Distortion levels in r.f. clipping

Theoretical intermodulation distortion levels from clipping a two-tone s.s.b. signal

by D. A. Tong, B.Sc. Ph.D.
Datong Electronics Ltd

Speech clipping at r.f. is now a well known and accepted method of increasing effective transmitter power. So far little experimental work has been done to determine the levels of intermodulation distortion generated by r.f. clipping. This study attempts to find the magnitude of distortion that can be expected when ideal clipping is applied. The analysis shows that quite high i.d. occurs with a two-equal-tone signal even for small amounts of clipping but, probably due to the waveform, this is not so objectionable when listening to speech.

In the use of a.m. and f.m. transmitters a.f. speech clipping has been widely used to raise the average-to-peak amplitude ratio of the speech waveform. Disadvantages of this technique are a high level of distortion, and the incompatibility with s.s.b. transmitters which are now the most common kind in use. If, however, the clipping is performed, at r.f. on a s.s.b. signal most of the distortion is eliminated and there is no incompatibility with s.s.b. transmitters. Radio frequency clipping by 15 to 20 dB can increase effective peak power by a factor of ten.

An add-on r.f. clipper can be placed in series with the transmitter microphone lead because the device generates its own s.s.b., clips it and demodulates it back to a.f. A single audio frequency sine wave, when processed in such a way, is reproduced as a sine wave with, assuming an ideal system, zero distortion, no matter what the degree of clipping. If the same signal is subjected to direct clipping it approaches a square wave and a set of odd harmonics is generated. If the wanted frequency is at the low end of the speech band, many of the harmonics are also in the speech bandwidth and cannot be filtered out. When several sine waves are simultaneously fed into a clipper the situation is more complex because a range of sum and difference products between various harmonics are generated.

The calculation which follows was undertaken in an attempt to find out how much distortion can be expected with ideal r.f. clipping applied to a typical two-tone s.s.b. test signal. Although a general calculation for the intermodulation produced by simultaneously clipping sine waves of varying amplitude would be complex, the special case of two equal-amplitude sine waves is more easily solved. The analysis shows that quite high intermodulation distortion occurs, even for small amounts of clipping. This is certainly not a condemnation of r.f. clipping for speech transmissions, but a point against in-band distortion, produced from a two-tone test signal, being used as an important parameter for characterizing a transmitter used for speech communications.

Calculation

A two-tone test signal can be regarded as a double sideband suppressed carrier signal produced by multiplying together two hypothetical signals. The test signal is defined as

$$f(t) = A_0\cos\omega_1 t + A_0\cos\omega_2 t, \qquad (1)$$

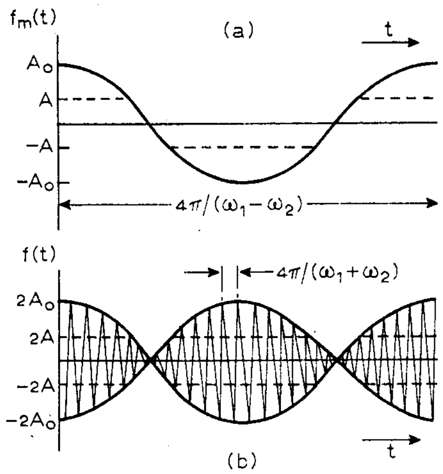


Fig. 1. (a) Modulating waveform, $f_m(t)$, which, when multiplied by cos $((\omega_1+\omega_2)/2)$ gives the two-tone test signal f(t). (b) Two-tone test signal f(t). Note its peak amplitude is twice that of the waveform in (a). The phase of the waveform with period $4\pi/(\omega_1+\omega_2)$ is ***shifted by π radians (180°) in each*** *successive section of the envelope. In both (a) and (b) the effect of amplitude limitation is shown by the dotted line.*

i.e. the sum of two signals. This can be written in the product form as

$$f(t) = 2A_0\cos(\omega_1+\omega_2/2)t\,.\cos(\Delta\omega/2)t \quad (2)$$

where $\Delta\omega = \omega_1 - \omega_2$. This second equation shows why the peak amplitude of the two-tone signal is twice that of each separate tone. It also shows that the envelope of the r.f. component, at angular frequency $\omega+\omega_2/2$ is amplitude modulated by a low frequency waveform which is equal to half the difference of the two original tones. The waveforms $f(t)$ and $2A_0\cos(\Delta\omega/2)t$ are shown in Fig. 1.

Because the envelope shape of $f(t)$ is directly related to that of the modulating waveform $2A\cos\Delta\omega t/2$, any amplitude limiting of the composite signal $f(t)$ can be regarded as due to similar limiting of the modulating waveform. This is indicated by the dotted lines in Fig. 1.

