Fuzzy Distortion in Analog Amplifiers: A Limit to Information Transmission?*

M. J. HAWKSFORD

University of Essex, Department of Electrical Engineering Science, Wivenhoe Park, Colchester, Essex, United Kingdom,

A theoretical model is introduced that attempts to emulate a low-level distortion mechanism inherent in bipolar junction transistor amplifiers and, as a consequence, suggests a low-level bound to the transmission of fine signal detail. The model gives positive support to the low-feedback school of design and proposes circuit techniques for maximizing signal transparency. The design principles have particular relevance to low-level signal stages, but should also find an association with all classes of amplifiers.

0 INTRODUCTION

The last decade has seen substantial debate concerning the relationship between objective and subjective assessment of amplifiers. Measurements have frequently been performed with often impressive results [1], yet on extended audition significant audible differences can still be perceptible.

Various investigations have cited, for example, the levels of harmonic distortion as a measure of excellence, where emphasis has been directed to the distribution and relative weights of the harmonic structure. Conclusions have been drawn suggesting that low-order harmonics exhibiting a smooth rolloff in amplitude with frequency [2], [3] are a useful indicator of an amplifier's performance. However, when on this basis the levels of distortion are critically compared, it is generally difficult to assert a high correlation between objective and subjective results. In fact auditioning of amplifier performance suggests that the absolute level of harmonic distortion is, within limits, only a second-order interest, as highlighted during valve/transistor comparison.

A second indicator of potential excellence depends on the assessment of transient intermodulation distortion (TID) [4], [5], a distortion that is prevalent in slow high-loop-gain feedback amplifiers. However, design criteria have been established [6], [7] which minimize the onset of TID. Clearly, TID is only part of the distortion repertoire and is probably of minimal consequence once the probability of its occurrence is low.

Primary and secondary crossover distortion, though predominant in power amplifier circuits, also occur in certain low-level operational amplifiers that use class AB output stages. However, although this nonlinear mechanism can lead to significant signal impairment, there are now a variety of design techniques [8]–[10] that successfully minimize the error signal.

A direct consequence of amplifier nonlinearity and signal interaction is partial rectification, which produces a dynamic shift in the quiescent bias state. If an amplifier incorporates energy storage elements (such as ac coupling and by-pass capacitors), then the error signal is filtered and exhibits "overhang," which is dominant in the lower midrange and bass frequency bands. Amplifiers should therefore minimize energy storage components and be designed to be near aperiodic within the audio band. Research has shown that an asymmetric pulse test is a sensitive method of assessment [11], [12].

Where amplifiers are operated at high signal levels, other mechanisms of dynamic distortion become significant. Nonlinear delay modulation (NLDM) of the

* This paper was the basis of a lecture to the British Section in 1982 October (see *JAES*, vol. 31, no. 3, pp. 164 and 166 (1983 March)). Manuscript received 1982 October 11; revised 1982 November 22.

signal will occur due to the dynamic variation of transistor parameters with signal: Modulation of collector–base capacitance with collector–base voltage, the shift of small-signal bandwidth with collector current, and general parametric changes when devices are thermally exercised are all contributory factors. However, after reviewing the many conventional forms of nonlinearity it is apparent that certain areas of subjective assessment still elude a satisfactory explanation, and it is unclear as to an optimum design strategy. Specifically the area of greatest concern is that of subjective clarity or what may be usefully described as signal transparency: the ability to resolve fine signal detail, especially in the presence of complex high-level signal components. There appears to be a distinction between distortion mechanisms that "color" the signal, thus adding their own character, and distortions that corrupt fine signal detail.

This paper addresses what is believed to be both a significant and a neglected factor of amplifier performance where two basic clues have emerged: first, that amplifiers using low or distributed feedback often audition with higher rank, even though they may exhibit higher levels of error signal, and second, that low-level amplifier stages appear particularly susceptible to signal impairment. A primitive theory is proposed and a design strategy presented as a means of performance optimization.

In preparing the work presented in this paper, a literature survey revealed an embryonic idea first published by West [13] in 1978. However, the idea was not developed to any extent, and its significance with respect to amplifier design was not established in depth. A later discussion by Curtis [14] dismissed the theory as a cause of "transistor sound." The author considers this dismissal somewhat premature and attempts in this paper to extend the theory in more detail, with respect both to the charge-control model of a transistor and to the application of the derived theory to amplifier design.

1 FUZZY NONLINEARITY: THE THOUGHT EXPERIMENT

Classical circuit theory represents current as a continuous function that flows smoothly and can be considered to have infinite precision within an uncorrelated random bound. This viewpoint is taken from a macroscopic stance of electromagnetism where the individual electrical fields of electrons merge to a nongranular continuum that allows near infinite precision in the transmission of information. Account is of course taken of the behavior of partial randomness of electrons, and this is introduced through linear noise analysis where the noise is seen as the limiting factor on low-level signal resolution. In fact basic calculations on the numerousness of electrons would suggest this to be perfectly reasonable and of little consequence to the audio circuit designer. We speculate here that this may well be an invalid assumption which disguises the true limit to the ultimate resolution of a low-noise amplifier stage.

Transistor operation depends in part on the transfer of charge from signal source to device, a theory first proposed by Beaufoy and Sparkes [15]. Essentially the theory shows that the level of collector current in a bipolar junction transistor (BJT) is a linear function of the local stored charge in the base region. The theory also proposes that the continual base current of a BJT provides a "top up" charge to compensate for recombination resulting from a finite carrier lifetime within the base. In equilibrium the rate of recombination is just balanced by the base current to maintain a constant average charge, which in turn determines the collector current.

