Eliminating Broadband Distortion in Transistor Amplifiers

By LEE C. THOMAS

(Manuscript received July 6, 1967)

This paper presents the results of a study directed toward understanding the basic distortion mechanisms in transistors. (i) We develop an analytic model for the transistor which describes small signal linear performance and nonlinear effects. The linear model is matched to the measured h-parameters of the device over a wide range of frequency and bias current. We superimpose three distinct nonlinear effects on this linear skeleton model, all approximated to third order terms. (ii) We show experimental confirmation that, for some bias-load conditions, the second order distortion can be minimized and we show that it is possible to simultaneously minimize both second- and third-order distortion under the same bias-load condition. This result also is confirmed experimentally. (iii) We derive and discuss in detail an analytic expression for the optimum load. Based on this expression, we present detailed procedures for finding this optimum condition for any transistor, and give experimental corroboration. (iv) We give a qualitative description of the interaction among these three nonlinear effects based on an analog computer simulation of the model. This description makes it easier to visualize the distortion cancellation phenomena derived in this paper, and indicates a technique for extending the effect to a broad band of frequencies. We conclude that proper use of the distortion cancellation effect can greatly improve intermodulation performance in existing transistors.

I. INTRODUCTION

System studies have indicated that very broad band (greater than 20 mHz) AM coaxial cable systems will be modulation-limited. Intensive investigations to understand and characterize the inherent modulation properties of devices and repeater circuits have been called for. We made one such study directed toward understanding the basic distortion mechanisms in transistors.

The history of transistor distortion literature can be characterized as an erosion process in which highly restricted parts of the total problem are attacked leaving fresh complexities exposed for future work. In early work by Akgun and Strutt, the analysis is restricted to nonlinearities in the emitter resistance assuming an ac short at the input and output.[1] Observed nulls in second and third order distortion do not correlate with the theory, which does, however, include frequency effects. Using many of the same assumptions, Mallinckrodt and Gardner extended this earlier work to account for a third order null at low frequencies when the nonlinear emitter resistance is dominant.[2]

More recently Riva, Beneteau, and Dalla Volta considered all important sources of distortion by breaking the problem into three distinct operating regions with expressions for minimizing second order distortion in each.[3] They do not treat of third order minimization, and they use a dc model. Reynolds analyzes third order minimization at particular nonzero frequencies for dominance of the emitter resistance nonlinearity.[4]

There are two reasons for the specialized nature of these efforts. First, transistors, as contrasted with vacuum tubes, have at least three dominant nonlinearities. It would be difficult to consider all of these in a general expression for second and third order distortion. Second, frequency effects can be important in many applications. In general, the analysis of nonlinear effects as a function of frequency requires the use of extremely powerful and, as a result, cumbersome analytic techniques. In the special case of an exponential input v-i relation it is possible to avoid a general analysis, which explains why analyses which include frequency effects have been limited to emitter nonlinearities. Even in this exponential case, however, the third order null predicted by Reynolds is a narrowband effect, applicable only at a particular frequency.

This paper extends these earlier efforts in four important respects.

(i) We conclude that the distortion measured at the terminals results from algebraic cancellation between distortion components produced by nonlinear effects within the transistor. This conclusion originated from empirical observations made on an analog computer simulation of a transistor. An analytic argument reinforces this conclusion by comparing plots of algebraic cancellation to measured distortion curves. Also, we give experimental support of the cancellation phenomenon.

(*ii*) We present a low-frequency analysis of a complete extrinsic model including three nonlinearities: emitter resistance, nonlinear current gain, and avalanche multiplication, all approximated by a third order polynominal. We avoid considerable complexity by directing the analysis strictly to the question of minimizing distortion and by not developing a general distortion expression. This analysis is independent of any assumptions concerning distortion cancellation, but yields the same results.

(*iii*) From this analysis we show that it is possible to *simultaneously* null both second and third order distortion under the same bias-load condition. The analytic technique we use to determine a null is linearization of the input-output relation up to and including third order, thus implying a minimum in harmonic distortion, intermodulation, or any other specialized figure of merit. The existence of this simultaneous null is verified in the laboratory.

(*iv*) Extension of the cancellation effect to a broad band of frequencies can be accomplished by external reactive compensation. This compensation maintains a 180° phase shift between the collector-base voltage and the real component of the emitter current, a relation that exists automatically at low frequencies where the rigorous analysis is performed. This phase shift is the fundamental requirement for total cancellation, based on the qualitative insight mentioned in item *i*.

PRINCIPAL SYMBOLS

A	Parameter in the $\beta(I_c)$ relation.
$\alpha(I_e)$	Current dependence of the dependent current source.
α_1 , α_2 , α_3	Taylor series coefficients in the expansion of $\alpha(I_e)$ around I_{eo} .
α_{max}	Maximum value of α with respect to I_e .
β	Common emitter ac gain.
β_{max}	Maximum value of β with respect to I_c .
I_c	Total collector current.
I_{cp}	Collector current where β_{max} occurs.
i_c	Small signal collector current.
I_e	Total emitter current.
I_{eo}	Emitter current bias level.
i_e	Small signal emitter current.
I_o	Collector current bias level.
I_{out}	Current in the load resistor.

$M(V_{cb})$	Voltage dependence of the dependent current source.
M_1, M_2, M_3	Taylor series coefficients in the expansion of $M(V_{cb})$ around V_o.
r_e	Emitter resistance.
r_1, r_2, r_3	Taylor series coefficients relating v_e to i_e.
R_L	Load resistance.
$R_{(2)\ opt}$	Load which minimizes second order distortion when third order distortion is negligible.
R_s	Source resistance.
V_{cb}	Total collector-to-base voltage.
v_{cb}	Small signal collector-to-base voltage.
v_e	Small signal voltage across r_e.
V_o	Collector-to-base bias level.
V_{out}	Voltage across the load resistor.

II. A QUALITATIVE MODEL FOR THE DISTORTION MECHANISM

Let us describe the qualitative insight (i) to get a broad look at the cancellation phenomenon before rigorous analysis obscures a simple concept.

An analog computer simulation of the model of Fig. 1 allows us to examine the interaction of the three nonlinearities by examining their effects one at a time. Thus, for example, we may allow only $\alpha(I_e)$ to be nonlinear and observe the second harmonic distortion components of the output voltage. If we then make α constant and allow $M(V_{cb})$ to vary, we observe that the resulting waveform is 180° out of phase with the first waveform as shown in Fig. 2. This is plausible since V_{cb} and I_e are inherently 180° out of phase at low frequencies. Thus any cancellation that we obtain between current dependent nonlinearities

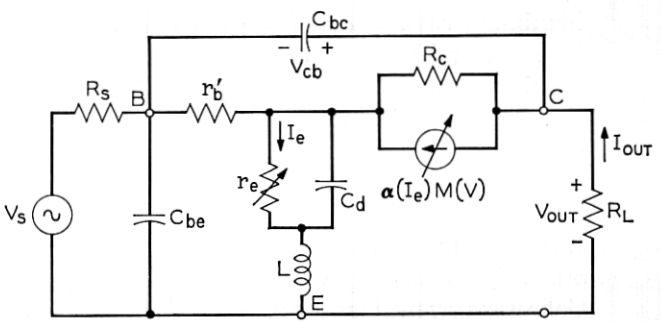

Fig. 1 — High frequency nonlinear model.

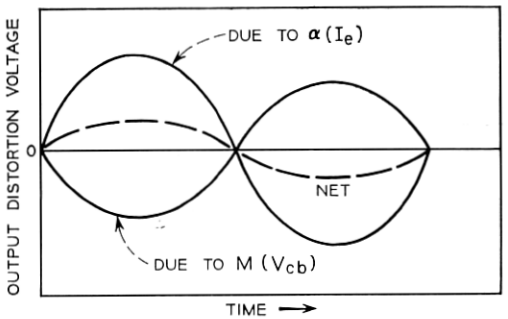

Fig. 2 — Cancellation of distortion components.

and voltage dependent nonlinearities will *not* require *phasing* I_e and V_{cb} properly, but will result from properly adjusting the relative magnitudes of V_{cb} and I_e.

