

## Loudness and the Power Series Transformation

### Preliminary Model and Measurements

T. D. Clack

Kresge Hearing Research Institute, University of Michigan, Ann Arbor, Mich., USA

**Abstract.** A nonlinear-algebraic approach to monaural intensity processing is proposed: the purpose is to integrate the conscious loudness attribute evoked by a pure tone with the power-series description of auditory distortion. The preliminary model consists of six postulates leading to a mathematical definition for loudness. To evaluate this equation, two task-specific sets of loudness judgments from each of seven subjects are examined. When linked to loudness-interval responses via the equisection assumption, the equation describes the behaviors quite well. Extrapolations into other intensity ranges predict similar responses with relatively slight overestimates. By relaxing the assumption that subjects will adjust the loudness proportions exactly as instructed, their ratio productions also can be described and predicted with sometimes surprising accuracy. Particularly striking is the prediction of overall levels and the curvilinearities of “doublings” from the “halvings”. In addition, the theory proposes absolute loudness measurement, an explanation for the growth of loudness including the principle underlying Steven’s Power Law, and might prove useful in examining some exceptions to this relationship. Several aspects of this model differ from traditional approaches to intensity processing, but it appears to warrant further critical evaluations.

### Introduction

Loudness is an obvious attribute of sensation when one ear is stimulated with a super-threshold pure tone of varied intensity. The listener is usually unaware that the auditory response also involves the generation of a harmonic series—the *subjective* overtones. Nonetheless, both the production of these aural harmonics and the loudness are integral aspects of the intensity processing within the system. Their interrelationship seems probable.

The influence of acoustic harmonics on loudness has been evident for some time. The first experimental demonstrations presented a pure-tone fundamental ( $f_1$ ) of about 100 Hz at 104 dB *SPL* while the phase of one of its initial four harmonics ( $2f_1 - 5f_1$ ) was shifted (Chapin and Firestone, 1934; Trimmer and Firestone, 1937). Marked loudness variations were found: as the phase of the  $2f_1$  was shifted, for example, the loudness of  $f_1$  could change by at least 10 dB. Loudness-phase interactions also have been reported by some listeners during measurements of  $2f_1$  thresholds with the 1000-Hz ( $f_1$ ) at 55–75 dB *SPL* (Clack, 1967, 1968).

Such observations are intriguing for at least two reasons. In the latter instances, the frequency separations are clearly greater than the critical band. Thus, the interactions apparently can occur regardless of the frequency differences. Second, the levels of the harmonics in all these experiments were from –20 to –50 dB with respect to the  $f_1$  intensities. With such relatively low-level harmonics, shifts in their phases would induce trivial changes in the amplitudes of the  $f_1$  waveforms. Thus, direct interference between the objective tones seems an unlikely cause for the findings.

Although aural harmonic amplitudes can approximate the levels of the acoustic harmonics with respect to the  $f_1$  intensities, the evidence linking them to loudness is less easily interpretable. Abnormal loudness growth (clinical loudness recruitment) has been correlated with decreases in the signal intensities at which aural harmonics are detected via the best-beat method (Opheim and Flottorp, 1951; Lawrence, 1958). These findings have equivocal meaning, however, because of a bias in the method (Egan and Klumpp, 1951; Clack and Bess, 1969). Other procedural weaknesses have thwarted attempts to establish a correlation in normal listeners (Clack and Erdreich, 1972; Erdreich and Clack, 1973).

The whole issue of how harmonic structure is related to loudness has been effectively ignored by

most psychophysicists. Wever (1929) observed, however, that the loudness-growth function co-varies with the relative amplitudes of aural harmonics [the electrophysiological (*CM*) phenomena, not the behavioral]. He hypothesized an additive interaction by assuming an integrator within the retrocochlear nervous system. Thus, the relation of loudness to signal intensity would be a function of the total internal stimulus complex. This notion has never been detailed sufficiently for testing, but much might be learned from further examination of this basic approach. To explore the summator idea requires several assumptions about underlying processing—six constitute the model proposed here.

### A Preliminary Nonlinear-Algebraic Model

To develop a preliminary thesis, we start with the description of aural harmonic generation (Clack et al., 1972; Clack, 1975). Thus, the initial postulate is a conversion of a monaural pure-tone input ( $s = \alpha \cos \omega t$ , where  $\omega = 2\pi f_1$ ) according to the power series—

$$S = \sum_{k=0}^{\infty} a_k s^k = a_0 + a_1 s + a_2 s^2 + a_3 s^3 + a_4 s^4 + \dots \quad (1)$$

This is a general nonlinear equation since the output ( $S$ ) is not a linear function of the input ( $s$ ), and was first applied to the auditory transformation by Fletcher (1929). Stevens and Davis (1938) later presented an electronic analogue that is helpful in conceptualizing how such an operator describes the transformation of a pure tone into a harmonic series (or vice versa).

Equation (1) can be simplified in a couple of ways before we proceed. First, as far as a sensation is concerned, it can be assumed that  $a_0 \approx 0$ , i.e., that without input ( $s$ ) there is no output ( $S$ ).

Description of the nonlinearity by  $\sum_{k=1}^{\infty} a_k s^k$  is compatible with existing evidence about aural harmonic behaviors. Second, the fully expanded equation can be delimited: here consideration will utilize terms only up to the cubic. This truncation is presently forced because the available measurements include only determinations of second and third aural harmonic amplitudes—the respective products of the quadratic ( $a_2 s^2$ ) and cubic ( $a_3 s^3$ ) terms. Thus, (1) is represented as a partial sum ( $S_p$ ):

$$S_p = a_1 s + a_2 s^2 + a_3 s^3. \quad (2)$$

To interconnect this truncated expression with the final loudness sensation, of course, requires several further presumptive hypotheses. One is that the effect of an input is not associated with its amplitude ( $\alpha$ ) until  $\alpha$  attains some minimal magnitude—the level of the absolute limen ( $\lambda_1$ ). For a signal with period equivalent to  $f_1$ , the effective level ( $\beta$ ) will be defined as  $\alpha - \lambda_1$ , where  $\lambda_1$  is the threshold of the  $f_1$  frequency. Substitution and expansion now yields

$$S_p = a_1 \beta \cos \omega t + \frac{1}{2} a_2 \beta^2 + \frac{1}{2} a_2 \beta^2 \cos(2\omega)t + \frac{3}{4} a_3 \beta^3 \cos \omega t + \frac{1}{4} a_3 \beta^3 \cos(3\omega)t. \quad (3)$$