To predict the effects of r.f. clipping on a two-tone signal, one can first calculate the Fourier components representing a clipped cosine-wave, and then consider the effect of modulating $\cos(\omega_1+\omega_2/2)t$ separately by each of these components. The results of each modulation process, i.e. each multiplication, can then be added together to give the final frequency spectrum.

Any complex modulating signal $(f_m(t))$ with a fundamental frequency of $\Delta\omega/2$ can be represented as the following general Fourier series 1,

$$f_m(t) = a_0 + a_1\sin\Delta\omega t/2 + a_2\sin 2.\Delta\omega.t/2 + \ldots \text{etc.} + b_1\cos\Delta\omega t/2 + b_2\sin 2.\Delta\omega.t/2 + \ldots \text{etc.}$$

This expansion simplifies if there is no d.c. component in the waveform, and the clipping is symmetrical. Under these conditions only the odd cosine terms are left

$$f_m(t) = b_1\cos\Delta\omega.t/2 + b_3\cos3.\Delta\omega.t/2 + b_5\cos5.\Delta\omega.t/ + \ldots \text{etc.} \qquad (3)$$

This is why every effort should be made, in any a.f. or r.f. process to achieve

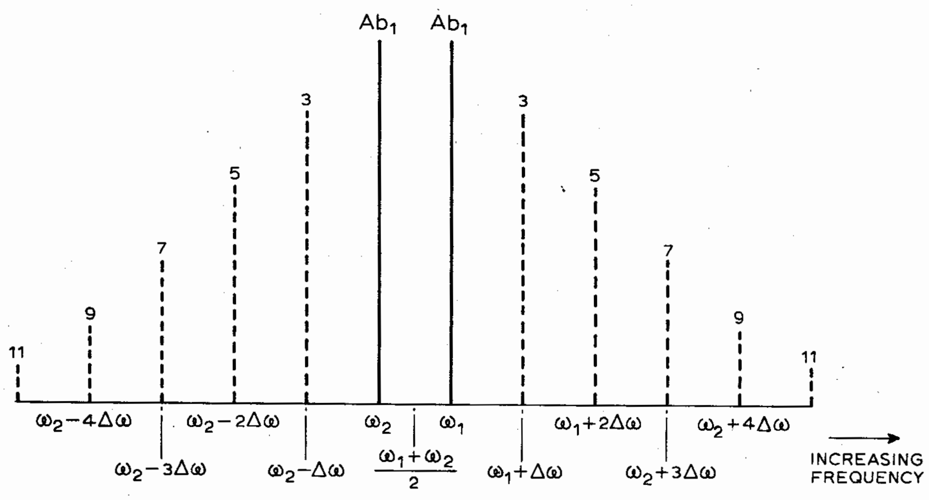
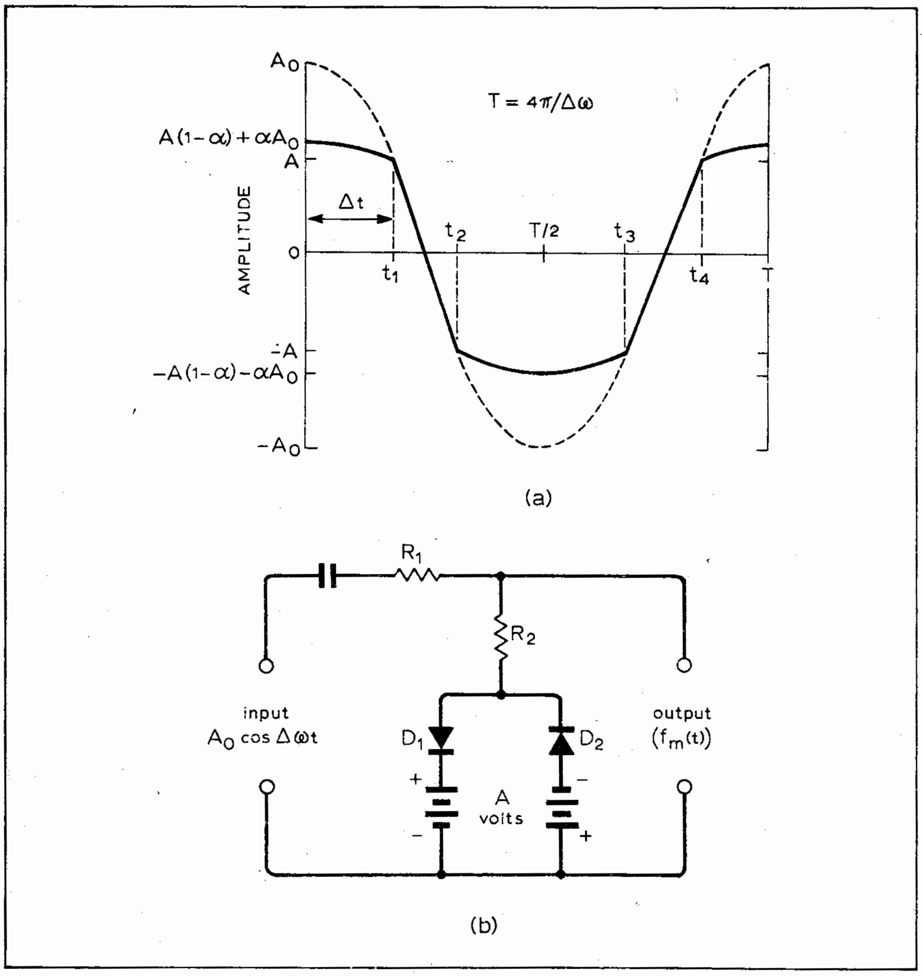