Since $I_e \approx I_{out}$, the most direct way to adjust the magnitude of $V_{cb} \approx V_{out}$ relative to I_e is to change the load resistance. Hence the strong dependence of distortion on R_L as shown in Fig. 3 for a fixed bias level of $V_{cb} = V_o$ and $I_{out} = I_o$. To obtain cancellation in second and third order distortion at the same time, not only the relative magnitudes are important but the absolute level must be correct. This cancellation model explains the sharpness of the null: since the net distortion is a small difference between large distortion components, a small percentage change in the ratio of the larger components will yield a large percentage change in the difference. Experimentally, as a null is passed the output distortion waveform changes phase by 180° as we would expect from one component's becoming dominant over the other.

It is important to notice that this cancellation effect is not some artificial phenomenon that we are forcing to occur. According to the model presented here, some degree of cancellation always occurs in any transistor at any level of distortion. We give a more quantitative argument supporting this exact cancellation model for visualizing the transistor distortion mechanism in Appendix C.

It has been the author's experience that a disturbingly large percentage of published technical material is exclusively concerned with presenting conclusions. In most cases, these conclusions were arrived at by the rigorous manipulation of symbols long after the original insight which prompted the investigation. The purpose of this section is to describe the insights first in the belief that the reader will have

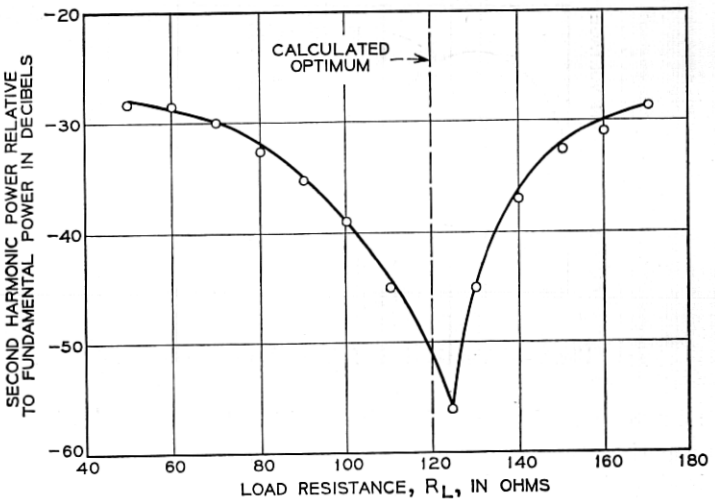

Fig. 3 — Experimental null in second harmonic distortion as a function of R_L, using a Western Electric 20J transistor with $V_e = 30$ volts, $I_e = 100$ milliamperes, and $R_s = 500$ ohms.

at least one less handicap if he is allowed to see the simple ideas on which the rather interesting conclusions of this paper are based. These ideas are:

(*i*) The nonlinearities of the transistor (including some, such as the base spreading resistance and the diffusion capacitance, which are not considered in this paper) are dependent on the emitter current, I_e, and the collector-base voltage, V_{cb}. At low frequencies I_e and V_{cb} are 180° out of phase.

(*ii*) As a result of this phase difference, distortion components resulting from these independent variables will subtract at low frequencies.

(*iii*) On an analog computer simulation, we observe the ability to extend this subtraction effect to the extent of total cancellation by manipulating external circuit parameters. Thus it should be possible to analyze a low frequency model by imposing the condition of zero distortion and solve for the required circuit parameters. We would expect the load resistance to be an important parameter in this analysis since it determines the ratio of V_{cb} to I_e.

(*iv*) Considering the low frequency phase difference between I_e and V_{cb} as the most important factor in achieving total cancellation, we suggest a technique for extending the low frequency results to a broad

band of frequencies. This extension is achieved by the simple expedient of compensating the load to achieve a constant real part, R_L, and still maintain the proper phase between V_{cb} and I_e as frequency increases.

The following sections develop the rigorous analysis (most of which is relegated to Appendix A) and examine in some detail the analytic conditions for a null and the implications of these conditions in the area of circuit and device design.

III. TRANSISTOR MODEL

The model in Fig. 1 has been matched closely to the h parameters of the Western Electric type 46A transistor over a wide range of frequency (5 to 100 mHz) and bias current (50 to 150 mA). Figs. 4 and 5 show a typical match, obtained from a general purpose optimization program. Three distinct nonlinear effects were then superimposed on this small signal linear skeleton model. The current dependence of the dependent current source is changed from αI_e to the expansion around the emitter current bias point, I_{eo},

$$\alpha(I_e) = I_o + \alpha_1(I_e - I_{eo}) + \tfrac{1}{2}\alpha_2(I_e - I_{eo})^2 + \tfrac{1}{6}\alpha_3(I_e - I_{eo})^3 + \cdots \quad (1)$$

where I_o is the quiescent collector current. The voltage dependence of the dependent current source is changed from the constant, M, to the expansion around the collector-to-base bias voltage, V_o,

$$M(V_{cb}) = 1 + M_1(V_{cb} - V_o) + \tfrac{1}{2}M_2(V_{cb} - V_o)^2 + \tfrac{1}{6}M_3(V_{cb} - V_o)^3 + \cdots \quad (2a)$$

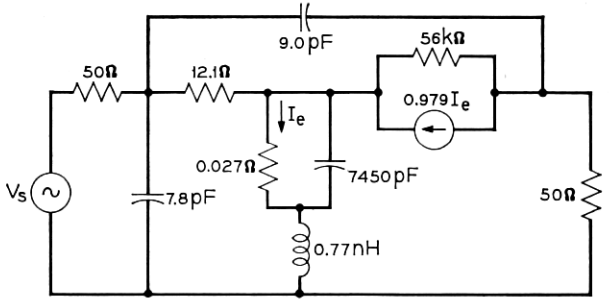

Fig. 4 — Linear model for $I_e = 150$ mA, $V_{cb} = 10$V. With the indicated element values, this model matches the measured h-parameters shown in Fig. 5. The quality of the match at this bias point ($I_e = 150$ mA, $V_{cb} = 10$V) is typical.