This second premise that periodic components must attain their respective thresholds ( $\lambda_1$ ) to make a non-zero contribution to the sum has two important aspects. First, notice that  $\beta$  is unlike the usual construct of sensation level (SL); the common practice of subtracting dB to obtain SL is equivalent to taking the ratio ( $\alpha/\lambda_1$ ) in linear measure<sup>1</sup>. Second, this postulate provides a basis for disregarding inaudible signal components. Thus, masked aural harmonics ( $\alpha \lesssim 90$  dB SPL) can be considered *blocked*, or otherwise

<sup>1</sup> All computations are carried out in dynes although the results are presented in dB equivalents to conform with traditional presentations

excluded from entry into further processing. Thus, (3) is reduced to only three remaining terms:

$$\frac{1}{2} a_2 \beta^2 + a_1 \beta \cos \omega t + \frac{3}{4} a_3 \beta^3 \cos \omega t.$$

The next major postulate is that components with equal periods combine, while those with unequal periods are differentially processed. The amplitudes of the first-order term ( $a_1 \beta$ ) and the contribution of the cubic term ( $3/4 a_3 \beta^3$ ) with the same frequency ( $f_1$ ), therefore, combine: this combination might be called the first-order dimension to differentiate it from a simple component.

The operation of a compressional transformation is usually assumed in rationalizing the ability of a sensory system for response over wide stimulus-amplitude ranges. Fechner, for instance, chose a logarithmic transform (Boring, 1956)—the mathematical equivalent of a specific power series. The general form of such a compression has the *even* terms negative, and the *odd* terms positive. Similarly, the present formulation will designate that the coefficients of the even-ordered terms are negative, and the odd-ordered are positive (this *compressional* postulate also is necessary because the parameter-estimation method yields only positive values for the  $a_k$ ).

Incorporation of these last two hypotheses produces a reformulation of the remaining terms as—

$$-\frac{1}{2} a_2 \beta^2 + [a_1 \beta + \frac{3}{4} a_3 \beta^3] \cos \omega t.$$

This expression contains a cosine factor, a ratio that varies over time ( $t$ ). The hearing of a tone, however, approximates that of a steady-state signal, at least for signals of sufficient duration ( $t \rightarrow \infty$ ). This implies some kind of AC-to-DC-like conversion, and a *period-averaging* postulate is necessary. The idea to be incorporated for now is analogous to the way that sinusoids are often resolved in conversion situations, e.g., use of rms to describe peak amplitudes. In such circumstances, the usual mathematical manipulation is the equivalent of multiplying amplitude by a constant, e.g., 0.707. Because of commutativity, the same end can be achieved through multiplying the cosine ratio by some factor to arrive at a constant value ( $C$ ).

In effect, then, the periodic components are weighted relative to the aperiodic through this multiplication by  $C$ . Such mathematical maneuvers imply underlying structural-functional mechanisms of considerable complexity involving attenuations and/or amplifications, etc. To guess at the nature of these processes seems presently unnecessary, except to indicate that at least two cues would have to be encoded. This seems to be the way the periphery operates (Zwislocki and Sokolich, 1973). With respect to the constraints on  $C$  or the other parameters, however, intuitions fail. Thus, an empirical attitude is the only reasonable alternative for the time being even though the mathematical development suffers some inelegance.

The final major premise defines the sensation of loudness mathematically. This “loudness-magnitude” postulate states that the sum ( $S_p$ ) of the remaining series of components is monotonically transformed by what will be assumed here to be a simple linear operator ( $\Psi$ ). Further, it will be presumed that loudness ( $L$ ) is equal to  $\Psi(S_p)$  where  $\Psi$  is a straight line with slope = 1 and an intercept at 0. Thus, “the loudness equation” becomes

$$L = -\frac{1}{2} a_2 \beta^2 + [a_1 \beta + \frac{3}{4} a_3 \beta^3] C. \quad (4)$$

With this derivation, the algebraic solution yields units isomorphic with the units of loudness measurement<sup>2</sup>. The dynamic range of loudness is expressed by the transformation of acoustic pressures given in (4)—or more accurately, in the extended form of which (4) is

<sup>2</sup> An absolute loudness function is defined by this equation since it incorporates a zero ( $L=0$  when  $\alpha/\lambda_1$ ) as well as units given by algebraic solution:  $a_2$  is in  $\text{dyn}^{-1}$  and  $a_3$  in  $\text{dyn}^{-3}$  when estimated from tone-on-tone masking measurements (Clack, 1975). Equation (4) yields a summation in dynes which is converted into psychological equivalents through the  $\Psi$  conversion. The units of measured loudness will be called “ $S$ ”

but a truncated version. Also because  $\Psi$  is assumed to be a linear function, the implication is that the initial power-series transformation is the only nonlinear stage in the auditory processing of intensity, and that the rest of the system weights the first-order dimension, then acts as a summator.

### Experimental Evaluations

Although the presumptions of this model (or any other such postulate set) might not be assessable directly, their consequence, (4), is testable once the coordinating definitions for the parameters are developed. Applications of the tone-on-tone masking procedures (Clack, *op. cit.*; Erdreich and Clack, 1972) enables estimation of

$\lambda_1$  = absolute threshold for the acoustic pure tone ( $f_1$ ),

$a_2$  = Coefficient of the quadratic term in the power series, the quadratic coefficient,

$a_3$  = Coefficient of the cubic term in the power series, the cubic coefficient.

A completely different method must be devised, however, to find values for

$a_1$  = Coefficient of the first-order (linear) term in the power series, the linear coefficient,

C = the weighting factor.

The tactic to be employed here involves first estimating  $a_2$ ,  $a_3$ , and  $\lambda_1$ , and then fitting (4) to sets of loudness measurements to obtain optimal values for  $a_1$  and C.