Fig. 2. Frequency components of the clipped waveform in Fig. 1(b). Note that there is no component with frequency $(\omega_1+\omega_2)/2$ despite the cyclic variation at this frequency in f(t).

Fig. 3(a). Function $f_m(t)$ used in the calculation of intermodulation levels (full line). Between t_1 and t_2, t_3 and t_4 it is an undistorted section of the waveform $A_0 \cos \Delta\omega t$. Elsewhere the latter has been attenuated to an extent dependent on the value of α. The full line is the waveform expected when feeding a cosine into the circuit of Fig. 3(b) which is an equivalent circuit of a practical diode clipper. Diodes D_1 and D_2 are assumed to have infinite reverse resistance, zero forward resistance, and zero offset voltage. The input signal comes from a source of zero internal resistance and the output is fed into an infinitely high load resistance. The parameter α is equal to $R_2/(R_1+R_2)$. **The input $A_0 \cos \Delta\omega t$, is** *an imaginary signal. In a real situation the clipper would have the same circuit except that the battery voltages would be 2A.*



symmetrical clipping. When each of these modulating components are used separately in equation (2), and the results are added to give the complete expansion for *f(t)*, the result is

$$f(t) = Ab_1 \cos \omega_1 t + Ab_1 \cos \omega_2 t$$
$$+ Ab_3 \cos(\omega_1+\Delta\omega)t + Ab_3 \cos(\omega_2-\Delta\omega)t$$
$$+ Ab_5 \cos(\omega_1+2\Delta\omega)t + Ab_5 \cos(\omega_2-2\Delta\omega)t$$
$$+ Ab_{1+2n} \cos(\omega_1+n\Delta\omega)t + Ab_{1+2n} \cos(\omega_2+n\Delta\omega)t.$$

The line spectrum corresponding to this expression is shown in Fig. 2. and the peak amplitude of each frequency component is shown above the corresponding spectral line.

Before discussing the actual magnitudes of b_1, b_3 etc., it is worth making a few points about Fig. 1 and 2. The higher the subscript, i.e. the higher the harmonic of $f_m(t)$ responsible for the particular component, the weaker the resulting component. Out of the unwanted intermodulation products, attention will be concentrated on the components at $\omega_1+\Delta\omega$ and $\omega_2-\Delta\omega$. These are the largest and also the closest in frequency to the desired signals and therefore the least likely to be filtered out. Fig. 1(b) is slightly misleading because the dotted lines, which represent the peak-to-peak amplitude after clipping, suggest that the waveform at frequency $(\omega_1+\omega_2)/2$ will be clipped. In fact, if this waveform is modulated by the clipped version of $f_m(t)$ shown in Fig. 1(a), i.e. a cosine wave of original amplitude A_o clipped to a peak amplitude of *A*, the sine shape of the high frequency waveform will be retained. This treatment therefore applies only to the case where clipping is achieved by an ideal r.f. compressor with threshold at *A* and with a long time constant compared to $4\pi/(\omega_1+\omega_2)$ but negligible compared to $4\pi/(\omega_1-\omega_2)$. Nothing is said about the spectral components which will appear centred on harmonics of $(\omega_1+\omega_2)/2$, but, because these are always discarded by subsequent filtering, this is not important. One assumes that for components centred on $(\omega_1+\omega_2)/2$, there will be negligible difference between true clipping and the hypothetical r.f. compressor mentioned above. The same assumption is also implicit in Kahn's treatment of hard clippers [2].

