity.

Balanced Transmission Line:

This member of our branching-chain family has $\lambda = \mu$ in Eq.(7). The expected number of progeny for each message-event passing through any given stage is then unity. No net gain or loss of message-events occurs between stages. But new variance is created continuously as transmission branches up and down, altering the total count from stage to stage. This point is evident from the propagation diagram in Figure 3c. The longer the chain (measured by t), the greater will be the output variance.

We derive our output distribution by modifying Eq.(8) to show λ and μ equal:

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

(53)

Equation (53) can be integrated directly:

$$\int_s^{G_t(s)}\frac{dy}{(1-y)^2}=\lambda\int_o^t d\tau,$$

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

The generating function for propagation based on equal gain and loss parameters, must be:

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

(54)

Equation (54) can be expanded as a power series in s. Output counting probabilities governing perfectly balanced transmission are then given as the coefficients of successive powers of s. This particular generating function happens also to be set up ideally for transformation into tails form:

$$Q_t(s) = \frac{1}{1-\lambda t(s-1)},$$
$$= \frac{1/1+\lambda t}{1-\left(\frac{\lambda t}{1+\lambda t}\right)s}\,.$$

(55)

We see a large chunk of probability in the zero state. Following Eq. (10a), the probability is shown to be:

$$p_t(o) = \frac{\lambda t}{1+\lambda t}\,.$$

(56)

We also see a low geometric tail extending over all non-zero states. As expected in view of perfectly balanced propagation parameters, the mean output count, given by $Q_t(1)$ in Eq. (55), is unity. The input event reproduces itself on the average at successive stages. This restriction carries straight through to the chain's output.

Except for notation, the tails generating function in Eq. (55) is the same as the one developed earlier, characterizing our so-called "baseline counting distribution", Eq.(38). The latter emerged from analysis of a log transform when incremental intensity *ax* is set equal to zero. Hence, the output distribution generated by Eq. (54) must be identical to one given previously as Eq. (40).

Multiple Progeny:

Now consider a birth process creating more than one new event each time it multiplies. We argued earlier that there is no special magic about pair production. Pairs are easy to visualize and convenient to handle, but they are not logically necessary in a "jump" process. (See Rosenblatt, 1962, pp.124-133.)

Countless conflicting influences affect the momentary states of sensory transmission. These influences create randomized data-flows that ordinarily defy simple explanation. If multiplication occurs repeatedly in order to boost transmission (as we believe it does), no rule demands that the boost be limited to one new message-event, or that the number created be constant. Perhaps the propagation operator changes configuration as excitability shifts along the network. Progeny might then have a characteristic probability distribution generated by changes in the operator.

We should certainly consider what happens when the pair-production restriction is relaxed. Consequently, turn now to a stochastic amplifier that *really* amplifies, one creating exactly m new message-events with each new multiplication. A single event enters a given stage, and at random either that same event or $m+1$ events emerge from it. Each of the latter becomes a vehicle for further multiplication with the same properties. Buildup is rapid and, if the multiplication factor m is sizable, buildup is also enormous.

The propagation diagram for a birth process with *m+1* progeny at each multiplication is illustrated in Figure *3d*. Classical pure-birth (Yule-Furry), presented earlier in this section, was restricted to pair production. Now that restriction is relaxed, and the Yule-Furry process becomes a case in which $m = 1$.

We begin by translating Figure *3d* into an instant-by-instant restriction on the time-dependent generating function:

$$G_\tau(s) = (1-\lambda\Delta\tau)\, G_{\tau-\Delta\tau}(s) + \lambda\Delta\tau\ G^{m+1}_{\tau-\Delta\tau}(s)\,,$$

$$G'_\tau(s) = \lambda G_\tau(s)\ \left(G_\tau^{m}(s) - 1\right).$$

(57)

A single message-event enters a given stage. Equation (57) tells us that with probability $(1 - \lambda\Delta t)$ it emerges as the same single event; or with probability $\lambda\Delta t$ it emerges in the company of m new events. These are the only allowable outcomes as message-events transit a given stage. The dropout parameter μ is assumed to be zero. Nothing can ever be lost.

If multiplication occurs, the generating function is raised to the power m+1, reflecting the summed effect of new chains created when a message-event multiplies. Evidently when $m = 1$, Eq.(57) above reduces to Eq.(49), our earlier expression for the stochastic restriction governing a Yule-Furry process.

Set up the differential equation as before:

$$\int_{s}^{G_t(s)} \frac{dy}{y(y^m-1)} = \lambda \int_{o}^{t} d\tau \tag{58}$$

Recall that the chain of events begins with a single input at time zero. Hence:

$$G_o(s) = s.$$

This furnishes the lower limit of integration for the left-hand side of Eq.(58). We can handle the integration itself easily via a device, i.e., Eq.(51), used earlier with the differential equation underlying pair production:

$$\frac{1}{y(y^m-1)} = \frac{y^{m-1}}{y^m-1} - \frac{1}{y},$$

$$\int_{s}^{G_t(s)} \frac{1}{m} \cdot \frac{m\, y^{m-1}}{(y^m-1)} dy - \int_{s}^{G_t(s)} \frac{dy}{y} = \lambda t,$$

$$\frac{1}{m} \cdot \ln\left(\frac{G_t^m(s)-1}{s^m-1}\right) - \ln\left(\frac{G_t(s)}{s}\right) = \lambda t,$$

$$\left(\frac{G_t^m(s)-1}{G_t^m(s)}\right) = e^{m\lambda t} \cdot \left(\frac{s^m-1}{s^m}\right). \tag{59}$$

A few additional steps with Eq. (59) bring us to the generating function of this birth process, but raised to the *mth* power. In other words we are looking at the sum of *m* distinct chains:

$$G_t^m(s) = \frac{e^{-m\lambda t} s^m}{1-(1-e^{-m\lambda t}) s^m}. \tag{60}$$

When m = 1, pair production governs the multiplication depicted in Figure 3d. Hence, Eq. (60) with $m = 1$ must be a classical pure-birth process (and we see that it is). Similarly, if *m* is an integer greater than unity, Eq. (60) remains geometric but data points are spread out over spaced intervals: m, 2m, 3m, ... Counts lying between these landmarks all have probability zero because they cannot occur given the multiplier.

Finally, the sought-after generating function corresponding to a single chain with multiple progeny is found by taking the *mth* root of Eq. (60):

$$G_t(s) = \frac{se^{-\lambda t}}{(1-(1-e^{-m\lambda t}) s^m)^{1/m}},$$

$$= \sum_{k=o}^{k=\infty} \binom{\frac{1}{m}+k-1}{k} e^{-\lambda t}(1-e^{-m\lambda t})^k s^{mk+1}.$$

$$k = 0,1,2,3,....$$

$$m = 1,2,3,4,.... \tag{61}$$

Coefficients of s^{mk+1} in the expansion of Eq.(61) are probabilities of counts that actually appear in the output. Hence, the coefficients are successive values of $p_t(mk+1)$. All other counts must have probability zero in virtue of the multiplier.

Evidently Eq. (61) is a negative binomial distribution, constructed over outputs at 1, m+1, 2m+1, 3m+1, and so on. Unit count occurs when an input manages to pass through the entire chain without multiplying. As the chain lengthens, this "no-multiplication" event becomes less and less likely, implying that $p_t(1)$ tends toward zero as t grows large.

Fractional binomial coefficients such as Eq. (61) are not as simple as they appear. However, if the multiplier m happens to be small, the expansion of any particular coefficient can be worked out, and several of the most useful are tabulated. For example, when $m = 2$ (three progeny at each multiplication), Jolley (1961, #166) shows how to expand the generating function:

$$p_t(1) = e^{-\lambda t},$$

$$p_t(2k+1) = \frac{1.3.5.7....(2k-1)}{2.4.6.8......2k} e^{-\lambda t}(1-e^{-2\lambda t})^k, \qquad k=1,2,....\infty. \tag{62}$$

Because Eq. (60) is a spaced out version of the geometric distribution, there is also an easy way to calculate the principal moments of a multiple-progeny birth process. Results are given in Table 1.

A remarkable thing happens when the multiplier becomes very large. Equation (61) tells us that a limit exists:

$$\lim_{m \to \infty} G_t(s) = s e^{-\lambda t}. \tag{63}$$

Only the unit state (the "no-multiplication" event) remains. Nearest next count is infinitely removed. This implies that the mean and variance of the counting distribution must both be infinite.

We have an outcome very similar to extinction in birth-death processes. The probability distribution in Eq. (63) fails to sum to unity. Its higher moments are infinite. Significant portions of the process lie beyond the reach of our analytical framework. In a literal sense, if the chain ever multiplies, it explodes.

<u>Immigration:</u>

Distribution family-members cited thus far produce transmission noise only by repeated branching. Single message-events passing along a channel are continuously transformed by one of the branching operators in Table 1. Input events will then generate an output distribution whose variance embodies the noise created in a branching chain. The process leaves a variety of clues to its structure because limits are imposed on our ability to detect small intensity differences. Moreover, the limits change in different test situations.

Classical branching theory also provides for a second type of internal noise. Suppose an outside event, a random pulse with no prior history in the channel, suddenly intrudes and mixes with message events owing their origin to an input signal. Intrusions of this type are illustrated in Fig. 2. The intruders, or immigrants as they are called, constitute a secondary noise process indistinguishable from message-noise, but deriving from non-message sources. Once in place, immigrants become subject to the rules governing passage of information in a branching chain. Mechanisms incorporating intruders are known as birth-death-immigration processes. The early work appears to have been done by Kendall (1948). Theory is discussed in Bharucha-Reid (1960, pp. 173-74), Bartlett (1955, p.112), and Diament and Teich (1992). There are many other sources.