There are several techniques suitable for obtaining loudness responses useful for such a fitting. Interval judgments have been chosen for two main reasons. One is that these behaviors appear relatively impervious to influences from other psychophysical tasks (Fagot et al., 1966). The second is that performance makes minimal demands of the subject: listeners need only discriminate whether the successively heard loudness intervals are smaller, equal to, or are larger than one-another (Torgerson, 1958). Of all the direct-loudness tasks attempted in this laboratory the listeners seem able to equate loudness intervals with the least instructions, and with the greatest expression of confidence that they had accomplished what they were asked.

Once all the parameters of (4) are obtained, its evaluation can proceed. To begin this testing, (4) can be fit to only a very limited range of interval judgement behavior, and then extrapolated to predict the responses of other segments. This yields a relatively weak assessment of course, since the behaviors described and predicted are so similar. A more rigorous approach would involve a test of the ratio properties inherent in the provisions of the zero and units of (4). Ratio production behaviors, therefore, will be analyzed and compared to predictions. To make these evaluations as stringent as possible, the behaviors of individuals will be examined.

### Procedure

#### Subjects

College students (ages 19–26) were paid as listeners. Each was given a screening sweep-frequency, Békésy-type audiogram before and after the experiment. The ears chosen had thresholds within 15 dB of audiometric zero (I.S.O., 1964 between 500 and 1500 Hz (except for L.R. whose 1500-Hz threshold showed an abrupt notch to +20 dB). None reported any history of hearing problems nor were they taking any medications worthy of remark. No significant changes in hearing were detected within the duration of this experiment.

#### Tone-on-Tone Masking

The tone-on-tone masking technique and analysis have been detailed previously, and need not be recounted here. Since the information about the aural harmonics of 500 Hz ( $=f_1$ ) was relatively scant, this frequency was chosen for these studies, and these results have been separately reported (Clack, *in press*).

#### Interval Equisection

Each listener sat before a table with a rectangular box on top. The face of this box contained five toggle switches in an upper row and three identical knobs directly below the middle three. Pressing each switch presented the 500-Hz tone and releasing it turned the tone off. Each correspondingly positioned knob controlled a continuously variable dB attenuator (with a maximum level randomly varied by a few dB from trial to trial). Subjects were instructed as follows:

By depressing one of the switches on the panel before you, you can listen to each of a series of five tones. The loudest tone is selected by depressing the switch farthest to the right and the softest tone is selected by depressing the switch farthest to the left. The loudness of the three middle tones can be varied with the controls. By turning the knob clockwise the loudness of the tone is increased and counter-clockwise makes the tone softer.

Your task is to set these three middle tones so that all five form a series of equal loudness increments. This can be done most easily by first setting the middle tone to lie halfway between the loudest and softest tone. Then, using the middle tone as a reference, set the tone between it and the softest tone to lie halfway between these two tones. Similarly, set the right hand adjustable tone to lie between the loudest tone and the middle tone. Before making your final judgement, listen to all five tones in the series in both descending and in ascending order several times to make any further adjustments you think may be necessary. When you feel that the tones make a series of equal loudness steps, raise your hand.

After a short practice, each set of recorded responses was obtained on a different run in which the intensity references were 35–60, 50–75, and 65–90 dB SPL delimiting the low, medium and high ranges, respectively.

**Table 1.** Test-retest estimations for the quadratic ( $a_2$ ) and the cubic ( $a_3$ ) coefficients together with the thresholds ( $\lambda_1$ ) of the 500-Hz tone for each of seven listeners: also presented are the means ( $\bar{a}_2$  and  $\bar{a}_3$ ), standard deviations ( $Sa_2$  and  $Sa_3$ ), and standard errors ( $S\bar{a}_2$  and  $S\bar{a}_3$ )

		1000 Hz							1500 Hz						
		60	65	70	75	$\bar{a}_2$	$Sa_2$	$S\bar{a}_2$	65	70	75	$\bar{a}_3$	$Sa_3$	$S\bar{a}_3$	$\lambda_1(SPL)$
L.R.	(1)	0.064	0.029	0.027	0.021	0.038	0.016	0.006	0.057	0.068	0.083	0.059	0.015	0.006	12.8
	(2)	0.059	0.035	0.024	0.043				0.054	0.045	0.044				
R.S.	(1)	0.028	0.014	0.105	0.019	0.038	0.030	0.011	0.018	0.016	0.016	0.020	0.009	0.004	5.3
	(2)	0.041	0.012	0.033	0.051				0.035	0.026	0.009				
C.B.	(1)	0.010	0.018	0.049	—	0.021	0.016	0.007	0.014	0.018	0.024	0.023	0.006	0.002	6.8
	(2)	0.014	0.014	—	—				0.029	0.020	0.029				
H.S.	(1)	0.027	0.028	0.025	0.025	0.030	0.006	0.002	0.023	0.031	0.024	0.026	0.004	0.001	12.5
	(2)	0.035	0.040	0.034	0.027				0.029	0.027	0.022				
A.D.	(1)	0.073	0.106	0.024	0.030	0.069	0.034	0.010	0.035	0.012	0.042	0.032	0.012	0.005	12.1
	(2)	0.046	0.086	0.024	0.045				0.033	—	—				
A.M.	(1)	0.025	0.036	—	—	0.029	0.009	0.004	0.007	0.007	0.009	0.010	0.006	0.003	13.9
	(2)	0.037	0.016	—	0.028				0.020	0.005	0.008				
K.A.	(1)	0.027	0.053	0.015	0.027	0.024	0.013	0.004	0.037	0.033	0.010	0.027	0.012	0.005	10.2
	(2)	0.024	0.021	0.013	0.015				0.039	0.019	0.023				

### Ratio Production

Pairs of 500-ms tone bursts were presented with the reference burst ( $S_r$ ) fixed at 10 dB increments from 30 to 90 dB SPL (38–98 dB SPL for one person). These  $S_r$  levels were presented in nonsystematic order within each run. The other tone-burst ( $S_v$ ) could be varied continuously by turning a knob hidden within a metal sleeve and attached to the listener's chair:  $S_v$  was variable over 50 dB with 3/4 turn lock-to-lock. Each ordered pair (i.e.,  $S_r:S_v$  or  $S_v:S_r$ ) was presented twice and each of the four  $S_v$  levels recorded whenever the subject signaled satisfaction with the adjustment.