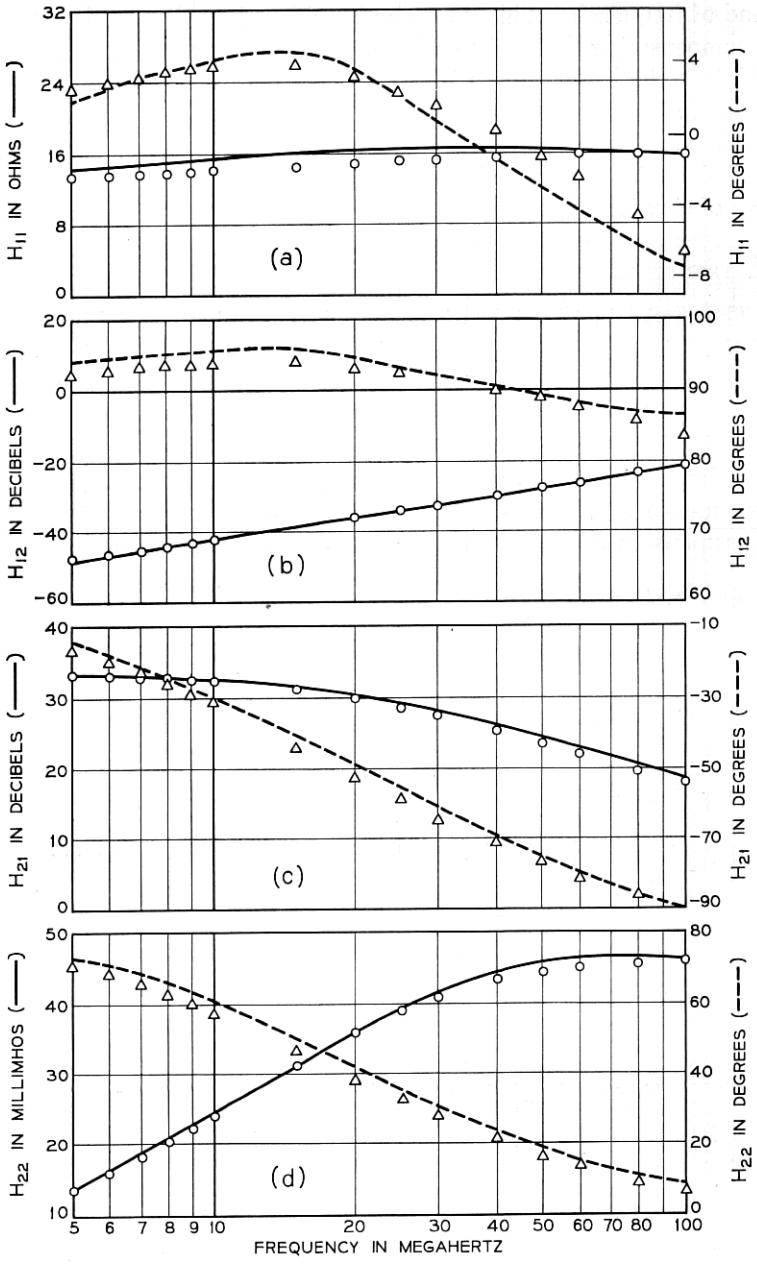

Fig. 5 — Magnitude and phase, as a function of frequency match to (a) h_{11}, (b) h_{12}, (c) h_{21}, and (d) h_{22}. Solid lines were measured; points were plotted from the model.

So that the total dependent current source relationship is

$$I_c = \alpha(I_e)M(V_{cb}). \tag{2b}$$

And finally the emitter resistance, r_e, is replaced by the expansion around I_{eo}

$$v_e = r_1(I_e - I_{eo}) + \tfrac{1}{2}r_2(I_e - I_{eo})^2 + \tfrac{1}{6}r_3(I_e - I_{eo})^3 + \cdots \tag{3}$$

where the coefficients in equations 1, 2, and 3 are the corresponding derivatives of the Taylor series expansion. Define the small signal quantities as

$$i_c = I_c - I_o \tag{4}$$

$$i_e = I_e - I_{eo} \tag{5}$$

$$v_{cb} = V_{cb} - V_o\,. \tag{6}$$

Using these relations, equations 1, 2, and 3 become

$$\alpha(I_e) = I_o + \alpha_1 i_e + \tfrac{1}{2}\alpha_2 i_e^2 + \tfrac{1}{6}\alpha_3 i_e^3 + \cdots \tag{7}$$

$$M(V_{cb}) = 1 + M_1 v_{cb} + \tfrac{1}{2}M_2 v_{cb}^2 + \tfrac{1}{6}M_3 v_{cb}^3 + \cdots \tag{8}$$

$$v_e = r_1 i_e + \tfrac{1}{2}r_2 i_e^2 + \tfrac{1}{6}r_3 i_e^3 + \cdots \tag{9}$$

Substituting equations 7 and 8 into 2b, and retaining third order terms

$$I_c = (I_o + \alpha_1 i_e + \tfrac{1}{2}\alpha_2 i_e^2 + \tfrac{1}{6}\alpha_3 i_e^3) \cdot(1 + M_1 v_{cb} + \tfrac{1}{2}M_2 v_{cb}^2 + \tfrac{1}{6}M_3 v_{cb}^3) \tag{10}$$

$$I_c - I_o = i_c = \alpha_1 i_e + I_o M_1 v_{cb} + \tfrac{1}{2}\alpha_2 i_e^2 + \tfrac{1}{2}I_o M_2 v_{cb}^2 + \alpha_1 M_1 i_e v_{cb} + \tfrac{1}{2}\alpha_1 M_2 v_{cb}^2 i_e + \tfrac{1}{2}\alpha_2 M_1 v_{cb} i_e^2 + \tfrac{1}{6}\alpha_3 i_e^3 + \tfrac{1}{6}I_o M_3 v_{cb}^3 \tag{11}$$

At this point we have developed a model for the transistor, indicating the nature and form of the particular nonlinearities considered in both the analog computer simulation of the complete, frequency-dependent model of Fig. 1 and the analysis of the dc model of Fig. 6.

IV. THE ANALYSIS

4.1 *Optimization Equations*

An analog computer simulation of the complete, frequency-dependent model just discussed suggests that a simpler model is sufficient to describe the distortion characteristics of the transistor at low

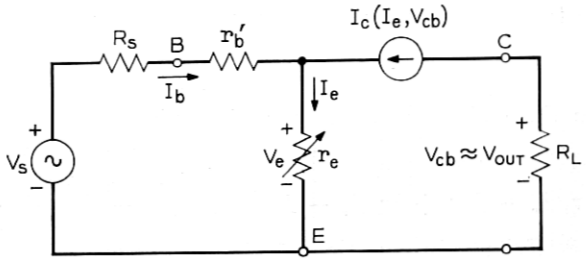

Fig. 6 — Low frequency nonlinear model.

frequencies. Fig. 6 shows this simplified dc model. The following analysis of this model is detailed in Appendix A.

(*i*) The incremental output voltage, v_{out}, is related to the input voltage, V_S, retaining third order terms as in equation 11.

(*ii*) This input-output relation is constrained to be linear, thus forcing both second and third order distortion to zero.

(*iii*) This constraint requires certain coefficients in the nonlinear v_{out} (V_S) relation to be zero. These coefficients are, of course, functions of the linear and nonlinear parameters of the system. Thus, when these functions are made zero, v_{out} is a linear function of V_S (to third order), and the derivation of the optimization equations is complete. These equations are:

$$-R_L^2(I_oM_2) + R_L(2\alpha_1M_1) + \delta - \alpha_2 = 0 \tag{12}$$

$$R_L^3(I_oM_3) - R_L^2(3\alpha_1M_2) + \xi - \alpha_3 = 0 \tag{13}$$

where

$$\delta = r_2/(R_s + r_b') < 0, \quad \text{since} \quad r_2 < 0, \quad \text{(equation 19)}, \tag{14}$$

and

$$\xi = r_3/(R_s + r_b') > 0, \quad \text{since} \quad r_3 > 0 \quad \text{(equation 20)}. \tag{15}$$

For the simpler case where the amplitude of third order distortion is sufficiently low so that third order terms are negligible, equation 13 is satisfied identically and only equation 12 remains, which is easily solved to yield

$$R_{(2)\,opt} = \alpha_1M_1/I_oM_2 + [(\alpha_1M_1/I_oM_2)^2 - (\alpha_2 - \delta)/I_oM_2]^{\frac{1}{2}}. \tag{16}$$

Thus $R_{(2)opt}$ is the value of load resistance which causes second order distortion to be zero for the case where third order terms are

negligible. Notice that the analytic technique used to determine a distortion null here is linearization of the input-output relation, and thus implies a minimum in harmonic distortion, intermodulation distortion, or any other specialized figure of merit. Of course, $R_{(2)\text{opt}}$ is a function of bias current and voltage because of the dependence of M_1 and M_2 on voltage and r_2, α_1 and α_2 on current. The implications of equations 12, 13, and 16 become more clear when the dependence of these parameters on bias is considered.

4.2 *Relating Parameters to More Directly Measurable Quantities*

It is revealing to express the parameters of equations 12 and 13 in terms of the bias variables and other directly measurable parameters of the transistor.