To formulate the intrusion phenomenon, begin with our rudimentary differential equation, Eq. (7). A branching message-process triggered by a single input event will continue to be identified with the generating function, $G_t(s)$. Now, however, we introduce a parallel and closely related process based on random immigration. This secondary noise process will be designated as $I_t(s)$. We can reconstruct the instantaneous behavior of $I_t(s)$ as follows:

$$I_\tau(s) = (1-\nu\Delta\tau)\ I_{\tau-\Delta\tau}(s) + \nu\Delta\tau\ (I_{\tau-\Delta\tau}(s) \cdot G_{t-\Delta\tau}(s)). \tag{64}$$

Immigration rate is given by ν, and the product $\nu\Delta\tau$ is the probability that a single immigrant will appear during $\Delta\tau$. Accordingly, Eq. (64) describes an instant-by-instant restriction on $I_\tau(s)$. At $\Delta\tau$ prior to τ, the process was either at $I_{\tau-\Delta\tau}(s)$ and no immigrant appeared, or a single immigrant arrived adding a new chain of events governed by the branching rules of the message-event process. The product of generating functions on the extreme right of Eq. (64) represents this new sum. We then have:

$$\frac{I_\tau(s) - I_{\tau-\Delta\tau}(s)}{\Delta\tau} = -\nu I_{\tau-\Delta\tau}(s)(1-G_{\tau-\Delta\tau}(s)).$$

Going to the limit as $\Delta\tau\rightarrow0$, yields a differential equation:

$$I_\tau'(s) = -\nu I_\tau(s)(1-G_\tau(s)). \tag{65}$$

An output generating function for $I_t(s)$ is obtained by solving Eq. (65) in the region between $\tau = 0$ and $\tau = t$, with $G_\tau(s)$ given either by Eq. (9) or one of the simpler versions in Table 1. A precise form of the general solution linked to Eq.(9) is available in several sources: notably Bharucha-Reid (1963, pp. 173-174), and Diament and Teich (1992).

The general solution of Eq. (65) is fairly complicated, involving rate parameters μ, λ, and ν, in addition to a time parameter t.

Table 1 offers a modest version of the solution for immigration into a balanced-line ($\lambda = \mu$). Here, we go the full distance toward simplification by developing only the immigration noise. Let:

$$G_\tau(s) = s.$$

This means that the message portion of the branching chain is taken to be noiseless. Any start-up or immigration event is delivered to the output without alteration.

Following Eq.(65) we then have:

$$\frac{dy}{y} = -\nu(1-s)\,d\tau,$$

where $y = I_\tau(s)$. Recall that $I_\tau(s)$ depicts a secondary noise source. When $t = 0$, all immigration noise must be concentrated at zero. Hence, $I_0(s) = 1$. Integrating both sides over values of τ between zero and t, we find:

$$\ln\left(\frac{I_t(s)}{1}\right) = -(1-s)\nu t,$$

$$I_t(s) = e^{\nu t(s-1)},$$

$$I_t(s) = \sum_{k=0}^{\infty} \frac{e^{-\nu t}(\nu t)^k}{k!}\cdot s^k. \tag{66}$$

Our result in Eq.(66) signals a Poisson distribution caused by random intrusions of immigrant pulses as the message-process runs its course from input to output. It follows that the product $G_t(s) \cdot I_t(s)$ describes the mixture of message and immigration events found in the output of branching chains subject to such secondary noise. In this special case, message noise is absent altogether. Accordingly:

$$\text{Output} = s \cdot I_t(s),$$
$$\text{Mean Output} = 1 + \nu t. \tag{67}$$

Shifting the Poisson one step to the right increases mean output by unity. A component νt of the mean output is due to immigration noise.

When machinery of this type is operating, immigrations occur with constant probability as messages pass through a branching chain. Evidently these extra pulses contribute a source of Poisson noise to the internal noise generated by the message-transmission process itself. But random intrusions do more than simply add a Poisson component to the output. Immigration noise introduced at any stage of the chain is subject to continuing modification by the mechanisms governing passage of information through the chain.

Thus, random intrusions will be amplified or lost right along with message-events. These interactions are typical of multiplication noise. They define one of the key differences between additive and multiplicative noise.

4. SEQUENCES OF BRANCHING CHAINS

If transmission is portrayed correctly as a stochastic process, an internal noise will accompany all output messages, making them less discriminable than signals entering the network. Limits on detectability might trace to this internal noise, especially if it is linked to stimulus intensity. The importance of these linkages was overlooked in early work on signal detection. Internal noise was treated as fixed and small, a minor contaminant rather than an integral part of the mechanism. We think it unwise to rule on such issues until more is known about the inner workings of the auditory system, particularly its handling of intensity information.

In this section we analyze transmission noise when stochastic chains performing different tasks are connected in sequence. For example, suppose receptor activity is processed by a neural network just behind the ear. The network acts as an amplifier, boosting weak signals by converting them into a pure birth process. A cascade of impulses results. This initial output consists partly of amplified signal and partly of noise created by the birth process. No way is known for teasing the components apart. Hence, each impulse is simply forwarded to a high-level decision center by a second network. The output of the latter now contains at least two distinct noise sources: 1) the noisy portion of the birth process, and 2) noise generated in transmission. Assume this transmission noise to come from a perfectly balanced branching-chain. Immigration noise might be another possibility, but we neglect it here.

We know how to formulate a birth process and a balanced-line separately, but how do we deal with their combined output when they are arranged in sequence?

The problem is important because a two-step detection mechanism is more realistic than a simple branching chain whose propagation parameters would be the same everywhere. In this section we show that such an equivalent branching chain exists, and can be substituted for a two-step amplifier-transmitter. Properties of the equivalent chain are not hard to calculate from the parameters of its two components.

This principle of equivalence implies that auditory detection is predictable from the properties of single branching chains as given by Eq.(14), even when the mechanisms governing transmission are far more subtle, involving sequences of processing networks performing basically different tasks.

Arbitrary Cuts:

To analyze the two-step transmission process, consider first an arbitrary cut in a branching chain. The cut divides transmission time into two distinct segments, one just prior to the cut and one after it. A parallel version of such segmentation was touched upon earlier when we sought to unravel discrete branching chains by making a cut at a particular stage. Here that earlier treatment is updated to handle continuous time and a stochastic process rather than successive stages in a discrete chain.

Call the time segment prior to the cut φt, where:

$$0 \leq \varphi \leq 1.$$

Evidently the second, right-hand time segment must be $(1 - \varphi)t$. An arbitrary cut will partition the generating function of the overall process into two distinct components just as in the case of discrete chains. The partition works out to be:

$$G_t(s) = G_{\varphi t}(G_{(1-\varphi)t}(s)). \tag{68}$$

Notice that as $\varphi \rightarrow 0$ we have:

$$G_t(s) \Rightarrow G_0(G_t(s)) = G_t(s),$$

and as $\varphi \rightarrow 1$:

$$G_t(s) \Rightarrow G_t(G_0(s)) = G_t(s).$$

These last relations identify cuts at the very beginning and very end of the chain. Such partitions were first encountered as we unravelled discrete branching chains in Eqs.(2) through (6). The effect of a partition is to nest the generating function following the cut inside the generating function prior to the cut. We have then fixed on the cut as a reference point, calculating the spread of events prior to it, and the evolution of events after it.

Now suppose a transmission process undergoes a fundamental change at the point of the cut. Equation (69) uses the nesting principle to construct a two-step transmission mechanism. The first chain's output serves as an input to a second step subsequent to the cut. This partition defines a two-step mechanism since the machinery before and after the cut is not the same. Partitioning the overall generating function, $G_t(s)$, via a cut at φt produces:

$$G_t(s) = H_{\varphi t}(J_{(1-\varphi)t}(s)). \tag{69}$$

$H_{\varphi t}(s)$ is our label for the generating function of the first step in a two-step transmission mechanism. Its duration is φt, and its output serves as input for a second step, labelled here $J_{(1-\varphi)t}(s)$. We indicate that the two steps are basically different by altering identities of the generating functions on either side of the cut. From Eq.(69) we see that an equivalent overall generating function, $G_t(s)$, also exists. It emerges whenever two such chains are connected in sequence. The form of the overall function will not necessarily conform to either of its components.

The critical point conveyed by Eq. (69) is that when $H_t(s)$ and $J_t(s)$ are established, $G_t(s)$ becomes fully determined. Specifically, if $H_t(s)$ is a pure birth process, and $J_t(s)$ is a perfectly balanced transmission line, the output of the second step, $J_t(s)$ proves to be equivalent to a single branching chain, $G_t(s)$ obtained by nesting a balanced-line generating function within a pure-birth-process. We do not yet know what the equivalent chain is, but we know how to find it.

Pure Birth With An Arbitary Cut:

To illustrate this reasoning, refer again to Eq. (52), the generating function for a pure birth process:

$$G_t(s) = \frac{e^{-\lambda t} s}{1-(1-e^{-\lambda t})\, s}.$$

Make a cut at $\varphi \cdot t$ and let:

$$p_1 = e^{-\varphi \lambda t},$$

$$p_2 = e^{-(1-\varphi)\lambda t}. \tag{70}$$

We label parameters in Eq.(70) as "1" and "2" in order to identify two separate steps, one on either side of the cut.

The cut defines a probability distribution of output counts from Step-1. For each possible count in Step-1, an identical number of chains is formed in Step-2. Hence, a Step-2 chain is set in motion by each Step-1 output event. Accordingly:

$$G_t(s) = \sum_{k=1}^{k=\infty} p_1 q_1^{k-1} \left(\frac{p_2 s}{1-q_2 s}\right)^k . \tag{71}$$

The summation in Eq. (71) omits zero because the latter is excluded from the pure birth process running in Step-1. Minimum count at the point of the cut must then be unity. Equation (71) in fact shows birth processes on both sides of the cut. The resulting geometric series sums out as follows:

$$G_t(s) = \frac{p_1\left(\frac{p_2 s}{1-q_2 s}\right)}{1-q_1\left(\frac{p_2 s}{1-q_2 s}\right)} . \tag{72}$$

The equivalent generating function, i.e., Eq.(72), displays a nested structure decreed by the rationale underlying Eqs. (68) and (69). Step-2 is sited neatly within Step-1, displacing the operating variable *s* normally found in the same location. In this case, both functions happen to be the same. Later we connect processes that change at the cut. In that event the nested generating function differs from its host, but the principle governing linkage between two successive branching chains remains the same.

After reduction, Eq.(72) becomes:

$$G_t(s) = \frac{p_1 \cdot p_2 s}{1-(1-p_1 \cdot p_2)s}, \tag{73}$$

where, of course:

$$p_1 \cdot p_2 = e^{-\varphi\lambda t - (1-\varphi)\lambda t} = e^{-\lambda t} .$$

We see that a birth process regenerates around a cut at φt. Parameters of the successive time segments simply multiply against each other, leading to easy recovery of the original birth parameter. Evidently this type of division can be extended indefinitely by cutting the time line over and over again. What we have done here reveals the repetitive inner workings of a birth process about as clearly as they can be shown. All branching chains operate under essentially the same rules.