The listeners were instructed to adjust  $S_v$  to a level of *one-half* or *double* the loudness of  $S_r$  in each run. These directions were elaborated as little as possible to allow each listener to operate on his/her own definitions. Practice was given until the listeners began to feel comfortable in their judgements and/or the data began to show some reliability.

### Design

The order of methods was unsystematic although the same task might be run one or more times during a session with rests and changes in earphone placements between runs.

### Results

For the seven listeners who completed the experiment, the estimates of the nonlinear coefficients ( $a_2$  and  $a_3$ ) are presented fully in Table 1. These values differ

slightly from those previously reported because the quantity  $\beta$  was used in the definitions of the aural harmonic amplitudes [see Equation (3) vs. the definitions in Clack, in press]. Since the standard deviations ( $Sa_2$  and  $Sa_3$ ) are usually small, the coefficients for each of the separate  $f_1$  intensities have been averaged to yield means ( $\bar{a}_2$  and  $\bar{a}_3$ ).

### Analysis of Loudness Interval Judgments

With  $\bar{a}_2$  and  $\bar{a}_3$  inserted, guessed values for  $a_1$  and  $C$  were substituted into (4) and  $L$ 's for  $\alpha$  at 65 and 90 dB SPL were calculated. On the assumption that the mathematical resultant loudness experiences ( $L_{90\text{dB}}$  minus  $L_{65\text{dB}}$ ) was divided into four equal segments (the equisection assumption), the differences in  $L$ 's was similarly divided and each of the corresponding equations then solved for the appropriate root. These three hypothetical responses ( $R_e$ ) could then be compared with the obtained settings ( $R_0$ ) using normalized sum of squared residuals ( $SSR_N$ ) defined as

$$SSR_N = \sum_1^3 \frac{[R_e - R_0]^2}{R_0^2}.$$

The  $a_1$  and/or  $C$  were then "corrected" and the whole process repeated until a minimal  $SSR_N$  was found.

The measured and theoretical responses are displayed at the top of each graph in Figure 1. The fit to these high-range data shown by connecting the  $R_e$  with the solid lines is obviously quite good. The corresponding  $a_1$  and  $C$  values, therefore, are given in Table 2. The dashed lines of Figure 1 are extrapolations of (4)

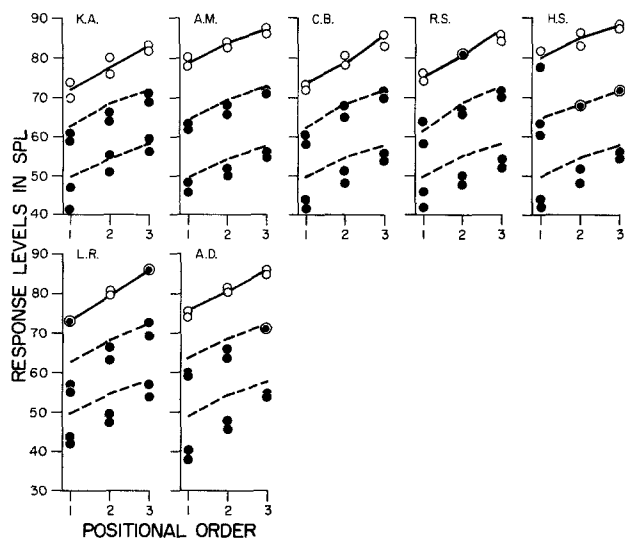


Fig. 1. Test-retest intensities adjusted by each of seven listeners instructed to equisection three 25 dB ranges into four intervals: solid line segments are fits of (4), and the dashed segments are extrapolations to the medium and low intensity ranges

Table 2.. Values of the first-order coefficients ( $a_1$ ), and weighting factors ( $C$ ) together with the normalized sum of squared residuals ( $SSR_N$ ) for each of the seven listeners

	L.R.	R.S.	C.B.	H.S.	A.D.	A.M.	K.A.
$a_1$	2.301	2.531	0.900	12.450	2.219	2.925	0.950
$c$	0.036	0.039	0.050	1.000	0.098	0.105	0.050
$SSR_N$	0.024	0.059	0.028	0.003	0.005	0.015	0.018

using the parameters of Tables 1 and 2. The medium (50–75 dB) ranges are predicted with some tendency to overestimation. The overestimation becomes even more evident in the lower-range (35–60 dB) *extrapolations* for most of the subjects. The magnitudes of these errors, however, can be given some prospective by considering the largest, i.e., 10.5 dB for A.D. in the lowest setting. The obtained  $\alpha$  was 0.0168 dyn and the predicted was 0.0568 dyn. This discrepancy (0.04 dyn) seems slight especially when it is remembered that the closest value used in the extrapolation (74.5 dB SPL = 1.0618 dyn) is 63.2 times greater than the value predicted. By such standards (4) seems to present at least a reasonable good basis on which to describe and predict equisection behavior.

#### Analysis of Ratio Productions

When instructed to produce loudness ratios, these listeners generated the results averaged in Table 3. The “halvings” generated dB-averaged decrements which vary with reference intensity: Starting relatively small (–8.9 dB) at 30 dB SPL, growing larger (–17.4 dB) at

60 dB SPL, and then decreasing again (–16.1 dB) at 80 dB SPL. The *doublings* exhibit a different trend and are unequal in magnitudes at comparable levels, e.g., 17.3 dB at 30 dB SPL and 7.8 dB at 80 dB SPL. Such observations have been reported before and these data are within the ranges of results found by others (Stevens, 1955).

With the parameters already determined from the equisection analysis, the loudness experiences defined by (4) may be assumed to be what each of the listeners adjusts. On this basis, the loudness proportions ( $p$ ) actually set by the subjects may be determined empirically. To find these  $p$ 's, the  $L$ 's for each of the reference SPL's is first calculated. Starting with a guessed proportion ( $pL$ ), the set of corresponding equations are then solved for the appropriate root. These guessed SPL's were next compared with the obtained responses (median  $R_0$ 's were used here) by again defining an  $SSR_N$ . A new value for  $p$  could then be substituted and the whole process repeated until a minimal  $SSR_N$  was attained.