However, in this paper we shall not be concerned directly with the mechanics of device operation, only a consequence of those mechanisms, namely, the level of charge transfer required in the amplification process. The probable importance of charge levels can be established by the following thought experiment.

In this discussion we shall evaluate the approximate levels of charge that are transferred to the base of a transistor under low-level signal excitation. Fig. 1 shows a basic zero feedback amplifier stage interfaced to a moving-coil transducer with source resistance r_c, where the input impedance of the amplifier is derived directly from the hybrid-π equivalent circuit of a transistor. In Fig. 1 r_{bb}' is the base bulk resistance, $r_{b'e}$ the dc input resistance (modeling small-signal recombination), and $C_{b'e}$ the base region capacitance storing the charge q_b which controls the collector current.

A value of the base storage capacitor can be estimated directly from a knowledge of f_β, the 3-dB bandwidth of h_{fe}, which is the collector–base current gain, assuming a first-order response,

$$C_{b'e} = \frac{1}{2\pi r_{b'e} f_\beta} \bigg|_{V_{CB} \to \text{constant}} , \tag{1}$$

this expression is derived from the observation that the reactance of $C_{b'e}$ is equal to $r_{b'e}$ at the frequency f_β, it also follows that $C_{b'e} \propto I_e$ (emitter current).

Let us further our argument by considering the output voltage v_i of a moving-coil cartridge,

$$v_i = \frac{f}{f_n} V_n \sin(2\pi f t) \tag{2}$$

where V_n is the nominal cartridge output amplitude at a normalized frequency f_n, typically 1 kHz. If the dynamic range of the system is DR, then the minimum

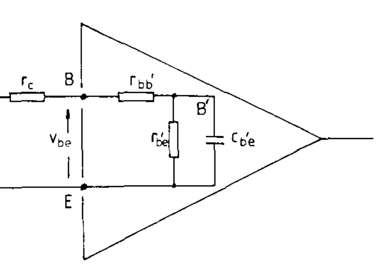

Fig. 1. Basic transistor amplifier stage.

resolvable signal level Δv_i is

$$\Delta v_i = \frac{f}{f_n} \frac{V_n}{DR} \sin(2\pi f t) . \quad (3)$$

In defining DR we refer to the smallest resolvable change in signal level that can exist without nonlinear corruption from complex high-level signal components. Ideally this change in signal level should be below the noise floor.

We next assign a minimum resolvable time period τ_m estimated by direct reference to the sampling theorem, which conveniently relates τ_m to the audible bandwidth f_a,

$$\tau_m \cong \frac{1}{2f_a} . \quad (4)$$

Assuming a sinusoidal input signal, an expression for the control base charge $q_b(t)$ for an input signal Δv_i is derived as

$$q_b(t) = \frac{1}{2\pi r_{b'e} f_\beta} \frac{f}{f_n} \frac{V_n}{DR} \sin(2\pi f t) \quad (5)$$

where

$$r_c + r_{bb'} << \left| r_{b'e} // \frac{1}{2\pi f C_{b'e}} \right| .$$

Hence the change in control charge that occurs over a time τ_m is

$$\Delta q_{b,\min} = q_b\left(t + \frac{\tau_m}{2}\right) - q_b\left(t - \frac{\tau_m}{2}\right) \quad (6)$$

where, aligning the difference equation to maximize Δq_b (in this sense our estimation is optimistically high) and assuming $\sin(\pi f \tau_m) \cong \pi f \tau_m$, we have

$$\Delta q_{b,\min} = \frac{V_n}{2 r_{b'e} DR} \frac{f^2}{f_n f_a f_\beta} . \quad (7)$$

We note from standard transistor theory that

$$r_{b'e} = (1 + h_{fe}) r_e \quad (8)$$

$$r_e = \frac{0.025}{I_e} , \ (I_e \text{ in amperes}) \quad (9)$$

$$f_T \simeq (1 + h_{fe}) f_\beta . \quad (10)$$

Hence

$$\Delta q_{b,\min} = \frac{20 V_n I_e}{DR} \frac{f^2}{f_n f_a f_T} .$$

To estimate typical values of changes in the base charge consider the following data base:

V_n	$= 200\ \mu V$	medium output moving-coil cartridge,
I_e	$= 10^{-3}$ A	transistor emitter bias current,
f_T	$= 50$ MHz	bandwidth to unity h_{fe},
DR	$= 10^4$	80-dB dynamic range,
f_n	$= 1$ kHz	normalizing frequency for cartridge,
f_a	$= 20$ kHz	audible bandwidth,
e	$= 1.96 \times 10^{-19}$ C	charge on electron,
h_{fe}	$= 500$	small-signal collector–base current gain (V_{CE} constant)

whereby the minimum change in base charge is evaluated as

$$\Delta q_{b,\min} \simeq (2 \times 10^{-6} f^2) e \qquad [\text{coulombs}] \quad (11)$$

Eq. (11) shows a remarkably low level of average charge transfer that occurs for small signals observed over the minimum resolvable time period (here assumed to be 25 μs).

It is also instructive to estimate the change in the number of electrons transferred into the base region through recombination over the minimum time period τ_m, due only to the minimum signal component Δv_i.