Assuming the standard exponential i-v relation at the emitter-base junction we can immediately derive from

$$I_e = I_s[\exp(\lambda q V_e/kT) - 1] \tag{17}$$

the following relations:

$$r_1 = kT/\lambda q I_o = r_o/I_o , \tag{18}$$

$$r_2 = -kT/\lambda q I_o^2 = -r_o/I_o^2 , \tag{19}$$

$$r_3 = 2kT/\lambda q I_o^3 = 2r_o/I_o^3 . \tag{20}$$

Similarly, if we assume that the avalanche effect in the common-emitter mode is described by an equation of the same form as Miller's[5]

$$M(V_{cb}) = [1 - (V_{cb}/V_A)^n]^{-1} \tag{21}$$

where V_A is the common-emitter breakdown voltage as shown in Fig. 7. Then, at $V_{cb} = V_o$:

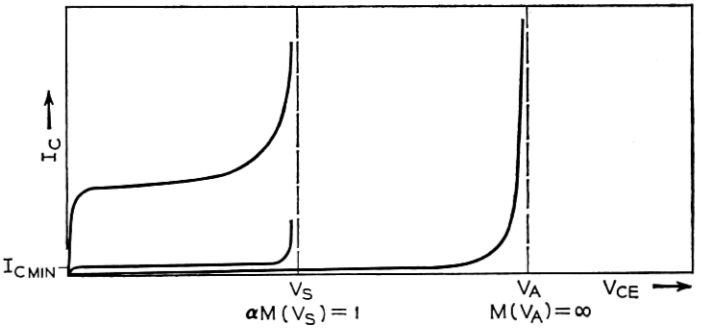

Fig. 7 — Avalanche characteristics.

$$M_1 = n(V_o/V_A)^n/V_o\,, \tag{22}$$

$$M_2 = n(n-1)(V_o/V_A)^n/V_o^2\,, \tag{23}$$

$$M_3 = n(n-1)(n-2)(V_o/V_A)^n/V_o^3\,. \tag{24}$$

The avalanche voltage, V_A, can be determined on a curve tracer oscilloscope by leaving the emitter open-circuited in a grounded base configuration and sweeping the collector-base voltage. The sustaining voltage, V_S, shown in Fig. 7, is obtaining with the transistor in the common-emitter mode and at least enough base current flowing to produce $I_{C\min}$ at the output. At V_S the avalanche factor $M(V_{cb})$ has increased above unity sufficiently so that $\alpha(I_C)M(V_S) = 1$. As a result the common-emitter current gain (β) at this voltage is infinite. Choosing the smallest α at which this occurs ($\alpha_{\min}$) allows us to determine the exponent, n, in equation 21:

$$\alpha_{\min}M(V_S) = 1 = \alpha_{\min}[1 - (V_S/V_A)^n]^{-1}. \tag{25}$$

Therefore

$$n \approx \log \beta_{\min}/\log\,(V_A/V_S), \tag{26}$$

where $\beta_{\min}$ corresponds to $\alpha_{\min}$ and may be determined from equation 27 using $I_o = I_{C\min}$. Notice that equation 21 constitutes an empirical relationship in this study and is not intended to be rigorously tied to any one of the various avalanche mechanisms. It is apparent, too, that the measurements determining equations 25 and 26 will be influenced by other voltage-dependent mechanisms (for example, the Early effect); hence they are not strictly related to the avalanche multiplication effect alone. Equation 21 has the virtue of mathematical tractability; equation 25 allows the parameters of 21 to be determined conveniently; and, finally, the excellent experimental agreement with the theory described in Section V provides adequate justification of the original assumptions. In any case, the derivation of equations 12 and 13 is based on a general power series expansion for $M(V_{cb})$ around V_o; hence it remains valid for any M_1, M_2, and M_3.

Finally we require α_2 and α_3. We show in Appendix B that β can be empirically related to collector bias by

$$\beta \cong \beta_{\max}/[1 + A\,\ln^2\,(I_o/I_{cp})] \tag{27}$$

where $\beta_{\max}$ is the maximum β which occurs at $I_o = I_{cp}$, as shown in Fig. 8, and A is a parameter of the equation. Determination of α_2,

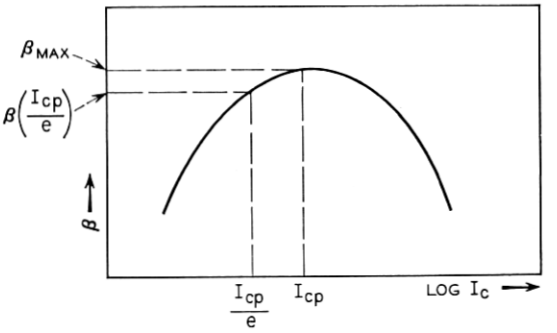

Fig. 8 — Current gain nonlinearity as a function of the bias current, I_o.

α_3, and A is derived in Appendix B. They can be expressed as

$$\alpha_2 \cong -\left(\frac{4A\alpha_{max}^2}{\beta_{max} I_o}\right) \ln\left(\frac{I_o}{I_{cp}}\right) \tag{28}$$

$$\alpha_3 \cong \left(\frac{6A\alpha_{max}^2}{\beta_{max} I_o^2}\right) \ln\left(\frac{I_o}{eI_{cp}}\right) \tag{29}$$

$$A = \left[\beta_{max} - \beta\left(\frac{I_{cp}}{e}\right)\right] / \beta_{max} \tag{30}$$

so that A may be determined by finding I_{cp} and measuring β at that current and at $1/e$ times that current. Thus equations 18 through 30 give the functional relations for the various parameters in equations 12 and 13 and indicate the method of measuring the more fundamental parameters such as n and A. In the next section we use these relations in existence conditions for a simultaneous null, in order to guide an experimental search for this condition.

4.3 *Existence Conditions for Realizability*

While the simultaneous solution of equations 12 and 13 has not been accomplished in closed form, it is possible to derive the conditions under which a solution exists. Expressed in terms of the bias variables, such conditions can then be used as a guide in an experimental search for simultaneous nulling of second and third order distortion.

Basically we require R_L to be real and positive. For the second order equation, solved in equation 16, this simply requires that

$$(\alpha_1 M_1 / I_o M_2)^2 \geqq (\alpha_2 - \delta)/I_o M_2 . \tag{31}$$

The condition for the existence of a positive, real solution to a cubic of the form

$$x^3 - px^2 + r = 0 \tag{32}$$

where

$$p = 3\alpha_1 M_2 / I_o M_3 \tag{33}$$

$$r = (\xi - \alpha_3)/I_o M_3 \tag{34}$$

$$x = R_L \tag{35}$$

is easily derived. Basically require

$$x^3 = px^2 - r. \tag{36}$$

Now, from equations 34 and 29, $r > 0$ for $I_o < eI_{cp}$. Thus, at $x = 0$, the parabola on the right side of equation 36 will be below the cubic on the left. There will be a positive intersection only if the equation is satisfied before the cubic term begins increasing more rapidly (larger slope) than the parabola. The slopes are equal at

$$x_1 = \tfrac{2}{3}p. \tag{37}$$

Therefore require

$$x_1^3 \leqq px_1^2 - r \tag{38}$$

or

$$\tfrac{4}{27}p^3 \geqq r. \tag{39}$$

Expressing this existence condition in terms of the problem variables and rearranging terms gives

$$(\alpha_1 M_2)^3 \geqq \tfrac{1}{4}(\xi - \alpha_3)(I_o M_3)^2. \tag{40}$$

Substituting in equations 31 and 40 with 18 through 29 and arranging terms we obtain

$$(V_o/V_A)^n > \text{the Greater of } [Q_1, Q_2] \tag{41}$$

where

$$Q_1 = \left(1 - \frac{1}{n}\right)\left[\frac{r_o}{I_o}(R_s + r_b') - \left(\frac{4A\alpha_{\max}^2}{\beta_{\max}}\right)\ln\left(\frac{I_o}{I_{cp}}\right)\right] \tag{42}$$

$$Q_2 = \left[\frac{r_o}{I_o}(R_s + r_b') - \left(\frac{3A\alpha_{\max}^2}{\beta_{\max}}\right)\ln\left(\frac{I_o}{eI_{cp}}\right)\right]\frac{(n-2)^2}{2n(n-1)}. \tag{43}$$

For most ranges of parameters and bias variables $Q_1 > Q_2$, thus, we will examine the condition

$$\left(\frac{V_o}{V_A}\right)^n > \left(1 - \frac{1}{n}\right)\left[\frac{r_o}{I_o}(R_s + r_b') - \left(\frac{4A\alpha_{\max}^2}{\beta_{\max}}\right)\ln\left(\frac{I_o}{I_{cp}}\right)\right] \quad (44)$$

in greater detail. This distinction between Q_1 and Q_2 is not critical, however, because they are similar in form. Thus, many of the qualitative considerations to be developed in the next section are the same for Q_1 or Q_2.

4.4 *Searching for a Simultaneous Null*

A careful examination of the existence condition (44) is useful in guiding an experimental search for a simultaneous null. Starting at the left side of the inequality, it is obvious that the bias voltage, V_o, must be as large as possible relative to V_A. Since, in any case, $V_o < V_A$, the exponent, n, should be as small as possible. The value of n, according to Rogers,[5] depends on whether the collector or base has the higher resistivity, and whether the high resistivity side is n or p type.