The magnitudes of coefficients b_1 and b_3 can be derived using Fourier analysis for the waveform shown by the full line in Fig. 3(a). An ideal limiter or clipper (R_2 in Fig. 3(b) equal to zero) allows no increase in input waveform amplitude after it has reached a certain threshold as shown by *A* in Fig. 3(a). Practical clippers approach this condition and therefore, the calculation has been carried out for the general case. The two results are

$$b_1 = A\left[\frac{2}{\pi}(1-\alpha)(1-x^2)^{1/2} +\right.$$

$$\frac{1}{x}(1-(1-\alpha)\frac{2}{\pi}\cos^{-1}x)\Big] \quad (4)$$

$$b3=\frac{4A}{3\pi}(1-\alpha)(1-x^2)^{3/2} \quad (5)$$

where $x=A/A_o$ and α is the gain of the clipper, see Fig. 3(a). It should be noted that in two-tone testing it is normal to measure the amplitude of the strongest intermodulation product relative to the product of the desired component. The intermodulation level in decibels will therefore be equal to

$$IP_{max}=20\log_{10}b_1/b_3 \quad (6)$$

where b_1 and b_3 are given by equations (4) and (5). In the case of infinite clipping with an ideal clipper, α and A/A_o are both zero. It can be shown that b_3 becomes $-4A/3\pi$ and b_1 becomes $4A/\pi$. For infinite clipping therefore, b_1/b_3 is 3. The minus sign is neglected because it merely implies a phase shift of π radians in the cos $3\Delta\omega t/2$ term relative to the cos $\Delta\omega t/2$ term. This result provides a check on the calculation because the waveform $f_m(t)$ is now a symmetrical square wave and the expansion for this is well known [1]. It also gives a useful result because the worst intermodulation product caused by infinite clipping of a two-tone signal will be at a level of −9.5dB relative to one of the wanted signals. In this special case it is easily shown that b_m is $4A/m\pi$, therefore the level of every intermodulation product can be derived, see Fig. 6.

A second check on the results can be obtained by making $A=A_0$ and $\Delta t=0$. It is found that $b_1=A_o$ and $b_3=0$ as expected because this is the case for no clipping. It was noted that with infinite clipping, b_1 is equal to $4A/\pi$. This amplitude is greater than the clipping threshold. If in addition α is non-zero, as in practical clippers, the wanted output will vary with the input amplitude even when the input is above the clipper threshold A. The degree of clipping in any situation is expressed as

$$B(\text{dB})=20\log_{10}(A_0/b_1) \quad (7)$$

and B represents the ratio of one wanted output to one of the inputs.

For a given degree of clipping it is useful to know how much the wanted output component has increased above the clipping level. This enables a.l.c. systems and other circuits following the clipper to be designed properly. The parameter is defined as

$$C(\text{dB})=20\log_{10}b_1/A. \quad (8)$$

In a practical system B and C are the important parameters from a performance point of view. From the design aspect A is important and it is worth noting that $B+C=-20\log_{10}x$ (9) where $x=A/A_0$ as before.