If we assume $r_{b'e}$ to be the dominant input resistance of the transistor, the base input current Δi_i associated with Δv_i is

$$\Delta i_i = \frac{f}{f_n} \frac{V_n}{r_{b'e} DR} \sin(2\pi f t) . \quad (12)$$

The charge Δq_r transferred from source to input due to recombination in time τ_m is calculated by integration,

$$\Delta q_r = \int_{t-\tau_m/2}^{t+\tau_m/2} \Delta i_i \, dt . \quad (13)$$

Aligning the integration window to maximize Δq_r and again assuming that $\pi f \tau_m$ is small,

$$\Delta q_r = \frac{20 V_n I_e}{(1 + h_{fe}) DR} \frac{f}{f_n f_a} . \quad (14)$$

Using the same data base,

$$\Delta q_r = 0.2 f e \qquad [\text{coulombs}]. \quad (15)$$

Eqs. (11) and (15) show that low-level signals in transistor stages are associated with an extremely small transfer of charge into the base of the input transistor. The basic analysis indicates that within τ_m the signal amplitude generally has greater effect on the charge transferred for recombination than that charge having direct control of the collector current (according to charge control theory). Nevertheless both calculations yield results of only a few electrons.

We therefore propose a theory that partial signal quantization is the fundamental process that sets an inherent bound to signal transparency through a transistor stage. Both Eqs. (11) and (15) support the probable existence of significant granularity where Eq. (11) suggests a form of amplitude quantization and Eq. (15) an association with $1/f$ noise.

It is also proposed that signal interaction with inherent nonlinearities in transistors, together with even small levels of interference from power supplies, neighboring circuitry, or undesired signal coupling (such as poor ground line design), can easily corrupt such minute signals and that such corruption should be interpreted as modifications to these low charge levels.

We conclude this preliminary discussion by giving in Table 1 typical levels of charge transferred to the base of a transistor within the minimum time period $\tau_m = 25\ \mu s$ against various signal levels to illustrate the potential dynamic range available. The example already cited in this section is used as a data base.

2 FUZZY MODELS

In this section we build upon the observations made of quantization and the relative magnitudes of low-level signals by introducing a basic model of the distortion process. It is emphasized that although the model is primitive, it is a natural extension of our thought experiment.

The proposed model is to be classed as "fuzzy" and the resulting distortion as fuzzy distortion due to its strong stochastic association. We commence by establishing two distinct groups of nonlinearity.

1) *Deterministic nonlinearity.* Classic system nonlinearity can be envisaged using a continuous model incorporating static or dynamic transfer characteristics. The main attribute of this broad distortion classification is repeatability where, assuming no time-dependent system parameters, the same error waveform will result under repeated tests. We note in particular that when measuring such distortion a degree of signal averaging is often used to suppress random events.

2) *Fuzzy nonlinearity.* A distortion process that results in an error signal with a strong stochastic element that does not include any uniform sampling function is defined here as fuzzy distortion. Such distortion will not exhibit exact error waveform replication under repeated tests. We note in particular that when measuring such distortion, any signal averaging will tend to mask the error waveform.

We proceed by further reference to the charge control model of a BJT [15] and attempt to produce a primitive model that matches the input impedance characteristic of a BJT transistor (see Fig. 1), exhibits the correct frequency response when observing h_{fe}, introduces a degree of charge quantization, and maintains the proper static relationship between base and collector currents.

The proposed model is illustrated in Fig. 2(b) and is configured so that it replaces directly the standard hybrid-π circuit shown in Fig. 2(a). A simplified notation is illustrated in Fig. 2(c).

Table 1. Typical levels of charge transferred.

	Charge Transfer in 25 μs	
	Δq_r	$\Delta q_{b,\min}$
Bias current of 1/500 mA	$2.56 \times 10^8 e$	—
Input signal of 200 μV	$2 \times 10^6 e$	$2 \times 10^4 e$
Input signal 80 dB below 200 μV at 1 kHz	$200e$	$2e$

The model consists of an integrator to convert input signal current to charge, cascaded with a uniform quantizer with an associated dither source $n(t)$ to scatter the quanta. The integrator and quantizer are enclosed within a negative-feedback loop, which together emulate the process of recombination and quantization of the stored base charge. The quantized base–emitter voltage $V_{b'e}$, which is proportional to the stored base charge, is converted to collector current by a transconductance stage with mutual conductance g_m. From standard transistor theory,

$$g_m = \frac{h_{fe}}{(1 + h_{fe})r_e} \tag{16}$$

$$r_e = \frac{\partial V_{BE}}{\partial I_E} = \frac{kT}{eI_e} \tag{17}$$

where k is Boltzmann's constant, T the junction temperature (kelvins), e the charge on an electron, and I_e the emitter bias current.

The model shown in Fig. 2(b) has a strong resemblance to certain classes of analog-to-digital encoder, in particular feedback (pulse-code modulation) and multilevel delta sigma modulation (DSM) [16], [17]. Since these encoding schemes combine integration and quantization within a feedback loop, they form useful vehicles for comparison. A major distinction between