Equation (73) also shows that birth processes are reproductive even when the parameter changes at the cut. If we set:

$$\begin{aligned} p_1 &= e^{-\varphi\lambda_1 t}, \\ p_2 &= e^{-(1-\varphi)\lambda_2 t}, \end{aligned} \tag{74}$$

the output of the second step will then be equivalent to a single-chain birth-process with a birth parameter defined by:

$$\lambda_{12} = \varphi\lambda_1+(1-\varphi)\lambda_2.$$

This is a simple example of a branching chain equivalent to a linked system with distinct components arranged in sequence.

Pure Birth Followed by Balanced Transmission:

Now turn to the generating function of a linked amplifier-transmitter as described at the beginning of this section. Incoming signals are processed first by a booster network yielding a pure birth process. The resulting impulse cascade is then transmitted to a decision center via a perfectly balanced transmission line.

These components were chosen for ease of calculation, but they offer a fairly realistic picture of sensory processing subsequent to receptor activity. Calculation is easy because mean values multiply when chains are connected in series. A perfectly balanced line with its mean of unity guarantees that overall output in the vicinity of a decision center will have a mean value identical with the local amplifier stage close to the ear.

On average, nothing is lost in transmission, although, of course, a balanced-line wipes out message-events as often as it creates progeny. The principal impact of a balanced-line is seen in the output variance. Noise created by events moving from the periphery to higher centers, adds to noise carried over from the initial birth process. The two sources seldom add up directly, due to covariance. Despite these subtleties, the combined variance can be found without difficulty . Just compute the equivalent single chain and calculate the variance of its output.

Following guidelines provided by Eq.(69), we now consider the generating function of a two-step detection-transmission device in which Step-1 is a pure birth process, and Step-2 is a balanced-line:

$$G_t(s) = \frac{p_1\left(1+\dfrac{s-1}{1-\gamma_2(s-1)}\right)}{1-(1-p_1)\left(1+\dfrac{s-1}{1-\gamma_2(s-1)}\right)}. \tag{75}$$

Parameters p_1 and γ_2 are the transmission constants of the two steps comprising the chain. Relative duration of each step is built into its parameter as before:

$$\begin{aligned} p_1 &= e^{-\varphi\lambda t}, \\ \gamma_2 &= (1-\varphi)\gamma t. \end{aligned} \tag{76}$$

The "parent" generating function in Eq.(75) is a birth process whose functional form was originally developed as Eq.(52) (see also Table 1). Nested within the parent is a balanced-line "offspring" derived earlier in Eq.(54). This nested structure is decreed by the logic underlying Eq.(69). It governs the construction of all two-step generating functions.

Subtract unity from both sides of Eq.(75) and switch over to tails form:

$$Q_t(s) = \frac{1}{1+p_1\gamma_2-(1+p_1\gamma_2-p_1)s},$$

$$Q_t(s) = \frac{1/1+p_1\gamma_2}{1-\left(1-\dfrac{p_1}{1+p_1\gamma_2}\right)s}.$$

(77)

The expression in Eq. (77) signals a birth-death process for the output of our two-step detector-transmitter mechanism. Its tails generating function conforms expressly to the version developed earlier in Eqs. (10a) and (10b). Based on guidelines provided by Eq.(10a), constants of the birth-death chain must then be:

$$p_t(0) = \frac{p_1\gamma_2}{1+p_1\gamma_2},$$

$$\beta = \frac{p_1}{1+p_1\gamma_2}.$$

(78)

We have established that this two-step detector-transmitter device, a pure birth process feeding a perfectly balanced line, is equivalent in output to a birth-death branching chain whose propagation parameters can now be deduced from the constants in Eq. (78). The output probability distribution is evidently:

$$p_t(0) = \frac{p_1\gamma_2}{1+p_1\gamma_2},$$

$$p_t(k) = (1-p_t(0))\cdot\beta\cdot(1-\beta)^{k-1},$$

k=1,2,3,...

(79)

All that remains is to calculate the mean and variance of the output in Eq. (79). The mean should be easy since output in a two-step mechanism is always the product of its component mean values. Balanced-lines deliver their input unchanged in expectation. The major alteration will be an increased variance. With a single input going in, mean output of a balanced line is unity. Accordingly, for the two-step process under consideration:

$$Mean\ Output = e^{\varphi\lambda t} \cdot 1.$$

To confirm this intuitive argument, go to Eq. (77) and calculate the mean directly:

$$Q_t(1) = \frac{1}{p_1} = e^{\varphi\lambda t}. \tag{80}$$

The result confirms intuition. A variance also follows when we differentiate $Q_t(s)$ with respect to s:

$$Q_t'(1) = \frac{1-p_1+p_1\gamma_2}{p_1^2}. \tag{81}$$

Now compute the variance via Eq. (17):

$$\begin{aligned} Variance &= \frac{1-p_1+2p_1\gamma_2}{p_1^2}, \\ &= e^{\varphi\lambda t}(e^{\varphi\lambda t}-1)+2e^{\varphi\lambda t}(1-\varphi)\gamma t. \end{aligned} \tag{82}$$

If φ, the fraction of the two-step mechanism devoted to pure-birth detection, approaches unity, output variance approaches a birth process. Similarly when $\varphi \to 0$, we get back the variance of a balanced line. Should immigrant pulses ever contaminate the balanced line, a secondary noise will be added to both mean and variance of the output. All the foregoing arguments are based on the supposition that such intrusions are negligible.

The existence of a single chain, equivalent to a two-step or multi-step transmission process, indicates that we are dealing with a robust device whose properties are not heavily dependent on details of operation assumed in advance.

Moreover, the variance due to branching, i.e., the magnitude of internal transmission noise, may easily become a dominant factor in any accounting of noise sources. This possibility mandates caution in adopting detection theories based exclusively upon stimulus properties. Everything learned thus far in our study of branching systems advises such caution. Problems associated with detecting small intensity changes may arise largely in transmission, rather than in acoustic signals, or auditory receptors. We should reserve judgment until more is known of the actual mechanisms of auditory transmission.

Our study of two-step detection mechanisms might well have concentrated on erosive transmission of message-events following initial amplification. This would not have altered our conclusions about output significantly. Omitting periodic boosts to message-events implies output-input ratios of less than unity in the second step of the process (see Table 1). Since the mean values of the two steps multiply against each other, central output after erosive transmission must be smaller than the input provided by the initial boost at the periphery. Such losses are possible, perhaps even likely; but if they occur, transmission is invariably more complex than with balanced lines at the second step.

5. ADAPTATION LEVEL

Ideally, alerting signals are produced by a single chain that has grown silent during continuing stimulation. The chain adjusts its propagation parameters to minimize amplification when dealing with a steady-state stimulus. This is accomplished by balancing gain and dropout parameters against each other. The resulting condition ($\lambda = \mu$) is identified as a balanced-line. It was studied in our earlier discussion of Eq. (38) through Eq. (41).

Balanced-lines establish a quiet baseline suitable for detecting incremental changes. Any alteration of the balance causes the chain suddenly to amplify, resulting in a cascade of impulses at the output. Average discharge varies with incremental intensity, but there is no necessary linkage between the shape of a discharge and the time course of a stimulus increment. Mechanisms of this type are not meant to be faithful copies of the input. They sound an alert, a warning that something new is happening out there. The format of increment detection relies on such warnings if the proper machinery is built into the sensory channel. Laming (1986) offers a variety of arguments suggesting that it is. Moreover, he shows that difference discrimination (comparison of two distinct stimuli) and increment detection (blip on a steady background) tend to follow different detection laws.

The balanced-lines described in this section resemble perceptual frames of reference that psychologists have traditionally identified as "adaptation-levels". (See Helson, 1947.) We first consider adaptation-levels established near absolute threshold, then turn to typical adaptation-levels involved in increment detection. In each case detection is mediated by alerting signals emerging from a two-step mechanism.

Ultralinearity in Two-Step Detection:

Two-step detection-transmission is the scheme described in our argument leading up to Eq. (79): a brief auditory stimulus sets off a peripheral birth process (Step-1); the latter then feeds a balanced-line conveying amplified stimulus information to a central decision region (Step-2). The mechanism is a branching chain constructed from two elementary building blocks connected end-to-end. Output is a brief warning signal pointing to a threshold auditory event.

Our tails generating function governing central output in this special case was developed in Eq. (77). It is:

$$Q_t(s) = \frac{1/1+p_1\gamma_2}{1-\left(1-\dfrac{p_1}{1+p_1\gamma_2}\right)s},$$

Where p_1 and γ_2 are constants associated with each step:

$$p_1 = e^{-\varphi\lambda t},$$

$$\gamma_2 = (1-\varphi)\gamma t,$$

and φ is the fraction of the entire chain occupied by the first step.

Evidently the energy level of a stimulus triggering this warning influences the size of the response cascade produced in Step-1.

The balanced-line in Step-2 limits further amplification to an amount just sufficient to compensate for losses in transmission. An efficient transmitter boosts the emerging impulse count until it is identical in expectation with the entering impulse count. This prescription implies fixed propagation constants in Step-2. Hence, Step-2 should be independent of stimulus influences.

This building-block approach to detection is based on a trio of independent operations: amplification, transmission, and decision. The approach is reminiscent of a black-box analysis, and is illustrated in Figure 4.

- FIGURE 4 HERE -

These black boxes, however, are not vaguely defined. They are very sophisticated. The manner of their connection has a decisive impact on the additivity of information passing through them.

To clarify this point put auditory stimuli through a log transform (see Eq. (28)) just prior to activating the pure birth process in Step-1. The parameter associated with Step-1 then becomes:

$$p_1 \Rightarrow \frac{1}{1+ax},$$

where λ_x now carries a subscript because it varies with stimulus intensity x. Parameters φ and t are absorbed into the units-constant.