Where the collector has the higher resistivity, the lowest values of n are for npn silicon, and for pnp germanium. A second, less important, advantage of small n is that the multiplier on the right side of the inequality is reduced. The first term in the brackets tends to be the major contributor to the right side of the inequality and is therefore the term which is most desirable to reduce. This term, which represents input distortion resulting from a nonlinear emitter resistance, can be reduced by increasing the bias, I_o, and by increasing R_s to approximate a current source drive, thereby reducing input distortion.

The second term in the brackets will favorably reduce the right side of the inequality only if the logarithm is positive. This will be true if the bias current, I_o, is greater than I_{cp}, which is consistent with the earlier requirement for a larger I_o. Finally, the multiplier λ, in equation 18 should be small in order to reduce r_o.

Thus, it appears that the most likely candidate for a simultaneous null is a silicon power transistor to allow large values of I_o and V_o. The structure should be either pnp or npn, depending on which type gives the smaller n.

V. EXPERIMENTAL PROCEDURE AND RESULTS

Let us illustrate the application of these existence conditions in an experimental determination of $R_{(2)\text{opt}}$ as well as a simultaneous null in second and third order distortion.

As a fundamental check on the theoretical results, we decided to determine the accuracy of equation 16 with 19, 22, 23, and 28 substituted for the Taylor series coefficients. Also, it was desirable to verify the existence of a simultaneous null using the existence conditions of the previous section. Because of the low frequencies involved (input frequency of 1 kHz), the simplest approach was to simulate the measurement apparatus on the analog computer, using the same oscillator and bandpass filters already available on the original simulation.[6] The transistor used was the Western Electric 20J, and npn power transistor.

Using this equipment, the parameters of the $\beta(I_c)$ characteristic curve of the transistor were measured:

$$\beta_{max} = 78$$

$$I_{op} = 15mA$$

$$\beta(I_{cp}/e) = 73.$$

From a curve tracer oscilloscope, the avalanche parameters were determined:

$$V_A = 60V$$

$$V_S = 35V$$

$$\beta_{min} = 45.$$

These measurements yield the information to compute

$$n = 7$$

$$A = .064$$

from equations 26 and 30. From the manufacturer's data, $r_e = 50$ mV and $r_b' = 50\ \Omega$. The output power was maintained at one watt.

These parameters give all the information required by equation 16 to compute the function $R_{(2)\,opt}(I_e)$ for various values of V_e. The curves in Figs. 9 and 10 show this computation compared to the plotted points which were measured. The agreement here is quite adequate. The quality of the match is further emphasized by comparing the computed values of $R_{(2)\,opt}$ indicated in Fig. 3 and Fig. 11 to the measured nulls. The computed value shown in Fig. 11 is based on a solution to equation 16 only.

A typical simultaneous null obtained in the laboratory is shown in Fig. 11. This data indicates the high voltages (to emphasize avalanche

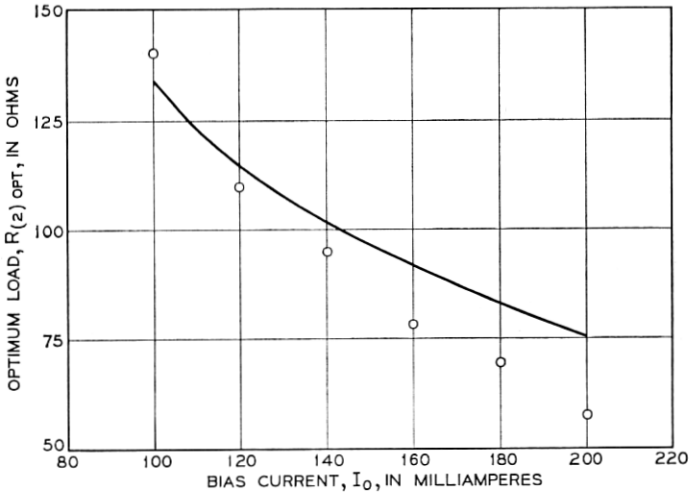

Fig. 9 — Measured (plotted points) and computed values (curve) of $R_{(2)opt}$ as a function of bias current, I_o, using a Western Electric 20J transistor with $R_s = 500$ ohms and $V_o = 29$ volts.

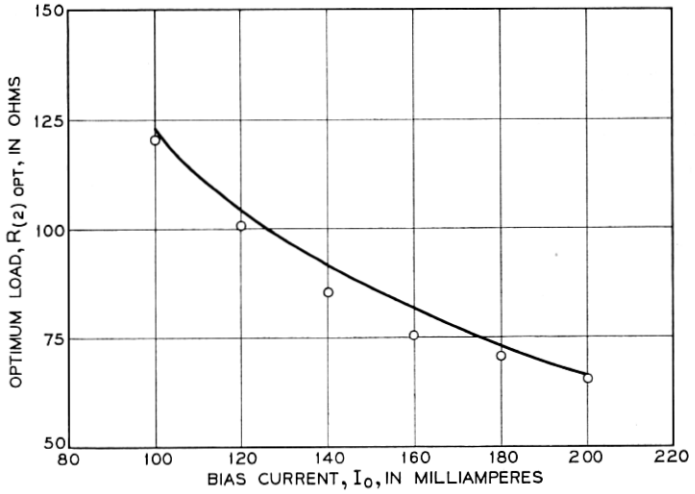

Fig. 10 — Measured (plotted points) and computed values (curve) of $R_{(2)opt}$ as a function of bias current, I_o, using a Western Electric 20J transistor with $R_s = 500$ ohms and $V_o = 25$ volts.

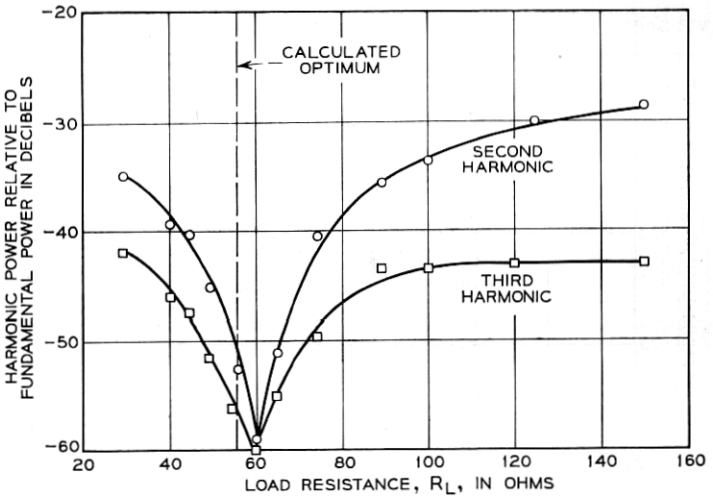

Fig. 11 — Simultaneous null in second and third harmonics as a function of load, R_L, with the same transistor and R_s as in Figs. 9 and 10. Here $V_o = 29$ volts and $I_o = 240$ mA.

distortion) and currents (to minimize input distortion) required for a simultaneous null. Conditions for a simultaneous null exist on an (R_L, I_o, V_o, R_s) surface, giving some redundant control to achieve desired power and impedance levels as well as minimum second and third harmonic distortion. It is apparent that a transistor manufactured with a lower value of n would allow a broader range of control over the bias voltage and current level required. Measurements on different units of the WE20J show a maximum spread of ±10 per cent in measured values of the optimum load for a simultaneous null.

Experimentally, as R_L is varied, the second harmonic displayed on the oscilloscope decreases in amplitude, goes to zero, and begins to increase in amplitude. As it goes through a null, the second harmonic changes sign, giving additional weight to the qualitative distortion model discussed in Section II.

VI. EXTENSION OF CANCELLATION TO A BAND OF FREQUENCIES

Up to this point, our discussion has been limited to low frequency effects. Now let us consider why the above results do not apply at high frequencies and look at a straightforward approach to extend the validity of all previous results to a broad band of frequencies.