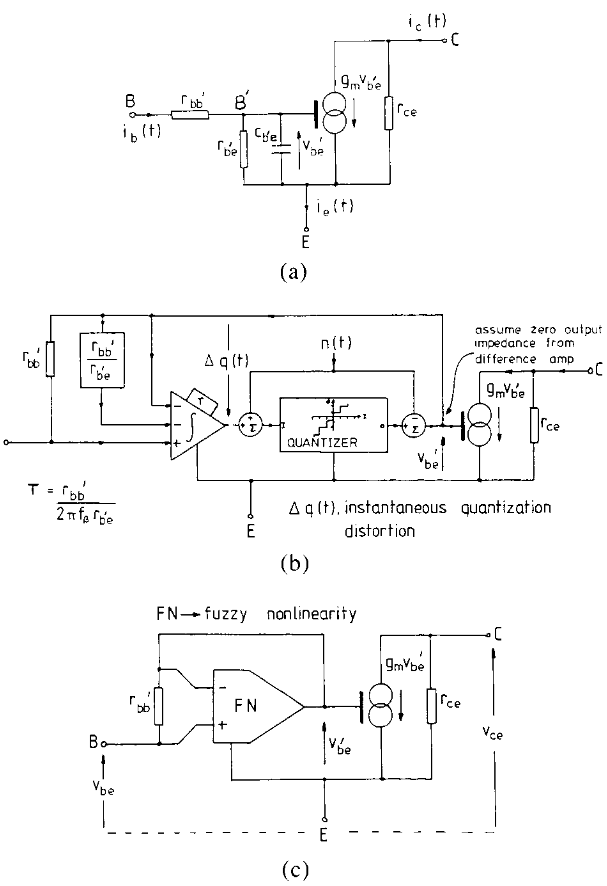

Fig. 2. Basic hybrid-π equivalent circuit of a BJT. (a) Standard circuit. (b) Circuit with modification to incorporate charge quantization. (c) Simplified functional presentation, including charge quantization.

the fuzzy model and digital encoders is that the former excludes a uniform sampling process. However, a random sampling function is permissible where the mean sampling frequency corresponds to the mean rate of recombination within the base of the transistor, which is determined by the base bias current (that is, a base bias current of 2 μA corresponds to a mean sampling rate of $\cong 10^{13}$ Hz). We note also from Eq. (2) that the mean sampling rate will undergo frequency modulation due to the instantaneous change in recombination current with change in base–emitter voltage.

We estimate the approximate frequency characteristic of the distortion spectra for the model of Fig. 2(b) by assuming the loop to be essentially linear and by representing the quantization distortion as a sinusoidal error signal added within the loop where, for purposes of analysis,

$$q(t) = Q\mathrm{e}^{\mathrm{j}2\pi ft} . \qquad (18)$$

Thus the collector error current $I_{q,c}\mathrm{e}^{\mathrm{j}2\pi ft}$ follows as

$$I_{q,c} = \frac{\mathrm{j}Qg_m f/f_\beta}{\left[1 + \frac{r_{b'e}}{r_c + r_{bb}'}\right]\left[1 + \frac{\mathrm{j}f}{[1 + r_{b'e}/(r_c + r_{bb}')]f_\beta}\right]} \qquad (19)$$

where r_c is the source resistance between base and emitter, as shown in Fig. 1.

From Eq. (19) we infer the basic form of error spectrum, which is illustrated in Fig. 3. Note the effect a low source impedance has on the break frequency in the approximate error spectrum.

The error spectrum shown in Fig. 3 compares with the general trend of pulse-code-modulation-type systems [16], [17] where quantization is dominant at high frequencies. The curve ignores other forms of random noise, such as the noise associated with r_{bb}'. Thus in general this effect will be at or below the device noise level.

The results show that the source resistance plays a dominant role in shaping the error spectrum where optimum performance is obtained when r_c is minimized. This compares favorably with the more common noise model of a transistor where the noise sources are represented as equivalent input noise voltage and current generators. An interesting by-product of the model structure is that it includes a mechanism that modifies the output noise as a function of source resistance, this being achieved by modifying the feedback factor around the internal feedback loop. Essentially when $r_c = 0$, maximum feedback is applied, and when $r_c = \infty$, minimum feedback results. Examination of the model schematic illustrated in Fig. 2(b) should clarify this operation.

In this section we have established a modification to the basic hybrid-π model of a transistor which includes the effect of charge quantization. We now proceed to examine some implications of these observations.

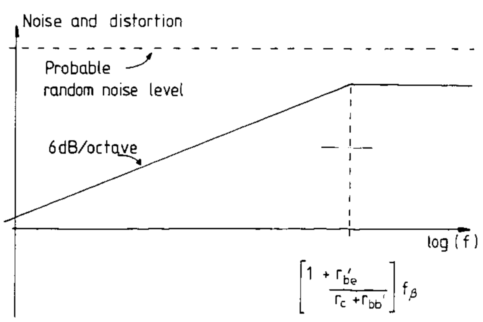

Fig. 3. Approximate error spectrum of collector current due only to quantization effects.

3 IMPLICATIONS OF FUZZY DISTORTION IN AMPLIFIER DESIGN

If we accept that a low-level nonlinear mechanism exists in transistors which has a different nature from deterministic nonlinearity, then we can make some basic observations as to the correct global strategy toward amplifier design.

Where a transistor operates with very-low-level signals that approach the noise floor, the artifacts of charge quantization will generate significant fuzzy distortion. The method to minimize this effect can be summarized as follows:

1) It is essential that low-noise devices be used that exhibit low r_{bb}' and low $1/f$ (recombination noise). The device should be chosen so as to maximize $C_{b'e}$. Thus large integrated arrays of transistors where many matched devices are paralleled should prove the best choice (such as the LM394).

2) Operate transistors so that $C_{b'e}$ is maximized. From Eq. (1) this infers a substantial level of emitter bias current which in turn will lower the device input impedance.

3) Eq. (1) infers that $C_{b'e}$ is an inverse function of $r_{b'e}$ (for given f_β). Thus a device should be chosen with a low value of h_{fe} [see Eq. (8)].

4) Selection of f_β is more complex. A low level of f_β will increase $C_{b'e}$, but at the expense of lowering the break frequency in the distortion spectra (see Fig. 3). It is suggested that f_β should be sensibly in excess of 20 kHz.