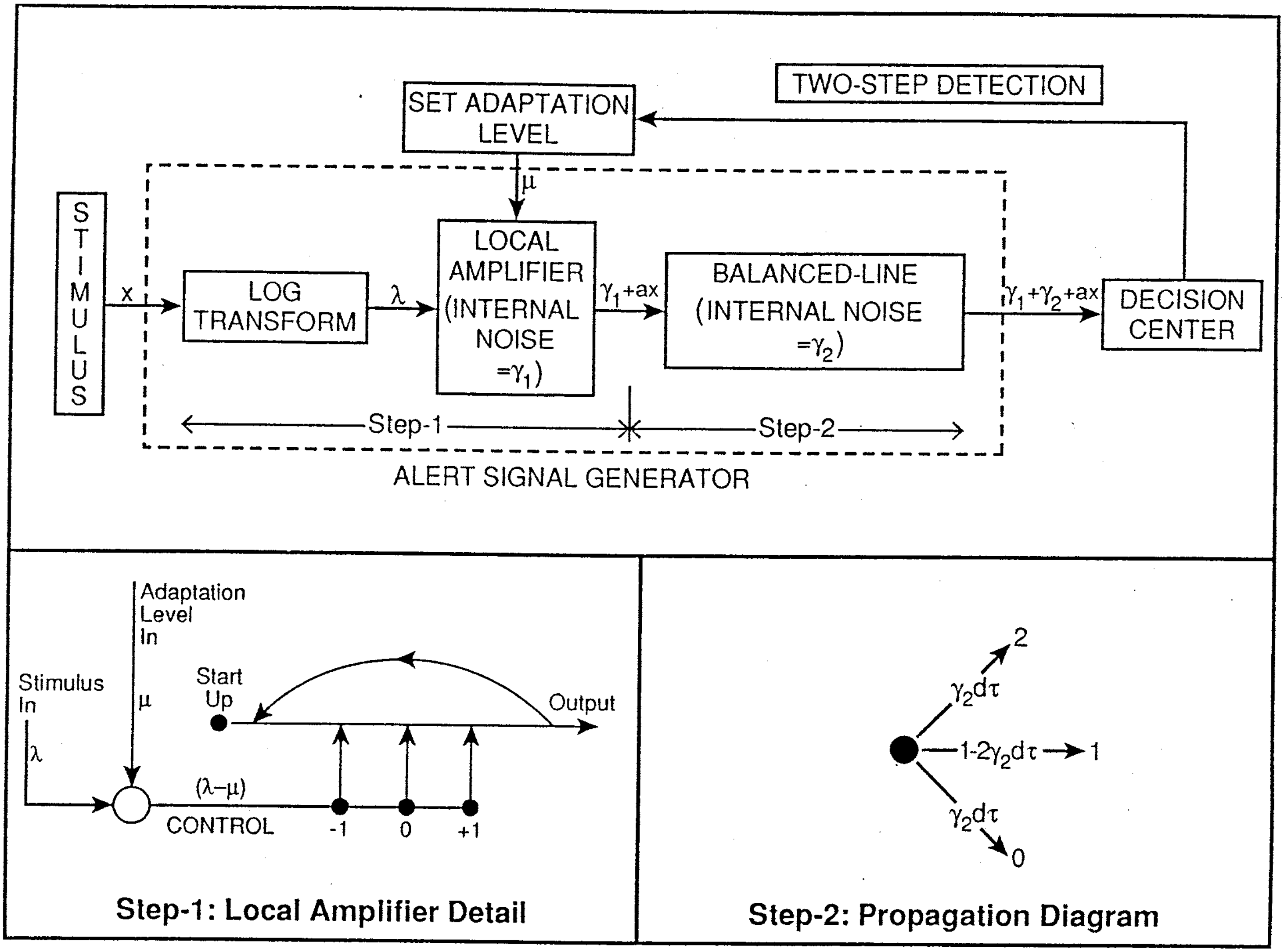
Because Step-2 is functionally separate from the initial amplifier, the transmission constant γ_2 remains unaffected by this maneuver. Accordingly:

$$Q_x(s) \Rightarrow \frac{\left(\dfrac{1}{1+\dfrac{\gamma_2}{1+ax}}\right)}{1-\left(1-\dfrac{1/1+ax}{1+\gamma_2/1+ax}\right)s},$$

$$Q_x(s) \Rightarrow \frac{1+ax/1+\gamma_2+ax}{1-(1-\dfrac{1}{1+\gamma_2+ax})s}.$$

(83)

Our tails generating function for the two-step detector shows the familiar form decreed earlier by Eq. (10b), signalling a birth-death process. Distribution constants of the latter are evidently:



Step-1: Local Amplifier Detail

Step-2: Propagation Diagram

Figure 4: Black box diagram showing two-step detection. Stimulus energy activates receptor. Output of latter undergoes log transform. Neural signal then builds up in local (i.e., peripheral) amplifier network. Output of amplifier varies with input intensity. Amplifier is a branching chain triggered by a single input event as shown in Step-1 diagram. Stimulus and background-level combine to set control ($\lambda - \mu$) on multiplication characteristics of amplifier. This regulates size of output cascade. Diagram shows equation parameters at output of processing units in which they are generated.

Amplifier output fed to balanced-line whose propagation operator is illustrated in Step-2 detail. Balanced-line delivers noisy but generally faithful version of amplifier cascade to decision center. Sudden burst of pulses at this center constitutes alert-signal, indicating change in background, and calling for response by decision center. Included in such response is possible resetting of adaptation level.

Operation of this mechanism is described in Eqs. (83)-(97) in text under section headed Increment Detection at Adaptation Level.

$$p_x(0) = \frac{\gamma_2}{1+\gamma_2+ax},$$

$$\beta = \frac{1}{1+\gamma_2+ax}.$$

(84)

With guidelines supplied by Eq. (84) it is now easy to construct the output counting distribution of a two-step detector, as well as the probability of exceeding any fixed output count as stimulus intensity changes. We omit these proofs. They were presented earlier in Eqs. (79) and (77). Interest centers now on measuring internal noise by connecting it with quantities we can measure.

The generating function in Eq. (83), as well as the constants in Eq. (84), depict auditory detection in a self-contained alerting device, working near absolute threshold. In view of Eq.(84), γ_2 is seen to be an (unmeasured) internal noise accompanying transmission. The effective intensity driving the mechanism must then be:

$$\gamma_2+ax,$$

which is our so-called "ultralinear" expression developed originally as Eq. (42).

Threshold intensity must be large enough to overcome the internal transmission noise. The latter is then measurable as an intensity added to any external signal, bringing it up to audibility. Noise magnitude is deduced; not measured directly. It does not sound like a noise. In most situations it is not audible at all. But our arguments show it is there and also measurable. The two stimuli, external signal and internal noise, are additive if the detector is configured as two distinct mechanisms in sequence.

Additivity is also evident from analysis of the output variance. Following Eq.(82) we have:

$$\textit{Output Variance} = (1+ax)(2\gamma_2+ax)\,. \tag{85}$$

A major component of the output variance is attributable to survivors, i.e., non-zero counts. It is given by:

$$\begin{aligned}\textit{Survivors Variance} &= \frac{1-\beta}{\beta^2}\,,\\ &= (\gamma_2+ax)(1+\gamma_2+ax)\end{aligned} \tag{86}$$

Equation (86) portrays a shifted, geometric distribution whose mean is $1 + \gamma_2 + ax$. Evidently γ_2 corresponds to the adaptation-level of the threshold detector. Incoming stimuli are added to this level in determining the survivors' counting distribution.

Full output variance in Eq. (85) requires weighting Eq. (86) by the survival probability:

$$\textit{Weighted Survivors' Variance} = (1+ax)(\gamma_2+ax) \tag{87}$$

Zero-count specifies a single state. It contributes no variance except for its mean-square difference from the overall mean value *1+ax*. There are two such mean-squares, one from the zero-state and one from the distribution of survivors. Adding them up, each weighted by its respective probability, we obtain:

$$\begin{aligned} m.s.d. &= \frac{\gamma_2}{1+\gamma_2+ax}(0-(1+ax))^2+\frac{1+ax}{1+\gamma_2+ax}((1+\gamma_2+ax)-(1+ax))^2, \\ &= \gamma_2(1+ax). \end{aligned} \tag{88}$$

Output variance will be the sum of these two components: Eqs. (87) and (88).

We see that transmission noise in Step-2 establishes the adaptation-level against which incoming stimuli are detected. Absolute thresholds are measured relative to this adaptation level.

<u>Additive Balanced-Lines</u>:

Adaptation-level concepts are usually found with stimulus conditions far removed from absolute threshold. The background is said to define a reference frame. Sensory, perceptual, and evaluative systems are seen as adjusting to the frame. Stimuli are perceived relative to the frame rather than directly.

An amplifier-transmitter mechanism adapts to high intensity backgrounds by adjusting propagation parameters at the intial amplifier step, neutralizing the background and converting operation to a balanced-line ($\lambda=\mu$). In that event, Step-1 cannot be a pure birth process. There is no way of getting to balance unless the mechanism can branch to the zero-state.

Accordingly, at non-zero background levels Step-1 must have a birth-death configuration. With the log transform, the initial amplifier step would be configured as follows:

$$H_x(s) = 1 + \frac{(1+ax)(s-1)}{1-(\gamma_1+ax)(s-1)} . \tag{89}$$

This generating function is adapted from Eq. (32). It describes the first, amplifier step of our increment detection process. As before, x symbolizes increment intensity, while γ_1 is an adaptation-level set by the reference frame in Step-1. Here we make a further assumption that background and increment are additive within the initial amplifier. Hence:

$$\gamma(x)=\gamma_1+ax.$$

Actually, nothing in this or subsequent arguments requires such within-step additivity. We could get by without it, and in the interest of generality perhaps should proceed that way. The assumption is added here because it simplifies analysis and makes our presentation much easier to follow.

Set $ax = 0$ in Eq. (89). The generating function now describes a balanced-line, but it remains the amplifier step of a two-step mechanism. Balance is created as the amplifier neutralizes a steady-state stimulus by setting up an adaptation level at γ_1.

Thus, with increment intensity at zero, a two-step detector-transmitter becomes two balanced-lines joined end-to-end. Intuition suggests they should be additive. Can we prove it?

Proof is easy. Recall our theorem (Eq.(69)) on arbitrary cuts in a branching chain. Apply the nesting principle with increment intensity $ax = 0$, while a second balanced-line displaces the operational variable s in Eq.(89):

$$G_x(s) = 1+ \frac{\left(1+\dfrac{s-1}{1-\gamma_2(s-1)}-1\right)}{1-\gamma_1\left(1+\dfrac{s-1}{1-\gamma_2(s-1)}-1\right)},$$

$$= 1 + \frac{s-1}{1-\gamma_2(s-1)-\gamma_1(s-1)},$$

$$= 1 + \frac{s-1}{1-(\gamma_1+\gamma_2)(s-1)}. \tag{90}$$

Equation (90) is the generating function equivalent of the two-step output. We see that sequenced balanced-lines are additive. At zero increment, the two-step system is equivalent to a single balanced-line with internal noise equal to the sum of noise parameters associated with the two steps.