If we accept the qualitative picture given in Section II, it is obvious that we should not expect to maintain exact cancellation as frequency increases, since the phase of V_{cb} relative to I_e will change. A small change in phase will have an effect similar to changing the relative magnitudes of the current and voltage-dependent distortion components: the amplitude of their difference (the net distortion) will change by a large percentage near a null. In fact, at higher frequencies (on the order of $f_T/100$), the null of Fig. 2 vanishes altogether. It is apparent, then, that a solution to this problem is to apply external reactive compensation in such a way as to keep V_{cb} and I_e 180° out of phase as frequency increases. In the model shown by Fig. 6 if we consider a capacitor, C_D, in parallel with r_e, it is straightforward to derive the relation,

$$-V_{cb}/I_e \cong \text{Re}\,\{Z_L\} + j\left[\frac{\omega r_b'}{\omega_T} + \text{Im}\,\{Z_L\}\right], \tag{45}$$

where

$$\omega_T = \frac{1}{C_D r_e}. \tag{46}$$

Ideally, we would desire $Z_L = R_L - j\omega r_b'/\omega_T$ but this would require a negative inductor. A simple first order approximation to this function would be to parallel R_L with a capacitor, C. Then

$$Z_L = \frac{R_L}{1 + \omega^2 R_L^2 C^2} - j\omega \frac{R_L^2 C}{1 + \omega^2 R_L^2 C^2}. \tag{47}$$

From equation 47 choose

$$C_{\text{opt}} = r_b'/R_L^2\omega_T\,. \tag{48}$$

Now

$$-\frac{V}{I_e} = \frac{R_L}{1 + \omega^2 R_L^2 C_{\text{opt}}^2} + j\frac{\omega^3}{\omega_T}\left[\frac{r_b' R_L^2 C_{\text{opt}}^2}{1 + \omega^2 R_L^2 C_{\text{opt}}^2}\right]. \tag{49}$$

For small angles the phase is given by

$$\varphi(\omega) = \omega^3 (r_b'/R_L\omega_T)^3. \tag{50}$$

Thus the phase is reduced below the uncompensated case up to the frequency

$$\omega_{\max} = \frac{R_L\omega_T}{r_b'}.$$

At which point the cubic dependence of equation 50 intersects the linear phase of the uncompensated transistor.

Obviously additional compensating elements can be used to cause higher derivatives of $\varphi(\omega)$ to be zero. A complication may arise if C_{opt} is less than the parasitic C_{CE} of the transistor. In this case we extend the required low-pass structure of the compensating network to include an inductor in series with R_L. In this case the required inductance is given by

$$L_{opt} = R_L^2 C_{CE} - r_b'/\omega_T$$

which is greater than zero for $C_{CE} > r_b'/R_L^2\omega_T = C_{opt}$.

VII. CONCLUSION

Our conclusions are based on simulation of the transistor on an analog computer, analysis, and experimental results. The rigorous analysis predicts the existence of a simultaneous null in second and third harmonic distortion under the same bias-load conditions. This null has been observed in the laboratory. In addition, experiments on the simulation provide qualitative insight into the nature of the distortion mechanism.

We conclude that this mechanism consists of the algebraic subtraction, at low frequencies, of distortion components from various sources within the transistor such as the nonlinear emitter resistance, current gain, and avalanche multiplication effect. This interaction between distortion components yields a net distortion which is the difference between the contributing components, and can be made zero by a proper choice of the bias and load.

With this mechanism in mind we developed a technique for extending the cancellation phenomenon to a broad band of frequencies. This technique consists of external reactive compensation which maintains 180° phase shift between the distortion components, a condition which exists inherently at low frequencies.

We have obtained experimental confirmation of the theoretical dependence of the optimum load for second order distortion on bias variables. The theory predicting a simultaneous null in second and third order distortion has been confirmed. We have also obtained experimental support for the distortion cancellation phenomenon. We discussed methods to aid future measurement efforts in implementing this distortion reduction phenomenon. These methods are based on interpretation of the theoretical expressions developed in the paper

which reveal the necessity for high levels of bias voltage and current to obtain a simultaneous null.

This study opens several fruitful areas for future work, both in device and circuit areas. Primarily, the phenomenon described uses circuit techniques to minimize distortion (optimizing the bias-load point). Additional effort in the circuit aspects of minimizing distortion should be directed toward desensitizing the null condition to variations in the bias-load point. For example, if the bias current is forced to change with R_L as shown in Fig. 9, optimum conditions could be maintained over a range of changes in the load.

In the realm of device design, effort should be directed toward adjusting device parameters to allow nulling in useful regions of the bias-load space. For example, a softer avalanche characteristic (lower value of n) would allow the use of lower bias voltages.

ACKNOWLEDGMENT

In a long term study of this kind, a large number of people share in the effort. There are certain individuals whose critical contributions are gratefully acknowledged.

Primarily, the author thanks S. Narayanan, who has analyzed the distortion problem using the Volterra Series representation which is amenable to digital simulation.[7]

J. S. T. Huang did much of the ground work on which the simulation is based and has been an invaluable consultant since that time.

Miss R. Mayorkas, M. J. Magelnicki, and R. Ukeiley took measurements for the linear and nonlinear characterization and Miss D. M. Bohling and Mrs. J. L. Murray helped with digital computer calculations. N. J. Chaplin contributed valued data and comments on initial results.

REFERENCES

1. Akgun, M., and Strutt, M. J. O., "Cross Modulation and Nonlinear Distortion in RF Transistor Amplifiers," IRE Trans. Elec. Devices, *6* (October 1959), pp. 457–469.
2. Mallinckrodt, A. J., and Gardner, F. M., "Distortion in Transistor Amplifiers," IEEE Trans. Elec. Devices, *10* (July 1963), pp. 288–289.
3. Riva, G. M., Beneteau, P. J., and Dalla Volta, E., "Amplitude Distortion in Transistor Amplifiers," Proc. IEE, *111* (March 1964), pp. 481–490.
4. Reynolds, J., "Nonlinear Distortions and their cancellation in Transistors," IEEE Trans. Elec. Devices, *16*, No. 11 (November 1965), pp. 595–599.
5. Phillips, A. B., *Transistor Engineering*, New York: McGraw-Hill, 1962, pp. 136–137.
6. Thomas, Lee C., "An Application of the Analog Computer to Electronic Circuit Simulation," IEEE Trans. Elec. Comp., *16*, No. 4 (August 1967).

7. Narayanan, S., "Transistor Distortion Analysis Using Volterra Series Representation," B.S.T.J., *46*, No. 5 (May–June 1967) pp. 991–1024.

APPENDIX A

Derivation of the Optimization Equations

In Fig. 6 the following relations hold

$$i_e = i_b + i_c \tag{51}$$

$$i_b = (V_s - v_e)/(R_s + r_b') \tag{52}$$

$$v_{out} = -R_L i_c \approx v_{cb} . \tag{53}$$

Let

$$V_s/(R_s + r_b') = I \tag{54}$$

$$r_1/(R_s + r_b') = \gamma \tag{55}$$

$$r_2/(R_s + r_b') = \delta \tag{56}$$

$$r_3/(R_s + r_b') = \xi. \tag{57}$$

Substituting (9) and (54) through (57) in (52)

$$i_b = I - \gamma i_e - \tfrac{1}{2}\delta i_e^2 - \tfrac{1}{6}\xi i_e^3 . \tag{58}$$

Combining (51) and (58)

$$i_c = i_e(1 + \gamma) + \tfrac{1}{2}\delta i_e^2 + \tfrac{1}{6}\xi i_e^3 - I. \tag{59}$$

Now i_c is given by equation 11. Therefore

$$\begin{aligned} i_e(1 + \gamma) + \tfrac{1}{2}\,\delta i_e^2 + \tfrac{1}{6}\xi i_e^3 - I \\ = \alpha_1 i_e + I_o M_1 v_{cb} + \tfrac{1}{2}\alpha_2 i_e^2 + \tfrac{1}{2} I_o M_2 v_{cb}^2 + \alpha_1 M_1 i_e v_{cb} \\ + \tfrac{1}{2}\alpha_1 M_2 v_{cb}^2 i_e + \tfrac{1}{2}\alpha_2 M_1 v_{cb} i_e^2 + \tfrac{1}{6}\alpha_3 i_e^3 + \tfrac{1}{6} I_o M_3 v_{cb}^3 . \end{aligned} \tag{60}$$