5) Design the transistor stage so as to maximize the device loading factor [18]. This will maximize the changes in charge for a given signal.

6) Minimize resistance in the input mesh of a transistor. This will reduce low-frequency fuzzy distortion. For the input stage of Fig. 1 this implies a low value of r_c. The effect of r_c can be observed by reference to Eq. (19) and the error spectrum in Fig. 3.

Consider by way of example a low-level disk preamplifier stage for use with a low-output moving coil cartridge. In Fig. 4 a moving-coil cartridge with source resistance r_c and generator signal $e_c(t)$ is interfaced to a disk amplifier which has an input resistance r_{in} and a voltage gain A_v.

The classical viewpoint would not expect r_{in} to play an important role other than providing an optimum load for the cartridge. (This may affect the frequency response, for example, when coupled with the generator source inductance.) However, for low-output-impedance moving-coil cartridges this generally has minimal effect. Indeed in selecting an "optimum" load resistance, it is normal to use an input shunt resistor.

However, fuzzy nonlinearity suggests that the level of i_{in} is of fundamental importance, and that this current must be maximized and flow into the base of the input transistor. A shunt input resistance is not an acceptable solution, as current will by-pass the transistor.

It therefore follows that the input signal must be considered in terms of both input voltage and input current. It is the input signal power that is fundamental.

3.1 Corollary 1

If we accept the notion of maximizing the signal power that flows into the base of the input transistor, then a transducer for an analog disk system must be selected such that

1) It converts a relatively high proportion of platter rotational energy into mechanical signal energy, as seen at the cantilever of the cartridge.

2) It exhibits a high mechanical-to-electrical power conversion.

It is possibly in these areas of performance where many moving-coil cartridges offer a significant performance advantage.

3.2 Corollary 2

In selecting a matching transformer/input circuit topology, the aim must be to maximize the flow of signal power into the base of the input transistor.

The proposal to maximize the input power is open to some debate. However, if it is realized that we wish both to maximize input signal current to the base of each transistor and to minimize source resistance r_c, then the notion of power maximization is a reasonable target.

Corollary 2 has profound ramifications in the choice circuit topology. Consider the classical amplifier configuration shown in Fig. 5. The circuit shows an input signal generator e_c with source impedance r_c. Again the amplifier has an input impedance r_{in} and voltage gain A_v, but a negative-feedback loop is included where the feedback factor is B with a Thévenin source impedance (seen by the inverting input) of r_f.

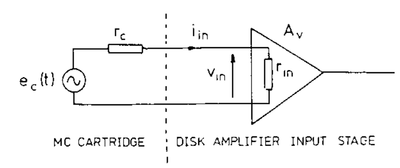

Fig. 4. Moving-coil cartridge–amplifier interface.

We proceed by calculating the instantaneous input signal power $p_{in}(t)$ to the differential input of the amplifier,

$$p_{in}(t) = \left[\frac{e_c(t)}{(r_c + r_f) + r_{in}(1 + AB)}\right]^2 r_{in} . \quad (20)$$

Differentiating $p_{in}(t)$ *wrt* r_{in},

$$\frac{\partial p_{in}(t)}{\partial r_{in}} = \left[\frac{e_c(t)}{(r_c + r_f) + r_{in}(1 + AB)}\right]^2 \times \left[1 - \frac{2r_{in}(1 + AB)}{(r_c + r_f) + r_{in}(1 + AB)}\right] \quad (21)$$

and setting $\partial p_{in}(t)/\partial r_{in} = 0$ to maximize the input power, the optimum r_{in} (for maximum input power) follows as

$$r_{in}\Big|_{opt} = \frac{r_c + r_f}{1 + AB} . \quad (22)$$

This gives the maximum input power as

$$p_{in}(t)\Big|_{max} = \frac{e_c^2(t)}{4(r_c + r_f)(1 + AB)} . \quad (23)$$

Eq. (23) shows the need to minimize all extraneous resistances within the input signal mesh, which is also a requirement for good noise design (that is, minimize r_f). [See also Eq. (19) and the discussion in Section 2 concerning input mesh resistance and fuzzy distortion.]

However, a more fundamental observation shows the maximum power flow to be an inverse function of the feedback parameter. Thus although classical feedback theory would suggest an improvement by operating the device well into its linear region of operation, it in fact forces the signal to within a relatively few quanta, thus exaggerating any effects of quantization.

To illustrate the process further, consider the combined systems of Figs. 2(c) and 5, as shown in Fig. 6.

Any amplification which follows the quantization process must by necessity amplify the quantized signal, together with additional random noise sources. The effect of negative feedback on a purely linear system will reduce the levels of additional noise resources that are injected within the feedback loop by a factor of $(1 + AB)$. However, this process is not true of a loop that includes quantization. In fact in this system the feedback will again reduce the additive noise, but it will only

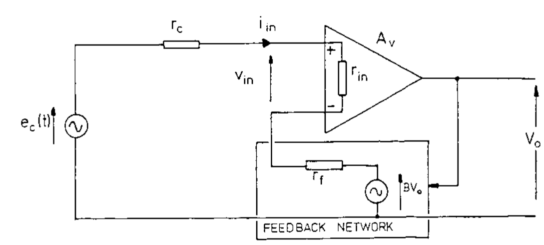

Fig. 5. Classical feedback amplifier structure.

partially reduce the effects of quantization. Thus fuzzy distortion components will be partly exposed by negative feedback, together with a complex process of intermodulation between signal additive noise and quantization distortion, which in general must include time smearing due to limitations in loop bandwidth.