Only the data from the *halvings* were thus fit. The obtained values for  $p$  and the accuracy of the fits are shown in the lower parts of Figure 2. The actual proportions derived are all less than 0.5, the ratios which would be expected if the listeners had followed the instructions exactly. Reasons for these differences could be several, but are irrelevant for present purposes. The significant observation is that the changes in the sizes of the decrements and the curvilinearities are often very well matched by the theoretical functions, especially for C.B., L.R., A.D., A.M., and R.S.

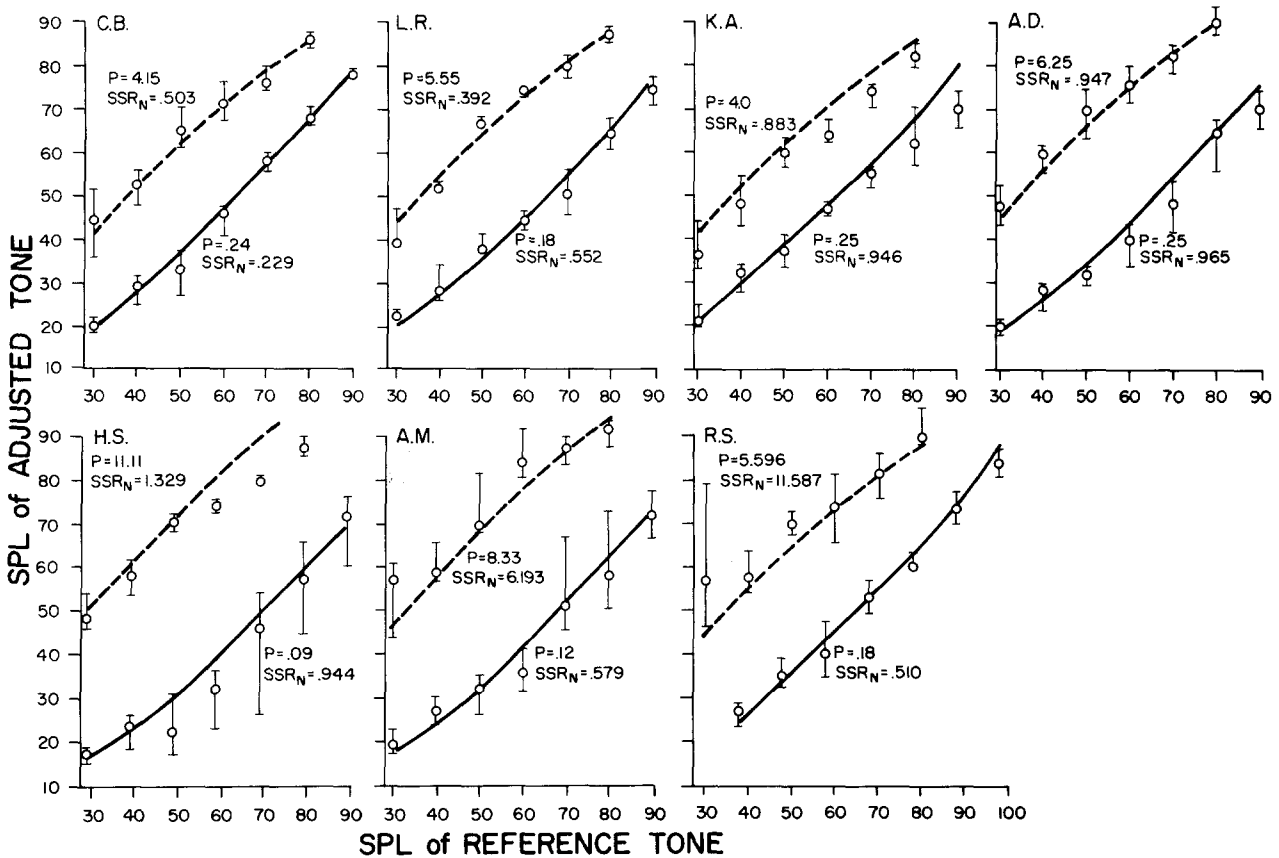
Solving for  $1/p$ , the results in the reciprocal task, “doubling”, can then be predicted for each of the listeners. The predicted and obtained results are displayed in the upper parts of Figure 2. Only one obvious departure from predictions is evident. The “doubling” responses of H.S. are fit well at the lower three intensities but at the higher levels she consistently set lower values than predicted—as much as 10.5 dB, the equivalent of 11.1 dyn which represents a significant theoretical overestimation. For the other six subjects, the predicted *doublings* fit with often startling precision, e.g., C.B., L.R., and A.D. In fact, the responses of subjects who are least variable in their ratio productions appear to be predicted with the greatest accuracy.

#### Discussion

The proposed integration of loudness with power-series nonlinearity yields information relevant to both aspects of intensity processing. A separation of these topics is artificial, but expositionally convenient even within such a framework.

**Table 3.** Mean half- and double-loudness adjustments for seven listeners who varied the intensity of a 500-Hz tone-burst with reference to the same signal fixed at 10 dB (SPL) levels

Production	Reference tone level (dB)							Average difference
	30	40	50	60	70	80	90	
"Halving"	- 8.9	-10.9	-16.5	-17.4	-17.3	16.1	-15.6	-14.7
"Doubling"	+17.3	+15.3	+17.5	+14.0	+10.3	+ 7.8	—	13.7



**Fig. 2.** Range and median ratio productions for each of seven listeners. The solid curves are fits of (4) to halvings and the dashed are predictions of doublings: the empirically determined response proportions (P) and the normalized sums of squared residuals (SSR<sub>N</sub>) are also shown

*The Nonlinear Transformation*

An adequate operational description of an auditory channel (an ear) would require determination of the relative weights ( $a_k$ ) of any internal conversion of the signal. The magnitudes of the higher-order ( $k > 1$ ) components can be estimated using methods based on tone-on-tone masking measurements. Now the development of the nonlinear-algebraic approach has suggested a procedure for evaluating the effective magnitudes of the first-order ( $a_1$ ) components.