Substituting (51) and (53) into (60) and gathering terms:

$$\begin{aligned} & i_c^3[\tfrac{1}{6} I_o M_3 R_L^3 + \tfrac{1}{2}\alpha_2 M_1 R_L - \tfrac{1}{2}\alpha_1 M_2 R_L^2 - \tfrac{1}{6}\alpha_3 + \tfrac{1}{6}\xi] \\ & + i_c^2[\tfrac{1}{2}\,\delta - \tfrac{1}{2}\alpha_2 - \tfrac{1}{2} I_o M_2 R_L^2 + \alpha_1 M_1 R_L \\ & + i_b(\tfrac{1}{2}\xi - \tfrac{1}{2}\alpha_1 M_2 R_L^2 - \tfrac{1}{2}\alpha_2 M_1 R_L - \tfrac{1}{2}\alpha_3)] \\ & + i_c[1 + \gamma - \alpha_1 + I_o M_1 R_L + i_b(\delta - \alpha_2 + \alpha_1 M_1 R_L) \\ & + i_b^2(\tfrac{1}{2}\xi - \tfrac{1}{2}\alpha_3 + \tfrac{1}{2}\alpha_2 M_1 R_L)] + i_b[1 + \gamma - \alpha_1] \\ & + i_b^2[\tfrac{1}{2}\,\delta - \tfrac{1}{2}\alpha_2] + i_b^3[-\tfrac{1}{6}\xi - \tfrac{1}{6}\alpha_3] - I = 0 \end{aligned} \tag{61}$$

The only approximation that we have made up to this point is that $v_{out} \approx v_{cb}$, the collector-to-base voltage, assuming that v_s is small. Now we would like to express the variables of (61) in terms of the independent driving voltage, V_s, and the output current, i_c, which is linearly related to the output voltage. To accomplish this, we start with (58) and make the approximation

$$i_b \approx I - \gamma i_c \tag{62}$$

where we have ignored the high order terms in (58) and used the linear relation $i_c \approx i_s$.

Notice that (62) certainly does not imply that we have fixed a linear relationship between i_b, I, and i_c. We are simply using this new approximate variable in the highly nonlinear (61) for convenience. The approximation is justified by the fact that second order and higher terms ignored in (62) would appear as fourth order and higher terms in (61).

Substituting (62) into (61) we have

$$\begin{aligned}
&i_c^3\{\tfrac{1}{6}I_oM_3R_L^3 - \tfrac{1}{6}\alpha_3(1-\gamma)^3 \\
&\quad + \tfrac{1}{6}\xi(1-\gamma)^3 + \tfrac{1}{2}\alpha_2M_1R_L(1-\gamma)^2 - \tfrac{1}{2}\alpha_1M_2R_L^2(1-\gamma)\} \\
&\quad + i_c^2\{\tfrac{1}{2}\,\delta(1-\gamma)^2 + \alpha_1M_1R_L(1-\gamma) - \tfrac{1}{2}I_oM_2R_L^2 - \tfrac{1}{2}\alpha_2(1-\gamma)^2 \\
&\quad + I[\alpha_2M_1R_L(1-\gamma) + \tfrac{1}{2}\xi(1-\gamma)^2 - \tfrac{1}{2}\alpha_3(1-\gamma)^2 - \tfrac{1}{2}\alpha_1M_2R_L^2]\} \\
&\quad + i_c\{(1-\gamma^2) - \alpha(1-\gamma) + I_oM_1R_L \\
&\quad + I[\delta(1-\gamma) - \alpha_2(1-\gamma) + \alpha_1M_1R_L] \\
&\quad + I^2[\tfrac{1}{2}\xi(1-\gamma) - \tfrac{1}{2}\alpha_3(1-\gamma) + \tfrac{1}{2}\alpha_2M_1R_L]\} \\
&\quad + I[\gamma - \alpha_1] + I^2[\tfrac{1}{2}\,\delta - \tfrac{1}{2}\alpha_2] + I^3[\tfrac{1}{6}\xi - \tfrac{1}{6}\alpha_3] = 0.
\end{aligned} \tag{63}$$

At the 100 mA bias levels where we are assumed to be operating, $r_1 \leqq 0.5\ \Omega$. Also $r_b' \approx 10$–$20\ \Omega$ and R_s can only increase the $R_s + r_b'$ sum in (55). Hence $\gamma \ll 1$ and will be ignored in (63). Thus we have effectively substituted I for i_b in (61) to obtain (63). This substitution is *not* justified by requiring the assumption $I \gg \gamma i_c$ in (62) (that is, a current source drive); but is justified on the grounds that the substitution of (62) into (63) did not generate new terms in (63) for $\gamma \ll 1$. Equation (63) is of the form

$$ai_c^3 + i_c^2(b + cI) + i_c(d + eI + fI^2) + gI + hI^2 + jI^3 = 0. \tag{64}$$

Now to force linearity we would like to require

$$v_{out} = kV_s, \qquad \text{where } k \text{ is a constant.} \tag{65}$$

But from (53) and (54), (65) can be expressed in terms of the variables of (64) as

$$i_c = -\frac{k(R_s + r_b')}{R_L} I = BI. \tag{66}$$

Substituting (66) into (64) and gathering terms

$$I^3[aB^3 + cB^2 + fB + j] + I^2[bB^2 + eB + h] + I[dB + g] = 0. \tag{67}$$

Now I is an independent variable so that this equation can hold only if each coefficient is simultaneously zero. In the linear term

$$dB + g = 0 \tag{68}$$

$$B = -\frac{g}{d}.$$

Ignoring terms in γ and noticing that $I_o M_1 R_L \ll 1$ in (63)

$$B \approx \frac{\alpha_1}{1 - \alpha_1 + I_o M_1 R_L}. \tag{69}$$

The constant B should be easy to identify. For small M_1 (low levels of V_o), $B = \beta_1$. However, at the higher values of I_o and V_o, $I_o M_1 R_L$ can be on the order of $(1 - \alpha_1)$. Thus, roughly speaking

$$B \geqq \tfrac{1}{2}\beta \gg 1. \tag{70}$$

Substituting (69) into (67) our final coefficients to be equated to zero in (67) become

$$aB^3 + cB^2 + fB + j = 0 \tag{71}$$

$$bB^2 + eB + h = 0. \tag{72}$$

Substituting for a, c, f, j in (71) by comparison between (64) and (63); ignoring terms in γ:

$$B^3[\tfrac{1}{6}I_o M_3 R_L^3 - \tfrac{1}{6}\alpha_3 + \tfrac{1}{6}\xi + \tfrac{1}{2}\alpha_2 M_1 R_L - \tfrac{1}{2}\alpha_1 M_2 R_L^2] + B^2[\alpha_2 M_1 R_L + \tfrac{1}{2}\xi - \tfrac{1}{2}\alpha_3 - \tfrac{1}{2}\alpha_1 M_2 R_L^2] + B[\tfrac{1}{2}\xi - \tfrac{1}{2}\alpha_3 + \tfrac{1}{2}\alpha_2 M_1 R_L] + [\tfrac{1}{6}\xi - \tfrac{1}{6}\alpha_3] = 0. \tag{73}$$

Gathering terms in R_L:

$$R_L^3[I_o M_3][\tfrac{1}{6}B^3] + R_L^2[-\alpha_1 M_2][B^3 + B^2] + R_L[\alpha_2 M_1][\tfrac{1}{2}B^3 + B^2 + B] + \xi[\tfrac{1}{6}B^3 + \tfrac{1}{2}B + \tfrac{1}{6}] - \alpha_3[\tfrac{1}{6}B^3 + \tfrac{1}{2}B^2 + \tfrac{1}{2}B + \tfrac{1}{6}] = 0 \tag{74}$$