4 MINIMIZATION OF FUZZY NONLINEARITY

We conclude our discussion on fuzzy distortion by suggesting a design method and basic circuit topologies that in principle meet the requirements of both high-level and low-level nonlinearities. In particular we emphasize low-level signal stages as these are potentially more susceptible to fuzzy nonlinearity.

Following the design aims discussed in Section 3, we must choose a low-noise transistor with a low value of collector–base current gain. This device should be operated at a collector current commensurate with noise considerations such that (ideally) the input impedance between base and emitter matches the source impedance of the transducer or presents an optimum load to the transducer. Provided the source signal is of suitable magnitude, the signal should be coupled directly to the base–emitter junction (assuming that high-level distortion will not be problematic), and preferably no ac coupling component should be used.

Coupled with this requirement, the input transistor should ideally use no feedback (local or overall), since Eq. (23) indicates a reduction in signal power. If the transducer output is too great, resulting in a high level of deterministic distortion, then a step-down transformer should be selected to permit using a zero feedback input stage. Ideally the input impedance should be designed to match the transformed source impedance of the transducer. This process will not change (in principle) the level of power extracted from the source (assuming a power match), but it will minimize high-level distortion and eliminate a loss of input power through the use of negative feedback. In many instances it will not be practical to design for a power match as high operating currents or many parallel devices may be necessary, though investigation into the LM394-type device should be encouraged.

In general an amplifier system will include several cascaded transistor stages within the signal path. Potentially each stage is a cause of low-level distortion, but as with noise design, the first transistor should be the dominant offender. To guarantee this ideal we must arrange for a progressive increase in input signal power as we proceed along the cascade, as well as attempting to minimize the number of series-connected transistors in the signal path.

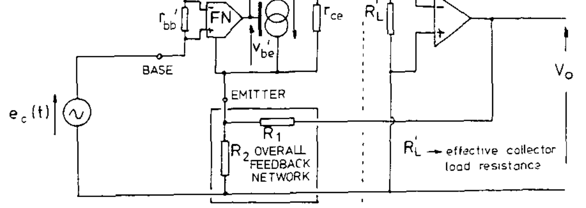

Fig. 6. Simplified feedback amplifier with quantizer model of Fig. 2(c).

In order to control deterministic distortion as signal levels are amplified, a degree of negative feedback will become mandatory, which will consist generally of a combination of distributed and multiple-feedback loops. However, in selecting the topology for the feedback structure, an increasing input-signal-power progression should be observed.

To examine this design strategy, consider $N + 1$ cascaded transistor stages, as shown in Fig. 7. Stages $1 \rightarrow N$ use distributed feedback, while the input stage 0 is optimized for fuzzy nonlinearity by using zero feedback. A single feedback loop encloses stages $1 \rightarrow N$, thus modeling a typical amplifier.

Let

$$\left.\begin{array}{lll} A_0, \ldots, A_N & = & \text{amplifier gains} \\ B_0, \ldots, B_N & = & \text{feedback factors}^1 \\ e_0, \ldots, e_N & = & \text{amplifier input signals} \end{array}\right\} \text{see Fig. 7}$$

$$\begin{array}{lll} r_0, \ldots, r_N & = & \text{amplifier input resistances} \\ p_0, \ldots, p_N & = & \text{input signal powers to amplifiers} \\ r_c & = & \text{transducer source impedance.} \end{array}$$

From Eq. (20) we calculate the rth-stage input signal power p_r. For stages $r = 1, \ldots, N$. (Assume that the source resistance is small compared with r_r.)

$$p_r = \left[\frac{e_r}{1 + A_r B_r}\right]^2 \frac{1}{r_r} \tag{24}$$

and for stage $r = 0$,

$$p_0 = \left(\frac{e_0}{r_c + r_0}\right)^2 r_0 \ . \tag{25}$$

4.1 Design Criterion

To minimize signal degradation caused by fuzzy nonlinearity in a cascade of transistor stages,

$$p_r = G_{fr} p_{r-1} \tag{26}$$

where ideally the interstage power gain $G_{fr} > 1$. We calculate the voltage gain relating e_r and e_{r-1}, for $r = 1$,

$$\frac{e_1}{e_0} = \frac{A_0}{\left\{1 + B_0 \prod_{p=1}^{N} [A_p/(1 + A_p B_p)]\right\}} \tag{27}$$

and for $r = 2, \ldots, N$,

$$\frac{e_r}{e_{r-1}} = \frac{A_{r-1}}{1 + B_r A_r} \ . \tag{28}$$

Hence we establish the constraints on the choice of feedback parameters by reference to Eqs. (24)–(28).

[1] The feedback networks are assumed to exhibit zero Thévenin source impedances.

If

$$\delta_{\mathrm{f1}} = \left\{1 + B_0 \prod_{p=1}^{N} \frac{A_p}{1 + B_p A_p}\right\}\left\{\frac{\sqrt{G_{\mathrm{f1}} r_0 r_1}}{r_c + r_0}\right\} \quad (29)$$

$$\delta_{\mathrm{f}r} = \sqrt{\frac{G_{\mathrm{f}r} r_r}{r_{r-1}}}\Bigg|_{r=2,\ldots,N} \quad (30)$$

we have

$$A_{r-1} = \delta_{\mathrm{f}r}(1 + B_r A_r)\Big|_{r=1,\ldots,N} \quad (31)$$

where we define $\delta_{\mathrm{f}r}|_{r=1,\ldots,N}$ as the set of fuzzy gain parameters of the amplifier system.

Examination of Eqs. (29)–(31) reveals the design criterion that will ensure a progressive power increase along the cascade of transistor stages (noting that calculated power levels refer to the input power to each transistor, not the associated circuitry).