Present findings indicate that these  $a_1$ 's can differ considerably from one ear to another. Some ears

apparently convert more of the signal amplitude directly into the hearing of the tone than do others. Since these  $a_1$ 's evidently vary also with respect to the  $a_2$ 's and  $a_3$ 's, there appear to be rather large differences in fidelity. Such results are consonant with our intuitions: most listeners with normal hearing, for example, report that a pure tone sounds pure. Furthermore, throughout the history of studying the behaviors of subjective tones, some listeners have heard them easily while others have not (Lawrence, op. cit.; Plomp, 1967). Thus, the development of a method for estimating  $a_1$  could represent an important step towards obtaining a full behavioral description of auditory distortion.

It would be premature to interpret the present  $a_1$  estimates too seriously since there are several methodological and theoretically substantive questions remaining to be explored. One example is the problem of which loudness responses to use in estimating  $a_1$ . There are various alternatives, e.g., the ratio productions instead of the interval behaviors. The answers vary somewhat depending on the choice of data. When the ratio productions are used in the fits, for instance, it is necessary to simultaneously estimate three parameters:  $a_1$ ,  $C$ , and the loudness proportion ( $p$ ) actually adjusted by the subject. Because the main present purpose is only to introduce the nonlinear-algebraic approach as an alternative to the traditional way of viewing loudness, a consideration of such matters seems inappropriate here. Nonetheless, the problem of which data will have to be resolved somehow if this approach is to attain its full potential descriptively.

Furthermore, the data for this report were gathered for other purposes leaving much to be desired. In making their loudness-interval responses, for example, it has been assumed that the listeners judged sense differences, and that they counterbalanced directional (hysteresis) biases as they were instructed. The ability of the model to fit the data suggests that these assumptions are reasonable to a first approximation. But it is also obvious that such assumptions need some kind of verification before the computed  $a_1$ 's can be accepted with any real confidence. The major contribution of this report, therefore, is not the exact values of the  $a_1$ 's found, but is the offering of techniques for their computation where none had been suggested before.

### Loudness

Coordinated by the assumption that the listeners equisect sense distances, the loudness-intervals are not only fit reasonably well by the nonlinear-algebraic model, but the extrapolations show relatively small deviations from the data. Whether these overestimation errors might be reduced if an expanded version of the truncated equation were used, remains to be seen. With the ratio productions, the coordinating assumption is that the subjects adjust their defined loudness experiences consistently to some proportion ( $p$ ), although not necessarily to that of the instructions. This presumes considerably less of the listener than is required in the more traditional scaling approach, i.e., the assumption that the listener performs as instructed. In fact, the present analysis only assumes that the listener can settle at adjustments of some interval which can then be transposed over a range of intensities. Even with this "weaker" assumption, and the parameters obtained from the loudness-interval data, the ratio

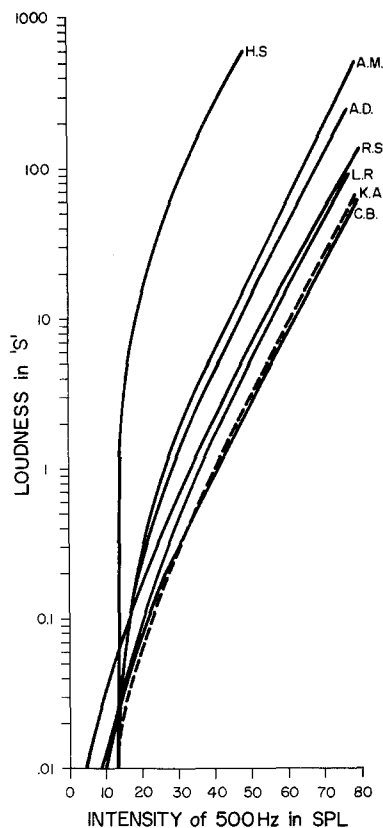


Fig. 3. Absolute loudness functions derived with (4) for each of the seven listeners;  $S=1000 \times L$  for illustrative convenience (also in Fig. 4)

productions appear described and predicted with often surprising accuracy. Particularly striking is the prediction of the curvilinearities as functions of the reference intensities. Thus, the mathematical definition appears to serve well as an intervening variable mediating the task-specific behaviors observed in interval and ratio productions.

There are other criteria that could be applied in trying to assess the worth of the preliminary nonlinear-algebraic model. One concerns the shape of the loudness functions that supposedly perform the mediations. Figure 3 displays such curves for each of the individuals sampled. The coordinates are log-log to facilitate comparisons with findings of previous investigations. There are several features of these functions that deserve attention.

The differences between the curves for individuals is one such feature. The absolute thresholds and the three other parameters of (4) are responsible: in order of importance,  $a_1$ ,  $C$ , and then the nonlinear coefficients,  $a_2$  and  $a_3$ . In the case of H.S., for example, the high  $a_1$  ( $=12.4$ ) appears mostly responsible for her unusually fast-rising function, e.g., a 40 dB tone is about 121 times louder to H.S. ( $S=0.0485$ ) than to

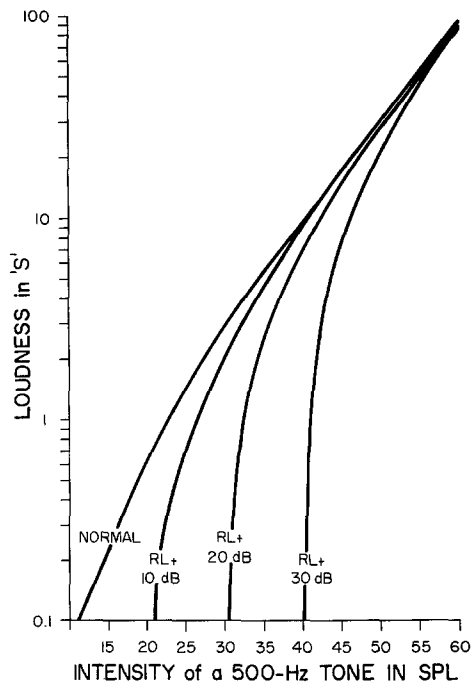


Fig. 4. Derived loudness functions for K.A. illustrating predicted changes accompanying various simulated hearing losses

C.B. (with  $a_1 = 0.9$ , and  $S = 0.0004$ )<sup>3</sup>. This very rapid growth in loudness deduced from analysis of the equisections of H.S. seems compatible with the observed aberrations in her ratio productions: it is not surprising that her *doublings* were adjusted to lower and lower relative levels as the reference intensities were increased to beyond 60 dB since loud pure tones are annoying to most of us, and tend to be avoided whenever possible.