Using (70) we can ignore lower powers of B, and $\alpha_2 M_1$, being the product of second order terms, is very small compared to the other coefficients in (74). Thus (74) becomes

$$(I_o M_3) R_L^3 - R_L^2(3\alpha_1 M_2) + \xi - \alpha_3 = 0. \tag{75}$$

Now substituting for b, e, and h in (72) by comparison between (63) and (64); ignoring terms in γ:

$$B^2[\tfrac{1}{2}\,\delta - \tfrac{1}{2}\alpha_2 + \alpha_1 M_1 R_L - \tfrac{1}{2} I_o M_2 R_L^2] + B[\delta - \alpha_2 + \alpha_1 M_1 R_L] + [\tfrac{1}{2}\,\delta - \tfrac{1}{2}\alpha_2] = 0.$$

Gathering terms in R_L:

$$R_L^2[-I_o M_2][\tfrac{1}{2}B^2] + R_L[\alpha_1 M_1][B^2 + B] + \delta[\tfrac{1}{2}B^2 + \tfrac{1}{2}B] - \alpha_2[\tfrac{1}{2}B^2 + \tfrac{1}{2}B] = 0. \tag{76}$$

Using (71), (76) becomes

$$-(I_o M_2) R_L^2 + (2\alpha_1 M_1) R_L + \delta - \alpha_2 = 0. \tag{77}$$

Equations (75) and (77) are the relations that must be satisfied to satisfy (67), which in turn results from the requirement of a linear input-output relation, (65).

APPENDIX B

Relating Current Gain Nonlinearities to the Bias Current

Riva[3] has shown that the small signal gain of a transistor can be closely matched to an expression of the form

$$\beta = h_{fe\max}[a \log_{10}^2 (I_c/I_{c\max}) + 2a \log_{10} e \log_{10} (I_c/I_{c\max}) + 1]^{-1}. \tag{78}$$

Where

$h_{fe\max}$ = maximum dc current gain
$I_{c\max}$ = collector current bias where $h_{fe\max}$ occurs
a = a constant characteristic of the transistor.

Differentiating the denominator of (78) reveals that the maximum ac current gain ($\beta_{\max}$) occurs for $I_c = I_{c\max}/e$. Call this current I_{cp}. Then

$$\beta = h_{fe\max}[a \log_{10}^2 (I_c/I_{cp}) - a \log_{10}^2 e + 1]^{-1}. \tag{79}$$

At the peak in the $\beta(I_c)$ curve, $I_c = I_{cp}$, and

$$\beta_{\max} = h_{fe\max}/(1 - a \log_{10}^2 e). \tag{80}$$

Substituting for h_{femax} in (79)

$$\beta = \beta_{\max}[1 + a(1 - a \log_{10}^2 e)^{-1} \log_{10}^2 (I_c/I_{cp})]^{-1}. \tag{81}$$

Then, for

$$A = a(1 - a \log_{10}^2 e)^{-1} \log_{10}^2 e \tag{82}$$

$$\beta = \beta_{\max}/[1 + A \ln^2 (I_c/I_{cp})] \tag{83}$$

$$\alpha = \frac{\beta}{1+\beta} = \frac{\alpha_{\max}}{1 + \dfrac{A\alpha_{\max}}{\beta_{\max}} \ln^2 \left(\dfrac{I_c}{I_{cp}}\right)}. \tag{84}$$

Where

$$\Delta I_c = \alpha \, \Delta I_e = \left(\alpha_1 + \frac{d\alpha}{dI_e} \Delta I_e + \frac{1}{2} \frac{d^2\alpha}{dI_e^2} \Delta I_e^2\right) \Delta I_e$$

$$\Delta I_c = \alpha_1 \, \Delta I_e + \frac{d\alpha}{dI_e} \Delta I_e^2 + \frac{1}{2} \frac{d^2\alpha}{dI_e^2} \Delta I_e^3$$

$$= \alpha_1 \, \Delta I_e + \tfrac{1}{2}\alpha_2 \, \Delta I_e^2 + \tfrac{1}{6}\alpha_3 \, \Delta I_e^3 .$$

Thus

$$\alpha_2 = 2 \frac{d\alpha}{dI_e} \tag{85}$$

and

$$\alpha_3 = 3 \frac{d^2\alpha}{dI_e^2} = 1.5 \frac{d\alpha_2}{dI_e}. \tag{86}$$

Now, taking $I_e \approx I_c$, from (84)

$$\alpha_2 = -\frac{4A\alpha_{\max}^2}{I_o\beta_{\max}} \frac{\ln (I_o/I_{cp})}{\left[1 + \dfrac{A\alpha_{\max}}{\beta_{\max}} \ln^2 \left(\dfrac{I_o}{I_{cp}}\right)\right]^2} \tag{87}$$

at $I_c = I_o$. In essentially all cases

$$0.04 < \frac{I_c}{I_{cp}} < 25$$

$$\beta_{\max} > 30$$

$$A\alpha_{\max} < 0.15.$$

Thus, to within 10 per cent in the most extreme case

$$\alpha_2 \cong -(4A\alpha_{\max}^2/\beta_{\max}I_o) \ln (I_o/I_{cp}). \tag{88}$$

Then, from (86)

$$\alpha_3 \cong (6A\alpha_{\max}^2/\beta_{\max}I_o^2) \ln (I_o/eI_{cp}). \tag{89}$$

Now to solve for A, notice that, from (78)

$$\beta\left(\frac{I_{c\,\max}}{e^2}\right) = h_{f\,\max}[a \log_{10}^2 e^2 - 2a \log_{10} e \log_{10} e^2 + 1]^{-1} \quad (90)$$

$$= h_{f\,\max}$$

But from (82) and (90)

$$h_{f\,\max} = \beta_{\max}/(1 + A). \quad (91)$$

Therefore,

$$A = (\beta_{\max} - h_{f\,\max})/h_{f\,\max} \quad (92)$$

where $h_{f\,\max}$ may be measured at

$$I_c = I_{c\,\max}/e^2 = I_{cp}/e. \quad (93)$$

APPENDIX C

The Qualitative Distortion Model

The purpose of this appendix is to support the qualitative picture of algebraic distortion cancellation given in the text. The development here is not intended to be rigorous, but rather to strengthen the reader's ability to share the author's insight into the cancellation mechanism. We have argued that the net distortion current, D, is the algebraic difference between positive and negative distortion current components, A and B, dependent on output voltage and current, respectively. Express this relation as

$$D = A - B. \quad (94)$$

But, for A and B monotonic in voltage and current, the ratio A/B is a measure of the load. Define this measure as

$$R = \frac{A}{B}. \quad (95)$$

Now

$$D = B(R - 1). \quad (96)$$

On a dB basis

$$20 \log \frac{D}{B} = 20 \log | R - 1 |. \quad (97)$$

Fig. 12 is a plot of 20 log $|R\text{-}1|$ as a function of R. Compare this plot with that of Fig. 2, which was measured in the laboratory. The simi-

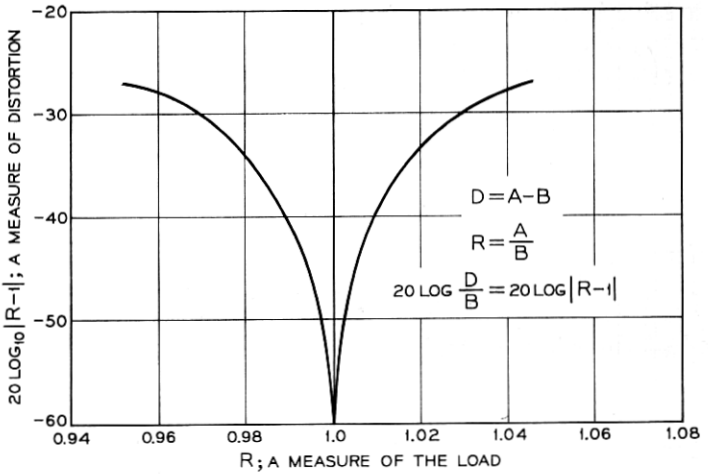

Fig. 12 — Decibel measure of the small difference between large numbers.

larity between the nature of these two minima adds additional weight to the idea that exact algebraic cancellation is involved in producing the net distortion frequencies. Thus any dependence of distortion on frequency should be compensated at distortion frequencies and not at input frequencies, since it is at the distortion frequency that cancellation takes place.