In practice there will be a limit to the input power to a transistor that will be dependent on the acceptable levels of deterministic distortion. We note that for a bipolar transistor which adheres to the form of Eq. (17) the fractional error component of emitter current is independent of I_{EO} for a given V_{BE} (where the subscript EO infers quiescent values),

$$V_{\mathrm{BE}} = V_{\mathrm{BEO}} + \Delta V_{\mathrm{BE}}$$

which corresponds to $I_{\mathrm{E}} = I_{\mathrm{EO}} + \Delta I_{\mathrm{E}}$. Then

$$\frac{\Delta I_{\mathrm{E}}}{I_{\mathrm{EO}}} = e^{(q\Delta V_{\mathrm{BE}}/KT)} - 1 \; . \quad (32)$$

Since the base–emitter voltage of a transistor is directly dependent upon the input power and input resistance, the input resistance should be minimized to reduce high-level distortion, for a given power level.

It is constructive to reflect upon a common circuit arrangement where a discrete transistor stage is cascaded with a BJT operational amplifier with local feedback. We will assume for simplicity that there is no overall feedback and proceed by suggesting typical circuit parameters:

Discrete stage		
input impedance	1 kΩ	(transistor)
voltage gain	20	
Operational amplifier stage		
input impedance	1 MΩ	(operational amplifier)
closed-loop gain	20	
open-loop gain	1000	(conservative estimate)

Hence from Eq. (31) $\delta_{\mathrm{f}r} \cong 0.4$, whereby the power gain follows from Eq. (30) as 1.6×10^{-4}.

This result shows a substantial reduction of input signal power presented to the second stage, the consequence being that any quantization effects will be significantly increased.

This circuit arrangement has often been used for disk preamplifiers and is a good illustration of a potential hazard of using high-gain high-input-impedance operational amplifiers.

We conclude this section by suggesting how it is possible to use a combination of low distributed feedback with feedforward error correction as a compromise to the distortion dichotomy existing between deterministic and fuzzy nonlinearities.

Germane to the design strategy is the selection of a distributed feedback system where the appropriate gains and feedback factors are calculated according to our now established fuzzy nonlinearity criterion. In so doing we accept that deterministic nonlinearity inherent in devices will potentially increase. However, by using nested feedforward error correction we can partially compensate the deterministic error signals and achieve acceptable linearity with high loading factors, even when local negative feedback is low.

In Fig. 8 we illustrate a two-stage feedforward amplifier where the error due to base–emitter nonlinearity in T_1 is partially corrected by the differential amplifier formed by T_2 and T_3. Further error-correction stages can be used to compensate for T_2 and T_3 nonlinearity using a nested configuration. The performance of such stages as a function of loading factor was considered in a previous paper [18].

The dominant advantages of this approach is that only very modest local negative feedback need be applied via R_1 and that the high-level distortion is partially compensated by the error amplifier. Such a technique allows good signal power coupling to T_1, yet permits an acceptable high-level distortion characteristic. It therefore follows that T_1 is exercised over a wide range of its operating characteristic while retaining good overall linearity.

The example just discussed illustrates how ideas of fuzzy nonlinearity could influence amplifier design. A second area of application concerns the construction and layout of circuits. Once the very small signal levels are appreciated and the point of view of "counting electrons" is taken, such factors as metal–metal con-

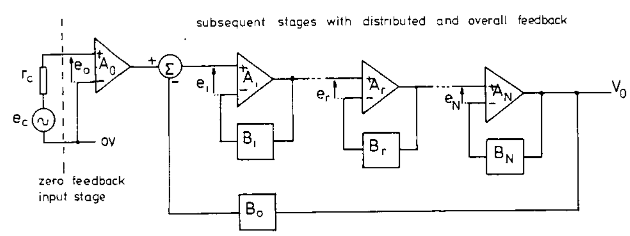

Fig. 7. Basic multiple-loop feedback amplifier topology.

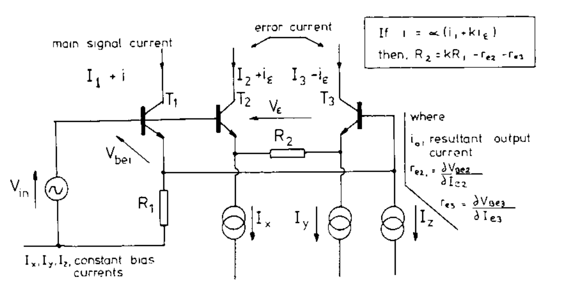

Fig. 8. Basic feedforward error-correction stage.

tacts, interference from adjacent circuits, and the displacement of charge through the dielectric of a capacitor require careful attention. These secondary factors will not be discussed in this paper, but they are influential in setting potential limits to signal transparency.

5 CONCLUSIONS

This paper has speculated on the existence of low-level nonlinearity inherent within BJT devices due to the quantization of charge carriers and has drawn attention to the relative magnitude of low-level signals. A model was introduced which forms a vehicle for comparison between an analog device and a class of digital modulation. This comparison is useful in that it is possible to speculate upon the nature and characteristics of the distortion.

Consideration of the mechanism of fuzzy distortion drew attention to the role of the input signal current at the base of a transistor and the need to maximize its value. This led directly to the usefulness of input signal power as a parameter in establishing levels of fuzzy distortion. A target design objective suggested the need to maximize this power flow, where the flow must be directly into the base–emitter junction and not into an external shunt resistor.