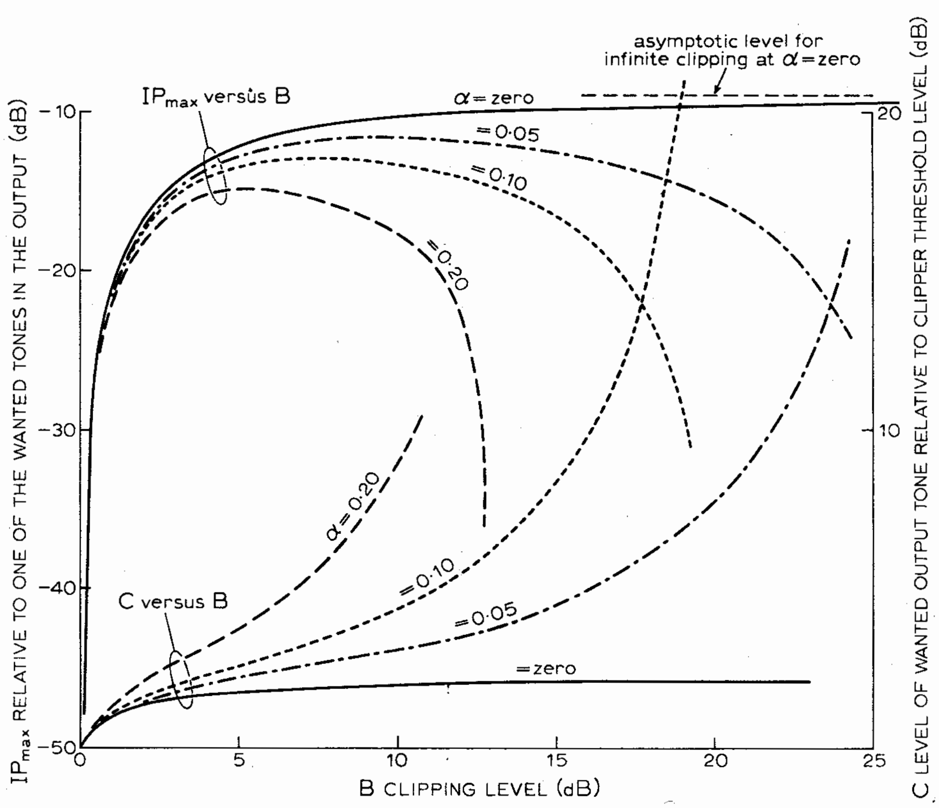


Fig. 4. The upper curves show the value of ***IP_{max} defined by equation (6), relative*** *to one of the wanted tones in the clipper output as a function of clipping level, defined by equation (7). The lower curves show the corresponding variations in the level of the desired output signal, defined by equation (8).*

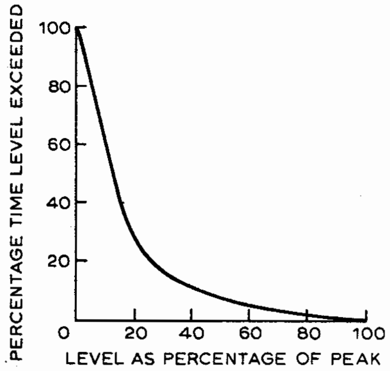


Fig. 5. Analysis of the instantaneous amplitude of phrases of speech spoken without pauses (taken from reference 5).

Numerical results

So that the expressions developed so far can be more easily used, the parameters IP_{max}, B and C as defined by equations (4) to (8) have been evaluated for a range of clipping levels and for various values of α using a calculator. The results are collected in Fig. 4 and 7. A most striking feature of these graphs is the high level of intermodulation distortion predicted for even small amounts of clipping. With an ideal clipper ($\alpha=0$) the amount of intermodulation produced, when one half of the r.f. signal is clipped, is obtained by putting $A/A_o=0.5$ in equation (9). This gives $B+C=+6$ dB. From the graph, B and C will then be 4.3 and 1.7 respectively and IP_{max} will be close to −13 dB. Note that the clipping level as defined in equation (7) will be 4.3 dB and not 6 dB. This situation is hardly improved by using a non-zero value of α. Although the IP_{max} curves for non-zero α bend over at high clipping levels to give an apparent improvement in i.p. levels, the curve for the level of wanted output rises steeply at the same time. This apparent improvement is therefore at the expense of large changes in output level. The explanation is, at high clipping levels with non-zero α, the distance between the levels A and $A(1-\alpha)+\alpha A_o$, in Fig. 3, has become so large relative to A, the latter has become negligible. The obvious steepness of the intermodulation curve at low clipping levels is a reflection of the flatness of the cosine function around the peaks. In this region a small increase in clipping level produces a large increase in the function affected by the clipper. For example, with only 1dB of clipping, Fig. 4 shows that $B=C=1$ and hence $A/A_o=0.794$. But cos $2\pi\Delta t/T=A/A_o$, and therefore $\Delta t/T=0.104$. This means that the flat sections of the clipped waveform in Fig. 3 account for as much as 41.6% of the total period T. The corresponding value for IP_{max} is −20.8 dB. As clipping increases, the discontinuities in the clipped waveform travel down ever steepening parts of the cosine function and the additional effect on i.p. level diminishes. At a clipping level of 10 dB, 84% of the waveform in Fig. 3 is flat and the intermodulation level is −10.4 dB, which is only 0.9 dB better than the ultimate value when the whole waveform is flat. Corresponding values for 3, 5 and 20 dB of clipping are 60% and −14.5 dB, 69% and −12.2 dB, 95% and −9.6 dB. It is striking that from only 3 dB of clipping to infinite clipping, the third order intermodulation products change by only 5 dB.

The percentage of a two-tone signal which has been flattened by clipping is easily measured using an oscilloscope. This parameter, called y, has been plotted against third order intermodulation level in Fig. 7 so that an upper limit for the intermodulation caused by flat-topping in a transmitter can be more easily estimated. Also shown is one quadrant of the cosine curve $x=\cos(\pi/2.y)$ which illustrates the comments made about the flatness around $y=0$. It also allows the ratio of clipper threshold to input amplitude, x, to be quickly related to y, and hence IP_{max}. It seems clear from these curves that a small amount of flat-topping in a transmitter will cause a lot of interference in adjacent channels if the transmitter is being driven by a two-equal-tone test signal. A similar point has also been made by L. A. Moxon (G6XN) in his recent review of r.f. speech clipping [8].