Notes on Notation :

With the introduction of two-step detection and its equivalent single branching chain, our notation grows unavoidably complex. We review it now in order to clarify current arguments as well as subsequent discussion of increment detection.

Parameters of the initial amplifier (see Fig.4) are labelled λ and μ respectively. When a branching chain is cut, parameters of steps on each of the cut side carry an indentifying subscript. Hence:

$$p_1 = e^{-\varphi(\lambda-\mu)t}$$

is a constant describing amplification in Step-1. An additional parameter φ denotes the fraction of the full chain length t allocated to Step-1.

Step-2 is a balanced-line transmission network receiving data from the first detection step, i.e., the initial amplifier, and delivering it to a high-level decision center. Balanced-lines have only one propagation constant. Gains and losses offset each other precisely. We label this single constant γ. Hence:

$$\gamma_2 = (1-\varphi)\gamma t$$

is the gain/loss propagation constant and the fraction of the full chain occupied by Step-2. A subscript on γ_2 indicates that these constants characterize Step-2, i.e., the transmission phase, of the detection process. We also refer to γ_2 as an "internal noise" accompanying transmission.

When our Step-1 amplifier listens to a continuous background, it adjusts gain and loss parameters to minimize amplification. We label the noise in this near-quiet state γ_1, where:

$$\gamma_1 = \varphi\lambda t,$$

and $\lambda = \mu$. Accordingly, γ_1 is the internal noise of the Step-1 amplifier working at an adaptation-level set by the intensity of a steady-state background.

Finally, the sum of internal noise factors when two balanced-lines are joined end-to-end, is identified as:

$$\begin{aligned}\gamma_{12} &= \gamma_1+\gamma_2\\ &= (\varphi\lambda+(1-\varphi)\gamma)t.\end{aligned} \tag{91}$$

Thus, γ_{12} is a noise constant characterizing a single balanced-line equivalent to our two-step detector with the Step-1 amplifier at adaptation level.

Increment Detection At Adaptation Level:

Earlier we showed that a two-step detector can neutralize steady-state background stimuli by adjusting parameters of the initial amplifier step so as to create a balanced-line condition. At this baseline, the two operations constitute a connected pair of balanced-lines with non-identical noise constants. The first, γ_1, depends on the intensity of the background stimulus, whereas the second, γ_2, is fixed by conditions within the auditory system. We have established that this pair of constants is additive if balanced-lines are connected end-to-end. (See Eq. (90)). In that event a single constant:

$$\gamma_{12} = \gamma_1 + \gamma_2$$

reflects the internal noise of a single chain whose output is identical to that of the two-step detector.

Using such additivity, we can write down the generating function of a two-step mechanism, detecting and transmitting increments above a steady-state background. Equation (89) is our guide. Begin by setting the initial adaptation level at γ_1, i.e., an arbitrary background intensity:

$$G_x(s) = 1 + \frac{(1+ax)\left(1+\dfrac{s-1}{1-\gamma_2(s-1)}-1\right)}{1-(\gamma_1+ax)\left(1+\dfrac{s-1}{1-\gamma_2(s-1)}-1\right)}. \tag{92}$$

Equation (92) shows a Step-2 balanced-line nested within the initial birth-death process. Stimuli entering our Step-1 amplifier are first put through a log transform (Eq. (28)). The resulting generating function was developed originally as Eq. (32). Within-step linearity was added in Eq. (89). We then nest a balanced-line corresponding to Step-2 in Eq.(89), producing Eq.(92), a single branching chain equivalent to our two step process.

Reduction of Eq. (92) yields:

$$G_x(s) = 1+\frac{(1+ax)(s-1)}{1-(\gamma_{12}+ax)(s-1)}, \tag{93}$$

where γ_{12} is the sum of internal noise constants as shown in Eq. (91).

Now switch over to tails form:

$$Q_x(s) = \frac{1+ax/1+\gamma_{12}+ax}{1-\left(\frac{\gamma_{12}+ax}{1+\gamma_{12}+ax}\right)s}. \tag{94}$$

Equation (94) is again in the standard format established by Eq. (10b). Except for the size of the internal noise, the expression is identical to Eq. (83) developed for a pure birth process at the initial amplifier.

Distribution constants of the two-step output are:

$$p_x(0) = \frac{\gamma_{12}}{1+\gamma_{12}+ax},$$

$$\beta = \frac{1}{1+\gamma_{12}+ax}.$$

(95)

Again, apart from the size of the overall internal noise, the parameters in Eq. (95) are the same as those found earlier with a pure birth process at absolute threshold. (See Eq. (84).)

Why should a two-step detector equate pure-birth at absolute threshold with birth-death for backgrounds above threshold? A link is built directly into the adaptation level. Differences between pairs of amplifier parameters ($\lambda - \mu$), contain a background setting as one term. (See Fig. 4.) When the background μ goes to zero, $\gamma_1 = 0$, and $\gamma_{12} = \gamma_2$. The equivalent branching chain (Eq.(94)), with adaptation level γ_{12}, then shades into a pure birth process (Eq.(83)) as background intensity drops out. A birth-death process limits on pure-birth when the background vanishes.

The probability that an alerting signal from this detector exceeds any fixed number of pulses can be calculated by expanding $Q_x(s)$ in powers of s. Equation (94) shows the expansion to be a geometric series:

$$Q_x(s) = \sum_{k=o}^{k=\infty} \left(\frac{1+ax}{1+\gamma_{12}+ax} \right) \left(\frac{\gamma_{12}+ax}{1+\gamma_{12}+ax} \right)^k s^k. \tag{96}$$

It is now easy to construct psychometric functions expressing the relation between detection probability and increment intensity for various values of a detection criterion, i.e., "critical number", in the burst of impulses emerging at the chain's output. These curves are built into the coefficients of s^k in Eq.(96).

Notice that our combined adaptation-level γ_{12} is not determined completely by the noise components of the first amplifier step. Adaptation-level includes an added source of noise from the balanced line in Step-2. False alarms can be suppressed more effectively by setting the amplifier's adaptation-level some distance above average background intensity. This is especially clear when the background happens to be an acoustic noise. Hence, our two-step detection mechanism controls false alarms by trading off its detection criterion k against its adaptation-level γ_{12}.

Based on the constants developed in Eq.(95), the output counting distribution of a two-step increment detector as depicted in Eqs.(93) and (94), i.e., a birth-death amplifier feeding a balanced transmission line, should be given by:

$$p_x(0) = \frac{\gamma_{12}}{1+\gamma_{12}+ax},$$

$$p_x(k) = (1-p_x(0))\cdot\beta\cdot(1-\beta)^{k-1}, \quad k=1,2,3,\ldots \tag{97}$$

where β is the branching chain parameter defined in Eq. (95).

Evidently, output counting distributions developed from all such two-step devices are intimately related. Family resemblances appear because all have the same relation to a simple branching chain. The parameters $p_x(0)$ and β differ from case to case, but each separate device generates an output having a characteristic geometric, or as it is often called, Bose-Einstein form. The latter is imposed by a characteristic noise added to message-events passing through a continuously branching Markov chain.

Increment-detection devices such as these work in essentially the same way at all background-levels (including zero background at or near absolute threshold). When background level increases, our internal noise parameter γ_{12} will increase along with it. These changes alter the resolving power of a detector, undercutting its ability to discriminate increments of fixed size as background intensity increases. Nearly all branching chains work in this way.

Hence, virtually the entire family of alerting signals described in this paper are Weber's-law devices.

Balancing Mechanisms:

We claim that two-step (i.e., amplify-transmit) detection devices should be able to neutralize steady-state backgrounds by adjusting parameters at the initial amplifier to establish balanced-line conditions. We say also that adaptation levels should not be set exactly to the background intensity, preferably some distance above. This would be an excellent scheme for suppressing false alarms when the background is an acoustic noise.

Requirements such as these are not trivial, and it is fair to ask whether a relatively automatic alarm device (which is what we have in mind for this increment detector) could know how or when to make such complex adjustments.

Of course, a high-level controller might be reading the stimulus environment continuously, sending out regular instructions to update adaptation-levels over an entire range of auditory processes. This does not seem beyond the ingenuity of a system renowned for its ability to organize and sharpen wholesale patterns of information flooding up from the two ears. But even a sophisticated controller would have to deliver its instructions to peripheral locations, directing structures there to make needed adjustments. Eventually we should be able to find and identify the mechanisms involved in such interventions.

Laming (1986, pp. 156-167) conjectures a balance between excitatory and inhibitory impulse trains, creating a quiet detection unit ready to respond to any incremental change in a continuous background. His detailed arguments suggest that our amplifier parameters, λ_x and μ_x, might represent opposed control systems in a branching chain, one excitatory and the other inhibitory.

A new stimulus would drive up λ_x causing the chain to amplify briefly. If the same stimulus continues, it becomes a background condition, and μ_x then increases to offset the newly-elevated level of excitation. A chain could be brought to a new equilibrium this way. The equilibrium is said to be new because it reflects a new internal noise-level. Sensing and neutralizing continuing excitation is not a complicated problem. We can think of several ways it might be done within the branching chain itself, requiring no elaborate or sophisticated intervention from high-level control mechanisms. Local management requires only an array of excitatory and inhibitory control sites on each branching chain, and a means for leaking continuous excitation into the counterpart control sites. Machinery such as this comes into equilibrium (adaptation) at a level set by any continuous background.

What if the stimulus happens to be a decrement? The line of reasoning just developed suggests that λ_x should drop immediately to a lower value. If this new level persists, a lower value of λ_x becomes the new background. This will tend to pull the counterpart μ_x down as equilibrium is restored.

The problem with such adjustments is that during the branching chain's initial reaction to a decrement, and prior to restoration of the balanced-line condition, our log transform in Eq.(28) would swing negative briefly. A shift such as that tends to block booster action at Step-1. There would be no output cascade. Hence, a detector using the log transform in Eq.(28) sets off an alarm only when it sees an increment.

The implied correction of Eq.(28) makes the log transform symmetric for both increments and decrements:

$$\ln(1+|ax|) = |\lambda_x - \mu_x|t. \tag{98}$$

Rectifier action such as that shown in Eq.(98) produces mirror-image psychometric functions above and below adaptation level. Each function starts from the same intensity. Each involves departures of the same magnitude, up or down.