The role of negative feedback was then debated, where it was shown that the input signal power was an inverse function of amplifier loop gain. It was therefore concluded that levels of feedback should be minimized and that extraneous resistance within the input mesh should also be minimized. However, it was noted that the role of negative feedback offers a contribution to high-level signal distortion, but that its application must be considered with great care from the viewpoint of fuzzy nonlinearity.

A brief discussion was presented where the bounds on the selection of feedback factor and forward amplification were established. Finally a circuit technique using low levels of distributed feedback with feedforward error correction was introduced as a means of circumventing the distortion dichotomy, thus allowing both good low-level and high-level distortion characteristics, that is, the dynamic range.

It is satisfying to see some of the design objectives compatible with established design techniques which are used to minimize the artifacts of TID and also as support to the low-feedback school of design, in particular since there are now several good-quality amplifiers which adhere in part to these design objectives and also have excellent subjective ratings.

Finally it must be emphasized that the ideas presented here are the extension of a thought experiment into the approximate nature and behavior of low-level signals in amplifiers. Clearly such parameters as the physical size of transistors and the relative amounts of total charge stored in the base region are of importance. However, such considerations put in doubt the application of BJT operational amplifiers with their high open-loop gains, very high differential input impedance due to the low collector bias currents in the input transistors, and low real estate. If such devices exhibit low-level quantum effects, they are not suitable for use in high-quality audio amplifiers where precision of control of fine signal detail is mandatory. In fact, applying the thought experiment discussed in Section 1, the implication of Eq. (23), and the example of the BJT operational amplifier in Section 4, the potential consequences should at least be of concern to the circuit designer.

6 REFERENCES

[1] N. Keywood, "Amplifier Review," *Hi-Fi News*, vol. 27, pp. 37–45 (1982 Aug.).

[2] J. Hiraga, "Amplifier Harmonic Distortion Spectrum Analysis," *Hi-Fi News*, vol. 22, pp. 41–45 (1977 Mar.).

[3] C. Ray, "Negative Feedback and Non-linearity," *Wireless World*, vol. 84, pp. 47–50 (1978 Oct.).

[4] M. Otala, "Transient Distortion in Transistorized Audio Power Amplifiers," *IEEE Trans. Audio Electroacoust.*, vol. AU-18, pp. 234–239 (1970 Sept.).

[5] M. Otala, "Circuit Design Modifications for Minimizing Transient Intermodulation Distortion in Audio Amplifiers," *J. Audio Eng. Soc.*, vol. 20, pp. 396–399 (1972 June).

[6] P. Garde, "Transient Distortion in Feedback Amplifiers," *J. Audio Eng. Soc.*, pp. 314–322 (1978 May).

[7] W. M. Leach, "An Amplifier Input Stage Design Criterion for the Suppression of Dynamic Distortion," *J. Audio Eng. Soc. (Engineering Reports)*, vol. 29, pp. 249–251 (1981 Apr.).

[8] J. Vanderkooy and S. P. Lipshitz, "Feedforward Error Correction in Power Amplifiers," *J. Audio Eng. Soc.*, vol. 28, pp. 2–16 (1980 Jan./Feb.).

[9] M. J. Hawksford, "Distortion Correction in Audio Power Amplifiers," *J. Audio Eng. Soc. (Engineering Reports)*, vol. 29, pp. 27–30 (1981 Jan./Feb.).

[10] S. Takahashi and S. Tanaka, "Design and Construction of a Feedforward Error-Correction Amplifier," *J. Audio Eng. Soc. (Engineering Reports)*, vol. 29, pp. 31–37 (1981 Jan./Feb.).

[11] Y. Hirata, M. Ueki, T. Kasuga, and T. Kitamura, "Nonlinear Distortion Measurement Using Composite Pulse Waveform," *J. Audio Eng. Soc. (Engineering Reports)*, vol. 29, pp. 243–248 (1981 Apr.).

[12] Y. Hirata, "Quantifying Amplifier Sound," *Wireless World*, vol. 87, pp. 49–52 (1981 Oct.).

[13] R. West, "The Great Amplifier Debate, 2-Transistor Sound," *Hi-Fi News*, vol. 23, p. 79 (1978 Jan.).

[14] S. Curtis, "Amplifier Noise and Clipping," *Hi-Fi News*, vol. 23, pp. 81–85 (1978 June).

[15] R. Beaufoy and J. J. Sparkes, "The Junction Transistor as a Charge Control Device," *ATE J.*, vol. 13, pp. 310–327 (1957 Oct.).

[16] R. Steele, *Delta Modulation Systems* (Pentech Press, London, 1975).

[17] M. J. Hawksford, "Unified Theory of Digital Modulation," *Proc. IEE*, vol. 121, pp. 109–115 (1974 Feb.).

[18] M. J. Hawksford, "Distortion Correction Circuits for Audio Amplifiers," *J. Audio Eng. Soc.*, vol. 29, pp. 503–510 (1981 July/Aug.).

THE AUTHOR

Malcolm Hawksford was educated at the University of Aston in Birmingham, England, from 1965 to 1971. In 1968 he obtained a first class honors degree in electrical engineering, and that same year was awarded a BBC research scholarship to investigate the application of deltamodulation to color television. In 1972 he obtained a Ph.D. degree. In 1971 Dr. Hawksford became a lecturer at the University of Essex, England, in the Department of Electrical Engineering Science. During his time at Essex he actively pursued research projects within the field of audio engineering, where projects on power amplifier design, loudspeaker crossover design, analog-to-digital conversion, and music synthesis have been undertaken. He has presented papers at conventions of the Audio Engineering Society. He is currently a member of the AES, IEE, and RTS.