Unfortunately there seems to be a shortage of experimental data with which to compare the results of the calculation. Data points taken from the only sources located have been superimposed on the graph in Fig. 7. W. Sabin's (WO1YH) measurements agree well with the theoretical curve but are restricted to 10 and 20 dB of clipping. P. E. Chadwick (G3RZP) measured intermodulation products as a function of the parameter y [10] but he does not state whether his i.ps. are third order or the sum of all intermodulation products. The four values quoted seem fairly consistent but, compared with the theoretical curve they appear to underestimate the distortion as the clipping increases. The four points supplied by P. J. Horwood (G3FRB) do not agree with the theoretical curve or Sabin's data. It would have been useful if the two test-tone frequencies had been quoted, (those used by Sabin were 600Hz and 1kHz) because the low intermodulation level quoted for 10 dB of clipping raises the suspicion that the intermodulation products were being reduced by the post-clipper filter.

At 20 dB of clipping the present calculation and the data quoted indicate the i.p. level is substantial and would be objectionable if one were listening to a two-tone test signal. When listening to speech the i.p. is not so objectionable; this is probably due to the nature of the waveform. The graph in Fig. 5 shows

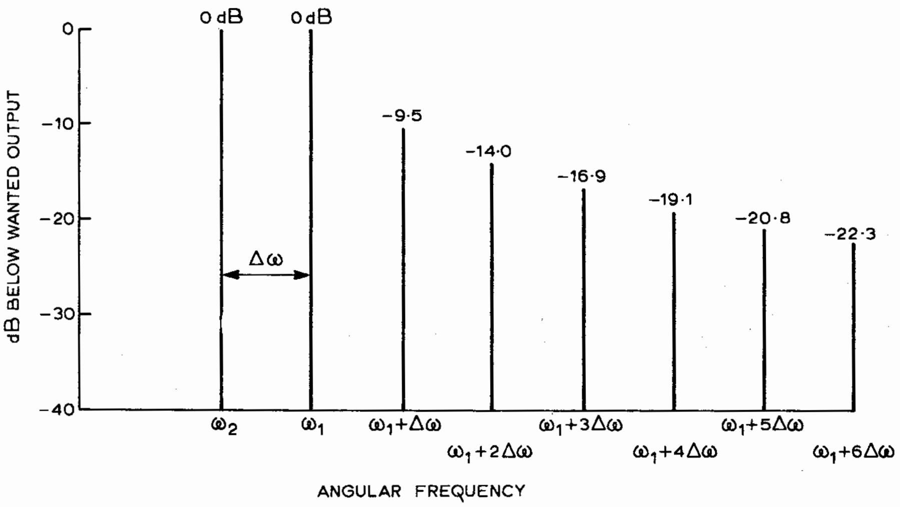
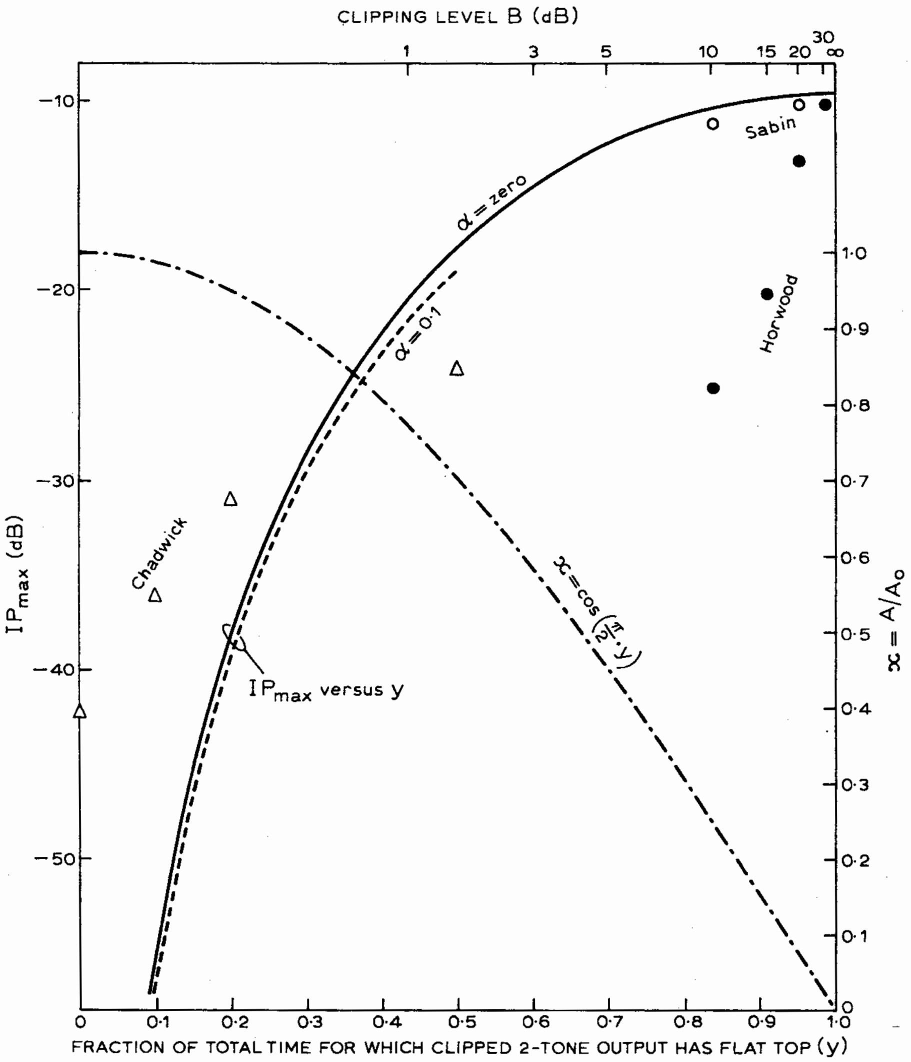