Another obvious feature of these functions is their shape: each shows that loudness grows very rapidly at near threshold intensities, and then decelerates to nearly a constant growth rate once the level exceeds about 30–40 dB. Inspection of these *linear* segments reveals that the nonlinear-algebraic definition predicts a 2:1 growth in the loudness sensation with each increment of about 6 dB. This rate differs from the 10 dB/2:1 promoted so persuasively by Stevens and colleagues. In fact, the 6 dB rate has been suggested before (Warren, 1958). Stevens (1963) marshalled three arguments for its rejection. One was that its basis, the *Physical-Correlate Theory*, ignored the *biological make-up of the sense organ*. This point is irrelevant to the nonlinear-algebraic approach since the 6 dB rate is the consequence of presumed steps in the auditory processing. Second, Stevens argued that most scaling

<sup>3</sup> Depending upon the number of units ( $S$ ) within the dynamic range of hearing, this factor may represent a relatively small or large loudness difference

experiments yield lower slopes. His conclusion had two bases: the first was his weighted averaging of the ratio productions taken from the work of many investigators. These results show 2:1 ranges from about 3–20 dB, and some might argue that with such variability, the average is meaningless. His other basis was the results from the magnitude estimation scales available at the time. When certain methodological biases are minimized, however, averages over large samples of subjects can be very close to the 6 dB rate (Warren, 1970, 1973). Even without such controls, subsequent studies have demonstrated that the 6 dB rate is within the ranges of individual results J. C. Stevens and Guirao, 1964). Finally, Stevens contended that the apparent transitivity demonstrated in cross-modality matches support to the 10 dB slope. Correct cross-modality results, however, are predictable with any slope if similar biases affect the scaling of the individual attributes involved. There is some evidence, for example, that the magnitude estimations extensively used in conjunction with cross-modality matches could be biased in any of several ways (Attneave, 1962; Banks and Hill, 1974; Curtis, 1970; Poulton, 1968). Thus, there are arguments in support of the 6 dB rate as well as for its rejection. The reasoning and evidence, however, remains sufficiently equivocal so that a hard stand pro or con might be a little premature for the time being.

What happens to the loudness functions when the zero-loudness level is changed? when the threshold is shifted? The predictions are illustrated in Figure 4 for K.A. These simulations show that as sensitivity is lost, loudness is predicted to grow more and more rapidly at near threshold intensities with relatively unaffected growth at sufficiently high levels. These results resemble findings well known in at least three circumstances, viz., equal-loudness matches at lower and lower frequencies, when a tone is masked by various levels of background noise, and in the complete recruitment of certain inner-ear pathologies. Conditions affecting both sensitivity and the relative magnitudes of relevant factors of (4) would generate functions resembling over and incomplete loudness recruitment as well. Thus, the nonlinear-algebraic theory provides a single framework for the analysis of conditions that can cause variations in the loudness growth patterns. While the adequacy of this aspect remains to be demonstrated, the nonlinear-algebraic approach appears to offer a much simpler explanation than those suggested from a different theoretical basis (Stevens, 1966; Stevens and Guirao, 1967).

In the final analysis then, the nonlinear-algebraic model predicts that nearly *equal stimulus ratios cause equal sensation ratios*. Thus, the principle underlying Steven's descriptive "power Law" is derivable from the



six postulates of the model. This formulation is more quantitative, and therefore, easily tested than prior attempts to explain this *law* (MacKay, 1963; Yilmas, 1967). For intensities below 30–40 dB, furthermore, the nonlinear-algebraic approach also predicts the significant departures from this *law* (Hellman and Zwislocki, 1963).

## Conclusions

Several presumptions about auditory intensity processing are required to interrelate the loudness of a tone with its generation of aural harmonics. One such set of six postulates constitutes the preliminary nonlinear-algebraic approach offered here. Accordingly, the pure tone input is initially converted via a power-series transformation—the only nonlinear stage in this model. Such a transformation produces aperiodic and periodic (aural harmonics) internal components. Subliminal periodic components are lost to further processing. Those with equal periods are combined into differentially weighted dimensions. These dimensions add algebraically with the aperiodic and *even*-ordered negative, and the *odd*-ordered positive. The resultant sum is the sensation of the pure tone called its loudness.

This set of presumptions is paralleled by a mathematical formulation that can be assessed once its parameters are estimated. The truncated version of the model developed here contains five such parameters: three are evaluated through results obtained in the tone-on-tone masking procedure, and two through fits of the expressions to limited samples of direct loudness judgments. Loudness-interval adjustments have been used, and the completed equation then extrapolated to predict other similar behaviors. Although an over-estimation tendency is apparent, the extrapolations are reasonably close to the findings from individual subjects. The completed equation also can be tested with other loudness judgments, e.g., ratio productions. In the present instance, the individuals were instructed to produce 2:1 loudness ratios. When told to adjust “half” loudnesses, the fits of the equation indicate that these subjects tended to produce ratios closer to one forth. Once these individual fractional ratios are determined, the responses to corresponding reciprocal proportions can be predicted. These predictions are often surprisingly accurate. Such successes indicate that the nonlinear-algebraic approach can interrelate task-specific behaviors—although the analysis sometimes conflicts with the assumptions that underly more traditional methods.

The nonlinear-algebraic approach has implications of broader theoretical interest as well. One is that it suggests a method for estimating the value of the first-

order (linear) coefficient of the power-series nonlinearity. Such values are basic to a functional description of auditory intensity processing. Another is that the theory predicts variations and the form of the loudness-growth function including Steven’s “Power Law”. Thus, the nonlinear-algebraic approach integrates formerly disparate theoretical considerations, and provides potential for many new insights into auditory intensity processing as well as loudness-judgment behaviors. Certainly this approach warrants further critical assessment.