6. REPRESENTATION OR ALARM SIGNAL?

The basic building block in our study of transmission is a single branching chain driven by a single start-up pulse. An initiating event multiplies and propagates through the chain, eventually becoming a cluster of events flowing toward the output. Governing the flow is a Markov operator applied continuously to every message-event. At random some are multiplied, becoming two or more pulses. Others drop out, disappearing before they can be recorded. The result is an evolving cluster of impulses, whose total moves up and down in a random walk as it approaches the chain's output. The process begins with a single input, and ends in a flurry of outputs. Output clusters are partly noise, reflecting random losses and multiplications encountered as progeny of the initiating event move through the chain .

Two Roads to Detection:

Our basic building block can be used in at least two different ways. A particular output can be viewed either as a tiny element in a large-scale representation of the stimulus environment, or as a completely self-contained alarm signal. In the first instance, thousands of building blocks form an amplified, noisy, but otherwise faithful, central copy of a peripheral stimulus pattern. The branching chains we have described constitute the elementary units used in constructing such representations. A sophisticated central analyzer might then refine this construction producing enhanced, sharpened images of the auditory environment.

We have concentrated on the second application of the basic building block: a single chain, or single-chain-equivalent, functioning as an alert-signal. The latter is discussed and evaluated throughout the paper, particularly in our treatment of adaptation-levels and two-step increment detection.

Part of our interest in alerting mechanisms traces to the novelty of the idea. These signals are psychological events conveying some form of alarm. They do not reflect subtle calculations imposed by decision theory. Laming (1986) was chiefly responsible for this outlook. He introduced alerting signals into sensory analysis, and developed their significance in increment detection.

Traditional treatments of auditory detection are representational, invoking the full panorama of the stimulus environment. Alerting signals are limited to a single narrow condition: a minute change in a continuous background stimulus.

Both outlooks are important in auditory analysis. The world we live in is filled with events to be represented and warnings to be heeded. When man hunted to survive, the slightest snap of a twig somewhere off in the distance conveyed instant, crucial information. Something out there had suddenly changed, and change was often deadly. Abrupt alarms of this sort are quite different from, say, the beauty of Placido Domingo's voice climbing up through the upper reaches of "*e lucevan le stelle*". The latter is a pure representational phenomenon. Analytical accuracy and broad band fidelity are the qualities required of a representational transmission system.

The faint sound of a twig breaking comes from a more primitive realm. Accurate detection of weak transients is the key requirement. The process draws attention away from whatever else may be going on, focusing it on sudden changes in the auditory environment. Increment detection is more likely to depend on reactions such as this than on hi-fi signal representations subjected to mathematical scrutiny.

Representational Detection:

A quintessential representational model is Hecht, Shlaer, and Pirenne's (1942) treatment of absolute visual thresholds. This important paper was mentioned earlier in our discussion of erosive transmission. We refer to it here as HSP.

According to HSP, photons absorbed in retinal rods give rise to a flux of nerve impulses. This neural flux is thought to be a faithful copy of an optical flux incident on the retina. Ordinary incandescent light is Poisson distributed. Hence, at absolute threshold, if the neural response is a copy of the optical input, we should find psychometric functions similar to those in Eq.(96) but based on the Poisson distribution. A template of such functions was constructed relating detection probability to intensity. Each function defined a criterion or "critical number" required for detection. Finally, HSP compared psychometric functions obtained experimentally with those on the template, and extracted a critical number telling how many photon-absorptions should be required for detection. They found the critical number to be quite small; of the order of 5-10 events.

For HSP, the pattern of impulses at the input of a visual transmission network was a faithful copy of the stimulus itself, diminished slightly by random losses. Auditory alert-signals, although triggered by stimulus events, are not copies of anything. They follow their own growth rules. Magnitude and time course bear no particular resemblance to sensory patterns. These are alert-signals, not representations.

HSP used an erosive transmission mechanism (random losses - no boosting) to convey photons from the surface of the cornea to a point at which detection occurs. In our own approach to representational detection, we employ balanced-lines to model the transmission stages mapping a peripheral stimulus pattern to a central locus. The pattern itself, not an accompanying alarm signal, is detected centrally.

We start as HSP did with a Poisson distribution of photons incident on the eye. A random subset of these is absorbed in the retina, generating a distribution of neural events just behind the receptor surface. The Poisson mean of this neural response is labelled m.

To simplify our argument, exclude local amplification. Simply connect each event in the Poisson start-up pattern directly to a balanced-line conveying a representation of the event to a decision area higher up in the system. There are as many balanced-lines as start-up events, each adding internal noise to the transmitted pattern. Hence, the system delivers a noisy, slightly boosted version of the Poisson input to the decision region. Our problem is to analyze this mapping.

Each start-up pulse in the stimulus pattern sets off a balanced-line transmission sequence, symbolized here by:

$$J(s) = 1+\frac{s-1}{1-\gamma(s-1)}. \tag{99}$$

The label of the generating function in Eq. (99) is changed to $J(s)$ because it will be nested in a Poisson distribution of start-up pulses, the initial condition proposed by HSP. Whenever branching chains are triggered by events having a probability distribution, output can be found by nesting the transmission generating function within the generating function of the triggering distribution. Accordingly:

$$G_m(s) = e^{m(J(s)-1)}. \tag{100}$$

Here $G_m(s)$ represents the output counting distribution at decision level. We know its mean to be *m* because individual balanced-lines connected to start-up events, reproduce the mean value of the Poisson input exactly. Combining Eqs. (99) and (100) yields:

$$G_m(s) = e^{\frac{m(s-1)}{1-\gamma(s-1)}}.$$

(101)

The central representation in this instance is a compound Poisson distribution (see Feller, 1957, p.270). It is not strictly Poisson unless the balanced-line constant $\gamma = 0$. The latter would imply noise-free transmission, an unlikely prospect. What else can be said about Eq.(101)?

Approximation When Transmission Dominates:

In view of the basic definition of e^z:

$$e^z = \lim_{m\to\infty}\left(1+\frac{z}{m}\right)^m,$$

if the Poisson mean *m* is of reasonable size, $G_m(s)$ can be approximated by:

$$G_m(s) \approx \left(1+\frac{s-1}{1-\gamma(s-1)}\right)^m = J^m(s).$$

(102)

Equation (102) shows the output as a sum over a fixed number of balanced-lines, each derived from a packet of energy making up the stimulus. Our approximation treats this number as fixed, whereas of course, it has a Poisson distribution.

Since the mean and variance of any single balanced-line are 1 and 2γ respectively, statistics of the overall output must be:

$$\text{mean} = m\,,$$
$$\text{variance} = 2m\gamma. \tag{103}$$

This type of approximation would be appropriate if our balanced-line noise constant γ were large in relation to stimulus variance. Noise introduced in transmission would be expected to swamp the stimulus variability. Hence, combined input and transmission noise levels could be represented adequately in the output by transit noise alone. The number of balanced-lines contributing to the total output would then be proportional to the average energy of the stimulus.

Balanced-lines mapping stimulus patterns to a central locus are viewed as common carriers in representational models such as this. The transmission noise constant γ characterizing each balanced-line is said to be fixed by conditions internal to the auditory system, not by intensity or indeed any other stimulus property.

Evidently the Cox-Smith (1954) superposition theorem governs output counting in Eq. (102). When many parallel transmission lines run side by side as they do here, superposition effects encountered earlier in Eq. (48) are soon rediscovered. Equation (103) finds the mean and variance of this many-line output to be proportional, implying square-root-law dectability.

Factorial Moments:

A little experimentation shows that branching-chain generating functions yield counting probabilities when expanded in powers of *s*, and factorial moments when expansion is carried out in powers of *s-1*. This unusual property is easy to prove. It constitutes an unexpected gratuity greatly enhancing the scope of any analysis. For example, expanding Eq. (99) in powers of *s-1* produces:

$$J(s) = 1+(s-1)+2\gamma\frac{(s-1)^2}{2}+6\gamma^2\frac{(s-1)^3}{3!}+\ldots.. \tag{104}$$

The k^{th} factorial moment of the balanced line distribution is then given by the coefficient of:

$$(s-1)^k/k!.$$

Our expansion of Eq.(104) provides still another way to calculate means and variances of balanced-lines. Evidently the mean is unity (coefficient of *s-1*); while the second factorial-moment (coefficient of $(s-1)^2/2!$), and hence also the variance, are given by 2γ.

To find the actual (not approximate) mean and variance of the compound Poisson distribution forming our stimulus representation, expand its generating function, Eq.(101), in powers of $s-1$:

$$G_m(s) = 1+\frac{m(s-1)}{1-\gamma(s-1)}+\left(\frac{m(s-1)}{1-\gamma(s-1)}\right)^2/2! + \cdot\cdot\cdot +,$$

$$G_m(s) = 1 + m(s-1) + (2m\gamma+m^2)\frac{(s-1)^2}{2!} + \cdot\cdot\cdot + \quad . \tag{105}$$

Mean output is evidently m in precise agreement with our earlier approximation in Eq.(103), but output variance is somewhat larger:

$$\begin{aligned} mean &= m; \\ variance &= 2m\gamma + m^2 + m - m^2, \\ &= m(2\gamma+1)\,. \end{aligned} \tag{106}$$

As the transmission noise constant γ goes toward zero, variance devolves on the Poisson input rather than on zero as in Eq.(103).

Moment generation becomes feasible when branching-chain generating functions can be expanded in powers of $s-1$. Accordingly, we can view $G_m(s)$ in Eq. (101) as a type of moment generating function. Let m grow large while γ approaches zero, so that the product remains fixed and finite. Now expand the exponent in Eq. (101) in powers of (s-1):

$$G_m(s) = e^{m(s-1)+2m\gamma\frac{(s-1)^2}{2}+\textit{terms approaching zero}} . \tag{107}$$

We see that $G_m(s)$, the output of a Poisson stimulus representation transmitted along balanced lines, can be represented as the sum of a Poisson input variable and a symmetrical distribution with mean zero and variance $2m\gamma$. The combination resembles a normal (gaussian) limit with mean m and variance $2m\gamma$, but the variances of the two components add to produce the result in Eq. (106).

Vision vs. Audition:

Balanced-line transmission of a Poisson stimulus pattern produces square-root-law discrimination at a central locus provided that the sole effect of intensity is to change the number of lines operating. This is a key restriction underlying Eq. (106). We have shown in effect that HSP's picture of Poisson detectability can be retained even after transmission to a central decision area if messages are conveyed on balanced-lines. Such central detection will be a bit noisier than its peripheral counterpart, but no other radical differences separate the two detection schemes.