Fig. 6. Frequency spectrum of an infinitely clipped two-tone s.s.b. signal. The high levels of unwanted components (all except those with frequency ω_2 or ω_1) show the need for a narrow band filter after a r.f. clipper. The spectrum is symmetrical about $\omega_2+\omega_1/2$ but only frequencies greater than this are shown.

Fig. 7. Data in Fig. 4. is shown replotted to illustrate how the level of third order intermodulation products varies with y. The relationship between y and x is also plotted which is merely one quadrant of a cosine curve. Also plotted are some published experimental data for third order intermodulation levels.



the percentage of total observational time during which the instantaneous amplitude of speech waveforms exceeds a given percentage of the peak level (reference 4). The 33% level is exceeded for only 14% of the time. If, therefore, the speech clipping begins at the 33% level (i.e. $A/A_o = 0.33$), up to about 8 dBs of clipping will occur, but only for a fraction of the total time. It seems quite reasonable therefore that neither r.f. nor a.f. clipping to this extent will cause much subjective deterioration of the speech, and this is what is observed in practice.

Because the remaining 86% of the speech has been raised in level by 9.5 dBs a large increase in talk power is achieved. Although this applies equally well to a.f. or r.f. clipping, the audible difference is less pronounced for the latter at high clipping levels because fewer distortion products are produced. For example, many pairs of frequencies in a speech waveform will have intermodulation products that fall outside the subsequent filter's pass-band. These two frequencies will appear undistorted whereas distortion products would still arise if a.f. clipping were used. Perhaps a more important point is that the level of intermodulation products produced when two tones are clipped at r.f. will reduce when their unclipped levels become unequal. In the limiting case of a single tone, no distortion is produced by an r.f. clipper. This is in direct contrast to audio clipping, where even a single tone would produce a series of odd harmonics which waste transmitter power.

Farkas and Gschwindt [5] have shown theoretically and empirically that for 20 dB of clipping, the output power from an audio clipper is 1.66 times that from a r.f. clipper. This extra 66% of power is wasted so far as wanted information is concerned. Furthermore, the extra power is being used to transmit interference within the information bandwidth and therefore, compared with r.f., a.f. clipping contains a built-in jammer which has 66% of the strength of a r.f. clipped signal.

Conclusion

In testing linear amplifiers for non-linearity at any frequency, a two-tone test input is widely used because of the relatively high levels of intermodulation products (compared to harmonics from a single tone of the same amplitude in the same situation) produced by a given degree of non-linearity [6]. Considering the extreme deviations from linear amplification produced by any form of clipper or amplitude limiter, the levels of intermodulation products calculated in Fig. 4 seem plausible. One feels, however, that maybe some other factor made the distortions figures, quoted in reference 3, so low.