*Acknowledgements.* This research has been supported by Public Health Service grant No. NS07352 and NS05785 from the National Institute of Neurologic Diseases and Stroke. The data reported here were collected by J. Erdreich.

## References

- Attneave, F.: *Psychology: A study of a science* Vol. 4, pp. 619–659. Koch, S., ed. New York: McGraw-Hill 1962
- Banks, W.P., Hill, D.K.: The apparent magnitude of number scaled by random production. *J. exp. Psychol. Mono.* **102**, 353–376 (1964)
- Boring, E.G.: *World of mathematics* Vol. 2, pp. 1148–1166. Newman, J. R., ed. New York: Simon and Schuster 1965
- Chapin, E. K., Firestone, F. A.: The influence of phase on tone quality and loudness; the interference of subjective harmonics. *J. Acoust. Soc. Amer.* **5**, 173–180 (1934)
- Clack, T.D.: Aural harmonics: The masking of a 2000-Hz tone by a sufficient 1000-Hz fundamental. *J. Acoust. Soc. Amer.* **42**, 751–758 (1967)
- Clack, T.D.: Aural harmonics: Preliminary time-intensity relationships using the tone-on-tone masking technique. *J. Acoust. Soc. Amer.* **43**, 283–288 (1968)
- Clack, T.D.: Some influences of subjective tones in monaural tone-on-tone masking. *J. Acoust. Soc. Amer.* **57**, 172–180 (1975)
- Clack, T.D.: Growth of the second and third aural harmonics of 500 Hz. *J. Acoust. Soc. Amer.* (in press)
- Clack, T.D., Bess, F.H.: Aural harmonics: The tone-on-tone vs. the best-beat method in normal and abnormal listeners. *Acta Otolaryngol.* **67**, 399–412 (1969)
- Clack, T.D., Erdreich, J.: Aural harmonics: A possible relation to loudness. *J. Acoust. Soc. Amer.* **51**, 113 (1972a)
- Clack, T.D., Erdreich, J., Knighton, R.W.: Aural harmonics: The monaural phase effects at 1500 Hz, 2000 Hz, and 2500 Hz observed in tone-on-tone masking when  $f_1 = 1000$  Hz. *J. Acoust. Soc. Amer.* **52**, 536–541 (1972)
- Curtis, D.W.: Magnitude estimations and category judgments of brightness and brightness intervals: A two-stage interpretation. *J. exp. Psychol.* **83**, 201–208 (1970)
- Egan, J.P., Klumpp, R.G.: The error due to masking in measurement of aural harmonics by the method of best-beats. *J. Acoust. Soc. Amer.* **23**, 275–286 (1951)
- Erdreich, J., Clack, T.D.: An algorithm for analysis of phase effects: Application to monaural distortion product estimation. *J. Acoust. Soc. Amer.* **52**, 1124–1126 (1972)
- Erdreich, J., Clack, T.D.: Aural harmonics: nonlinearity and loudness. *J. Acoust. Soc. Amer.* **53**, 312–313 (1973a)
- Fagot, R.F., Eskildsen, P.R., Stewart, M.R.: Effect of rate of change in physical intensity on bisection and fractionation judgments of brightness. *J. exp. Psychol.* **72**, 880–886 (1966)
- Fletcher, H.: *Speech and hearing*. New Jersey: Van Nostrand 1929

- Hellman, R.P., Zwillocki, J.: Some factors affecting the estimation of loudness. *J. Acoust. Soc. Amer.* **33**, 687—694 (1961)
- Lawrence, M.: Audiometric manifestations of inner ear physiology: The aural overload test. *Trans. Amer. Acad. Ophthalm. Otolaryng.* **62**, 104—119 (1958)
- MacKay, D.M.: Psychophysics of perceived intensity: A theoretical basis for Fechner's and Steven's laws. *Science* **139**, 1213—1216 (1963)
- Opheim, O., Flottorp, G.: The aural harmonics in normal and pathological hearing. *Acta Otolaryng.* **45**, 513—531 (1955)
- Plomp, R.: Beats of mistuned consonances. *J. Acoust. Soc. Amer.* **42**, 462—474 (1967)
- Poulton, E.C.: The new psychophysics: Six models for magnitude estimation. *Psychol. Bull.* **69**, 1—18 (1968)
- Stevens, J.C., Guirao, M.: Individual loudness functions. *J. Acoust. Soc. Amer.* **36**, 2210—2213 (1964)
- Stevens, S.S.: The measurement of loudness. *J. Acoust. Soc. Amer.* **27**, 815—829 (1955)
- Stevens, S.S.: The basis of psychophysical judgments. *J. Acoust. Soc. Amer.* **35**, 611—612 (1963)
- Stevens, S.S.: Power group transformations under glare, masking, and recruitment. *J. Acoust. Soc. Amer.* **39**, 725—735 (1966)
- Stevens, S.S., Davis, H.: *Hearing*. New York: Wiley 1938
- Stevens, S.S., Guirao, M.: Loudness functions under inhibition. *Percept. Psychophys.* **2**, 459—465 (1967)
- Torgerson, W.S.: *Theory and methods of scaling*. New York: Wiley 1958
- Trimmer, J.D., Firestone, F.A.: An investigation of subjective tones by means of the steady tone phase effect. *J. Acoust. Soc. Amer.* **9**, 24—29 (1937)
- Warren, R.M.: A basis for judgments of sensory intensity. *Amer. J. Psychol.* **71**, 675—687 (1958)
- Warren, R.M.: Elimination of biases in loudness judgments for tones. *J. Acoust. Soc. Amer.* **48**, 1397—1403 (1970)
- Warren, R.M.: Quantification of loudness. *Amer. J. Psychol.* **86**, 807—825 (1973)
- Wever, E.G.: *Theory of hearing*. New York: Wiley 1949
- Yilmaz, H.: Perceptual invariance and the psychophysical law. *Percept. Psychophys.* **2**, 533—538 (1967)
- Zwilslocki, J.J., Sokolich, W.G.: Velocity and displacement responses in auditory-nerve fibers. *Science* **182**, 64—66 (1973)

Received: October 27, 1977

Dr. T. Dean Clack  
Kresge Hearing Research Inst.  
Medical School  
University of Michigan  
Ann Arbor, MI 48109, USA