Earlier we excluded local amplification prior to balanced-line transmission in order to simplify our comparison of erosive and balanced-line transmissions. But now we see that preamplification can have a profound effect on detection. With branching chains, the initial boost of a local amplifier creates a Weber's Law device. Amplification affects both the size and variance of the cascades transmitted by each separate line or chain.

Incremental intensity changes can then be viewed as producing relatively small alterations in the number of lines operating, but relatively large changes in the variance along each line. Hence, local amplification transforms the detection law governing incremental change from square-root form to Weber's Law. It is probably important to point out again that vision research has uncovered a region of square-root-law detectability, some 4-6 log units wide, just above absolute threshold (see Bouman, 1961, p. 385).

It appears that a case can be made for square-root-law detection in vision, at least in the vicinity of absolute threshold. We know of no comparable findings in audition. The point is important for deciding whether auditory alert-signals are real. Is there any evidence of alert-signalling in the visual field? Suitable evidence might come from finding Weber's law with certain sudden changes in the visual periphery.

Difference Discrimination:

In the transmission framework we have developed, alert-signals are triggered by incremental shifts: abrupt departures from a steady-state background. What happens when this same framework is applied to difference discrimination: pairs of stimuli presented against a background of silence?

The format of difference discrimination puts adaptation-levels at or near zero. Each member of the stimulus pair must then trigger off its own alert-signal as well as its stimulus representation. Detection requires a comparison, either between these alert-signals or the transmitted patterns themselves. Such pair-comparisons impose entirely new rules on auditory detection. They force comparative judgments rather than simple reactions to a change of state.

In an earlier paper we calculated detection probabilities from comparisons of auditory alerting signals involving simple increments, pedestals, and pairs of isolated signals; all in two alternative forced-choice format (McGill and Teich, 1991b). Higher detection probabilities were found with pedestals and increments. The differences appeared to be substantial.

These changes in detectability interact with technical problems posed by energy scattering in brief, isolated, auditory signals. Together, they insure that relatively few empirical studies of genuine difference discrimination will be carried out. Increment detection is much easier to set up experimentally, and is found to be much sharper than paired comparisons of separate signals.

One of life's little paradoxes is that for years we have turned to decision theory in order to interpret increment detection although narrow-band alerting signals offer a much simpler framework. Perversely, we have generally avoided the more complex difference discrimination experiments in which decision theory may be significantly involved.

An unspoken principle of auditory detection is that the "ear" reconstructs stimulus patterns with high precision. Signal analysis and the search for optimum decision regions then become the names of the games. These quantitative delights are less frustrating than the drudgery of analyzing data propagation in auditory networks. They might also be very productive if we knew the key problems could be solved entirely within the stimulus domain. But that is the point of this paper. We do not know.

We do know that a case can be made for narrow-band alerting signals, and for balanced-lines as common carriers of sensory information. Each contributes an interesting new twist to our ideas on detection, but not quite in the stimulus domain.

REFERENCES

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

Bartlett, M. (1955). *Stochastic Processes.* Cambridge, England: Cambridge University Press.

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

Bouman, M. (1961). History and present status of quantum theory in vision." In W. Rosenblith (Ed.), *Sensory Communication* (Chapter 21, pp. 377-401.) Cambridge, MA: MIT Press and Wiley.

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

Çinlar, E. (1972). Superposition of point processes. In P. Lewis (Ed.), *Stochastic Point Processes: Statistical Analysis, Theory, and Applications*, 549-606. New York: Wiley.

Cox, D., & Smith, W. (1954). On the superpositon of renewal processes. *Biometrika*, **41**, 91-99.

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

Feller, W. (1957). *An introduction to probability theory and its applications V.1, 2nd Ed.* New York: Wiley.

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

Green, D. (1967). Additivity of masking. *Journal of the Acoustical Society of America*, **41**, 1517-1525.

Green, D. (1960). Detection of a noise signal. *Journal of the Acoustical Society of America,* **32**, 121-131.

Green, D. (1960). Psychoacoustics and detection theory. *Journal of the Acoustical Society of America,* **32**, 1189-1203.

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

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.

Jolley, L. (1961). *Summation of Series*, 2nd Ed. New York: Dover.

Kendall, D. (1948). On the generalized "birth and death" process. *Annals of Mathematical Statistics*, **19**, 1-15.

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

McGill, W. (1963). Stochastic latency mechanisms. In R. Luce, R. Bush, & E. Gadanter (Eds.) *Handbook of Mathematical Psychology* Vol. I, (Chapter 6, pp. 309-360.) New York: Wiley.

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

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

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

McGill, W., & Teich, M. (1991b). *Simple Models of Sensory Transmission*. CHIP Technical Report No. 132. Center For Human Information Processing, University of California, San Diego. 67pp.

Miller, G. (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.

Rosenblatt, M. (1962). *Random Processes*. New York: Oxford University Press.

Rushton, W. (1961). Peripheral coding in the nervous system. In W. Rosenblith (Ed.), *Sensory communication.* New York: Wiley.

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

Teich, M. (1992). Fractal neuronal firing patterns. In McKenna (Ed.) *Single Neuron Computation.* (Chapter 22, pp. 589-625). San Diego: Academic Press.

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

Teich, M., Prucual, P., Vanucci, G., Breton, M., & McGill, W. (1982). Multiplication noise in human visual system at threshold: 1. Quantum fluctuations and minimum detectable energy. *Journal of the Optical Society of America,* **72,** 419-431.

Townsend, J. (1976). Serial and within-stage parallel model equivalence on the minimum completion time. *Journal of Mathematical Psychology*, **14**, 219-238.

Viemeister, N. (1971). Intensity discrimination of pulsed sinusoids. The effect of filtered noise. *Journal of the Acoustical Society of America,* **51,** 1265-1269.

Vorberg, D., & Ulrich, R. (1987). Random search with unequal search rates: serial and parallel generalizations of McGill's model. *Journal of Mathematiccal Psychology*, **31**, 1-23.

CHIP Technical Report List

1. David M. Green and William J. McGill. *On the equivalence of detection probabilities and well known statistical quantities*. October 1969.
2. Donald A. Norman. *Comments on the information structure of memory*. October 1969.
3. Norman H. Anderson. *Functional measurement and psychophysical judgment*. October 1969.
4. James C. Shanteau. *An additive decision-making model for sequential estimation and interference judgments*. October 1969.
5. Norman H. Anderson. *Averaging model applied to the size-weight illusion*. October 1969.
6. Norman H. Anderson and James C. Shanteau. *Information integration in risky decision making*. November 1969.
7. George Mandler, Richard H. Meltzer, and Zena Pearlstone. *The structure of recognition. Effects of list tags and of acoustic and semantic confusion*. November 1969.
8. Dominic Massaro. *Perceptual processes and forgetting in memory tasks*. January 1970.
9. Daniel Graboi. *Searching for targets: The effects of specific practice*. February 1970.
10. James H. Patterson and David M. Green. *Discrimination of transient signals having identical energy spectra*. February 1970.
11. Donald A. Norman. *Remembrance of things past*. June 1970.
12. Norman H. Anderson. *Integration theory and attitude change*. August 1970.
13. A.D. Baddeley and J.R. Ecob. *Reaction time and short-term memory: A trace strength alternative to the high-speed exhaustive scanning hypothesis*. November 1970.
14. A.D. Baddeley. *Retrieval rules and semantic coding in short-term memory*. December 1970.
15. Roy D. Patterson. *Residue pitch as a function of the number and relative phase of the component sinusoids*. March 1971.
16. George Mandler and Marilyn A. Borges. *Effects of list differentiation, category membership and prior recall on recognition*. May 1971.
17. David E. Rumelhart, Peter H. Lindsay, and Donald A. Norman. *A process model for long-term memory*. May 1971.
18. David E. Rumelhart and Adele A. Abrahamsen. *Toward a theory of analogical reasoning*. July 1971.
19. Martin F. Kaplan. *How response dispositions integrate with stimulus information*. August 1971.
20. Martin F. Kaplan and Norman H. Anderson. *Comparison of information integration and reinforcement models for interpersonal attraction*. August 1971.
21. David M. Green and R. Duncan Luce. *Speed-accuracy tradeoff in auditory detection*. September 1971.
22. David E. Rumelhart. *A multicomponent theory of confusion among briefly exposed alphabetic characters*. November 1971.
23. Norman H. Anderson and Arthur J. Farkas. *New light on order effects in attitude change*. March 1972.
24. Norman H. Anderson. *Information integration theory: A brief survey*. April 1972.
25. Donald A. Norman. *Memory, knowledge, and the answering of questions*. May 1972.
26. David J. Weiss. *Averaging: An empirical validity criterion for magnitude estimation*.
    Norman H. Anderson. *Cross-task validation of functional measurement*. June 1972.
27. David E. Rumelhart and Patricia Siple. *The process of recognizing tachistoscopically presented words*. August 1972.
28. Ebbe B. Ebbesen and Richard Bowers. *The effects of proportion of risky to conservative arguments in a group discussion on risky shift*. September 1972.
29. Ebbe B. Ebbesen and Michael Haney. *Flirting with death: Variables affecting risk taking on our nation's highways*. September 1972.