For reasons noted already the two-equal-tone test signal is not suited for testing in-band distortion in a speech transmitter using a clipper. The result is too pessimistic when compared with the audible deterioration. On the other hand, when testing for out-of-band interference in a s.s.b. system using r.f. clipping, a two-tone signal is valuable because it represents a worst-case situation, which is what adjacent channel users are interested in. The present calculation could easily be carried further to evaluate the amplitudes of higher intermodulation products more likely to be outside the information bandwidths defined by the post-clipping filter. It is probably safer to assume the worst-case situation of infinite clipping where the amplitudes of intermodulation products produced by the two-equal-tone input are given by $b_m = |(4A/m\pi)|$. These components are plotted in terms of decibels below the wanted output in Fig. 6. A similar diagram appears in reference 7. The high level in adjacent channel regions shows why the filter following an r.f. clipper has to do more than remove harmonics at $2f$, $3f$, etc., where f is the carrier frequency, as has been implied in at least one publication. With a speech input Fig. 6 would appear more as a continuous spectrum. This figure emphasises the recommendation that the post-clipper filter should be of comparable specification to the original sideband selection filter. Such a filter also helps to reduce the carrier and unwanted sidebands to their original pre-clipping level.

Crossover distortion

The solid line in Fig. 3(a) can also be regarded as an exaggerated output from a class B audio amplifier suffering from cross-over distortion and where the input signal is a single sinusoid. The same calculation can therefore be used for determining the amplitudes of the harmonics produced by such an amplifier. Similarly the curves given in Fig. 4 for low values of α can be interpreted as giving the amplitude of the strongest inter-modulation product produced by a class B r.f. power amplifier with an arbitrary amount of cross-over distortion when fed with a two-equal-tone test signal. This treatment cannot be used to calculate the intermodulation products for a two-tone audio-frequency signal subjected to clipping or cross-over distortion for reasons given earlier. Qualitatively, however, the result of clipping such a signal will be a series of overlapping sets of products as in Fig. 2, with each set centred on a multiple of $(\omega_1 + \omega_2)/2$. Thus, the total number of unwanted products within the information bandwidth will be much larger than if the compression of dynamic range is carried out by heterodyning to r.f., clipping, filtering harmonics, and heterodyning the result back to a.f.

References

1. *Higher Mathematics*, J. W. Mellor, Dover Publications Inc., 1954.
2. L. R. Kahn, "Analysis of a limiter as a variable-gain device", *Electrical Engineering*, 72, pp 1106-1109, 1953.
3. P. J. Horwood, "A clipper and filter for transistorised SSB exciters", *Radio Communication*, 48, pp 88-89, February 1972.
4. P. F. Raven, "The use of double sideband suppressed carrier modulation for portable VHF radio telephones", *I.E.E. Conference Publication* No. 64, *Signal Processing Methods for Radiotelephony*, 1970, pp43-46.
5. B. Farkas, R. A. Gschwindt, "An investigation on dynamic compression for the transmission of speech", *I.E.E. Conference Publication* No. 64, *Signal Processing for Radiotelephony*, 1970, pp97-101.
6. H. M. Nodelman and F. W. Smith, *Mathematics for Electronics*, McGraw-Hill Book Co. Inc., 1956, chapter 10.
7. J. Daguet and K. Gilabert, "La parole a niveau constant dans les emetteurs a bands laterale unique", *L'Onde Electrique*, XLI, No. 410, pp498-509, 1961.
8. L. A. Moxon, "Performance of r.f. speech clippers", *Ham Radio*, pp26-33, November 1972.
9. W. Sabin, "R.f. clippers for S.S.B.", *QST*, pp13-18, July 1967.
10. P. E. Chadwick, "Technical Topics", *Radio Communication*, 48, pp457, July 1972.

Sixty Years Ago

Our occasional "Letter from America" was a feature of the journals as long ago as sixty years. The Letter in the October 1916 issue — incidentally written by David Sarnoff who later became chairman of RCA — includes among other things the small drama of a stolen crystal detector:

"The operator in charge of a wireless station installed on a vessel left his cabin for a few minutes, during which time an unscrupulous person entered the radio cabin and purloined the crystal detector from the tuner, as well as the spare crystal detectors which were lying in a box on the operating table. The ingenuity of this particular operator was not all that could be desired, and as a result the vessel was unable to receive wireless signals for a period of two and a half days."

The report then goes on to quote *The Electrical World:*

"A condition in which the operation of a wireless telegraph outfit must be entirely suspended on account of the lack of a spare bit of carborundum is not healthy. That it can exist reflects no credit upon either the operator himself, the company which trained and employed him, or the naval examiners who granted his licence certificate. This is quite over and above that part of the responsibility of the operator to construct a temporary detector of at least enough sensitiveness and reliability to keep the ship radio in operation.

" . . . Had the operator, who was forced to sit idly by his dead receiver because the crystal had been stolen, known a little of what expedients were made use of in the early years of wireless telegraphy there would have been no failure to protect his ship by radio service. Two needles and a pencil, a knife-blade and a broken incandescent lamp, a piece of dry-cell carbon and an iron wire — any of these could be used as a microphonic detector which would take the place of the crystal and receive signals from fifty to one hundred miles."