30. Norman H. Anderson. *Algebraic models in perception.* November 1972.
31. Norman H. Anderson. *Cognitive algebra: Information integration applied to social attribution.* December 1972.
32. Jean M. Mandler and Nancy L. Stein. *Recall recognition of pictures by children as a function of organization and of distractor similarity.* January 1973.
33. David E. Rumelhart and Donald A. Norman. *Active semantic networks as a model of human memory.* Marc Eisenstadt and Yaakov Kareev. *Towards a model of human game playing.* June 1973.
34. George Mandler. *Memory storage and retrieval: Some limits on the reach of attention and consiousness.* July 1973.
35. Kent L. Norman. *A method of maximum likelihood estimation for stimulus integration theory.* August 1973.
36. Yaakov Kareev. *A model of human game playing.* August 1973.
37. Donald A. Norman. *Cognitive organization and learning.* August 1973.
38. The Center for Human Information Processing: A Five Year Report - 1968-73.
39. Larry D. Rosen and J. Edward Russo. *Binary processing in multi-alternative choice.* October 1973.
40. Samuel Himmelfarb and Norman H. Anderson. *Integration theory analysis of opinion attribution.* December 1973.
41. George Mandler. *Consciousness: Respectable, useful, and probably necessary.* March 1974.
42. Norman H. Anderson. *The problem of change-of-meaning.* June 1974.
43. Norman H. Anderson. *Methods for studying information integration.* June 1974.
44. Norman H. Anderson. *Basic experiments in person perception.* June 1974.
45. Norman H. Anderson. *Algebraic models for information integration.* June 1974.
46. Ebbe B. Ebbesen and Vladimir J. Konecni. *Cognitive algebra in legal decision making.* September 1974.
47. Norman H. Anderson. *Equity judgments as information integration.*
    Arthur J. Farkas and Norman H. Anderson. *Input summation and equity summation in multi-cue equity judgments.* December 1974.
48. George Mandler and Arthur Graesser II. *Dimensional analysis and the locus of organization.* January 1975.
49. James L. McClelland. *Preliminary letter identification in the perception of words and nonwords.* April 1975.
50. Donald A. Norman and Daniel G. Bobrow. *On the role of active memory processes in perception and cognition.* May 1973.
51. J. Edward Russo. *The value of unit price information. An information processing analysis of point-of-purchase decisions.* June 1975.
52. Elissa L. Newport. *Motherese: The speech of mothers to young children.* August 1975.
53. Norman H. Anderson and Cheryl C. Graesser. *An information integration analysis of attitude change in group discussion.* September 1975.
54. Lynn A. Cooper. *Demonstration of a mental analog of an external rotation.*
    Lynn A. Cooper and Peter Podgorny. *Mental transformations and visual comparison processes: Effects of complexity and similarity.* October 1975.
55. David E. Rumelhart and Andrew Ortony. *The representation of knowledge in memory.* January 1976.
56. David E. Rumelhart. *Toward an interactive model of reading.* March 1976.
57. Jean M. Mandler, Nancy S. Johnson, and Marsha DeForest. *A structural analysis of stories and their recall: From "Once upon a time" to "Happily ever after".* March 1976.
58. David E. Rumelhart. *Understanding and summarizing brief stories.* April 1976.
59. Lynn A. Cooper and Roger N. Shepard. *Transformations on representations of objects in space.* April 1976.
60. George Mandler. *Some attempts to study the rotation and reversal of integrated motor patterns.* May 1976.
61. Norman H. Anderson. *Problems in using analysis of variance in balance theory.* June 1976.
62. Norman H. Anderson. *Social perception and cognition.* July 1976.
63. David E. Rumelhart and Donald A. Norman. *Accretion, tuning and restructuring: Three modes of learning.* August 1976.
64. George Mandler. *Memory research reconsidered: A critical view of traditional methods and distinction.* September 1976.
65. Norman H. Anderson and Michael D. Klitzner. *Measurement of motivation.*
    Michael D. Klitzner and Norman H. Anderson. *Motivation x expectancy x value: A functional measurement approach.* November 1976.

66. Vladimir J. Konecni. *Some social, emotional, and cognitive determinants of aesthetic preference for melodies in complexity*. December 1976.

67. Hugh Mehan, Courtney B. Cazden, LaDonna Coles, Sue Fisher, Nick Maroules. *The social organization of classroom lessons*. December 1976.

67a. Hugh Mehan, Courtney B. Cazden, LaDonna Coles, Sue Fisher, Nick Maroules. *Appendices to the social organization of classroom lessons*. December 1976.

68. Norman H. Anderson. *Integration theory applied to cognitive responses and attitudes*. December 1976.

69. Norman H. Anderson and Diane O. Cuneo. *The height + width rule in children's judgments of quantity*.
Norman H. Anderson and Clifford H. Butzin. *Children's judgments of equity*. June 1977.

70. Donald R. Gentner and Donald A. Norman. *The FLOW tutor: Schemas for tutoring*. June 1977.

71. George Mandler. *Organization and repetition: An extension of organizational principles with special reference to rote learning*. May 1977.

72. Manuel Leon. *Coordination of intent and consequence information in children's moral judgments*. August 1977.

73. Ted Supalla and Elissa L. Newport. *How many seats in a chair? The derivation of nouns and verbs in American Sign Language*. November 1977.

74. Donald A. Norman and Daniel G. Bobrow. *Descriptions: A basis for memory acquisition and retrieval*. November 1977.

75. Michael D. Williams. *The process of retrieval from very long term memory*. September 1978.

76. Jean M. Mandler. *Categorical and schematic organization in memory*. October 1978.

77. James L. McClelland. *On time relations of mental processes: A framework for analyzing processes in cascade*. October 1978.

78. Jean M. Mandler, Sylvia Scribner, Michael Cole, and Marsha DeForest. *Cross-cultural invariance in story recall*.
Jean M. Mandler and Marsha DeForest. *Developmental invariance in story recall*. November 1978.

79. David E. Rumelhart. *Schemata: The building blocks of cognition*. December 1978.

80. Nancy S. Johnson and Jean M. Mandler. *A tale of two structures: Underlying and surface forms in stories*. January 1979.

81. David E. Rumelhart. *Analogical processes and procedural represenations*. February 1979.

82. Ross A. Bott. *A study of complex learning: Theory and methodologies*. March 1979.

83. Laboratory of Comparative Human Cognition. *Toward a unified approach to problems of culture and cognition*. May 1979.

84. George Mandler and Lawrence W. Barsalou. *Steady state memory: What does the one-shot experiment assess?* May 1979.

85. Norman H. Anderson. *Introduction to cognitive algebra*. June 1979.

86. Michael Cole, Edwin Hutchins, James Levin, and Naomi Miyake, Editors. *Naturalistic problem solving and microcomputers. Report of a Conference*. June 1979.

87. Donald A. Norman. *Twelve issues for cognitive science*. October 1979.

88. Donald A. Norman. *Slips of the mind and an outline for a theory of action*. November 1979.

89. The Center for Human Information Processing: A Description and a Five-Year Report (1974-79). November 1979.

90. Michael Cole and Peg Griffin. *Cultural amplifiers reconsidered*. December 1979.

91. James L. McClelland and David E. Rumelhart. *An interactive activation model of the effect of context in perception. Part I*. April 1980.

92. James L. McClelland and J.K. O'Regan. *The role of expectations in the use of peripheral visual information in reading*. February 1980.

93. Edwin Hutchins. *Conceptual structures of Caroline Island navigation*. May 1980.

94. Friedrich Wilkening and Norman H. Anderson. *Comparison of two rule assessment methodologies for studying cognitive development*. June 1980.

95. David E. Rumelhart and James L. McClelland. *An interactive activation model of the effect of context in perception. Part II*. August 1980.

96. Jean M. Mandler. *Structural invariants in development.* September 1980.
97. David E. Rumelhart and Donald A. Norman. *Analogical processes in learning.* October 1980.
98. James A. Levin and Yaakov Kareev. *Personal computers and education: The challenge to schools.* November 1980.
99. Donald A. Norman and Tim Shallice. *Attention to action: Willed and automatic control of behavior.* December 1980.
100. David E. Rumelhart. *Understanding understanding.* January 1981.
101. George Mandler. *The structure of value: Accounting for taste.* May 1981.
102. David E. Rumelhart and Donald A. Norman. *Simulating a skilled typist: A study of skilled cognitive-motor performance.* May 1981.
103. Jean M. Mandler. *Representation.* June 1981.
104. Donald R. Gentner. *Skilled finger movements in typing.* July 1981.
105. Edwin L. Hutchins and James A. Levin. *Point of view in problem solving.* August 1981.
106. Michael Cole. *Society, mind and development.*
     Michael Cole. *The zone of proximal development: Where culture and cognition create each other.* September 1981.
107. Laboratory of Comparative Human Cognition. *Culture and cognitive development.* November 1981.
108. Donald R. Gentner. *Evidence against a central control model of timing in typing.* December 1981.
109. Robert M. Boynton and Allen L. Nagy. *The La Jolla analytic colorimeter: Optics, calibrations, procedures and control experiments.* December 1981.
110. Jonathan T. Grudin and Serge Larochelle. *Digraph frequency effects in skilled typing.* February 1982.
111. Dimensional analysis and the locus of organization. Donald R. Gentner, Jonathan T. Grudin, Serge Larochelle, Donald A. Norman, David E. Rumelhart. *Studies of typing from the LNR Research Group.* September 1982.
112. Donald A. Norman. *Five papers on human-machine interaction.* May 1982.
113. Naomi Miyake. *Constructive interaction.* June 1982.
114. Donald R. Gentner. *The development of typewriting skill.* September 1982.
115. Margaret M. Riel. *Investigating the system of development: The skills and abilities of dysphasic children.* February 1983.
116. David E. Rumelhart and Donald A. Norman. *Representation in memory.* June 1983.
117. George Mandler. *Consciousness: Its function and construction.* June 1983.
118. Norman H. Anderson. *Schemas in person cognition.* September 1983.
119. Norman H. Anderson. *A theory of stereotypes.* December 1983.
120. Norman H. Anderson. *Psychodynamics of everyday life: Blaming and avoiding blame.* December 1983.
121. Donald R. Gentner. *Expertise in typewriting.* May 1984.
122. Alessandro Duranti. *Intentions, self, and local theories of meaning: Words and social action in a Samoan context.* June 1984.
123. Geoffrey Hinton and Paul Smolensky. *Parallel computation and the mass-spring model of motor control.* June 1984.
124. Donald R. Gentner. *Skilled motor performance at variable rates: A composite view of motor control.* October 1985.
125. George Mandler, Carmen Overson, Peter Graf, and Dolores Kraft. *Activation and elaboration: Indirect sources and effects.* November 1985.
126. Geoffrey M. Boynton. *Documentation of computer programs for experiments on categorical color perception.* September 1987.
127. Conrad X. Olson. *A shortcut method for naming OSA color specimens.* March 1988.
128. John S. Houston and Pamela A. Sample. *Programmable automated perimetry: UCSD user's guide.* April 1988.
129. William J. McGill and Malvin C. Teich. *Auditory signal detection and amplification in a neural transmission network.* May 1989.
130. Yrjö Engeström. *Developing thinking at the changing workplace: Toward a redefinition of expertise.* October 1989.
131. Michael Cole. *Cultural psychology: A once and future discipline?* January 1990.
132. William J. McGill and Malvin C. Teich. *Simple models of sensory transmission.* January 1991.
133. Harold Pashler. *Dual-task interference and elementary mental mechanisms.* February 1991.
134. William J. McGill and Malvin C. Teich. *Alerting signals and auditory detection in branching chains.* December 1992.