Which Tube Shall I Use ?

GEORGE FLETCHER COOPER

The author presents a lucid yet simplified description of the use of tube characteristic curves for the selection of the proper tube, from the standpoint of distortion, for a given application.

I am always reading, and for that matter writing, articles on how to design this and that, especially in the way of audio amplifiers. My staff of trained statisticians estimates that if all the audio amplifiers which have been designed were set to operate at full output simultaneously, one in four of the population would be off to have their eardrums pierced. One thing, however, is scarcely mentioned: although I explain why I take feedback round from end to beginning, and my friend X tells you why he uses only local feedback, we rarely explain why we chose to use a particular tube in a particular place. Oliver Heaviside said somewhere, "Even Cambridge mathematicians deserve justice," so I must rapidly plead that we do usually tell you why we chose our output stage tubes. My own personal choice for the 20-100 watt output stage is the EL34, even if it is a little harder to get than old faithfuls like the 6L6, just because I have found it very easy to drive. Maybe some of the newer tubes are even easier, but at the moment I have no reason to try to find out.

The tube at the front end of the amplifier may choose itself, too, especially in portable equipment. My own personal feeling is that microphony is best treated by using the tube you like in its ruggedized version and providing some mechanical anti-shock system, while hum is rarely troublesome if the heater line is lifted up to about 20 volts above ground and is kept balanced. I am tempted to rectify the heater supply rather than use special tubes, but so far I have never been that desperate. As a result I am pretty free to choose the tubes right through the amplifier. There are some special considerations to be taken into account which will not affect most readers: when you are working on a whole range of equipment, spares inventories must be remembered or you find you need more space for spare tubes than for the operating installation.

Designers of studio-frequency amplifiers usually adopt a figure of merit based on the ratio of transconductance to capacitance. Just how you combine input and output capacitance depends on the interstage coupling circuit. This is discussed in Terman and I do not propose to look it up and copy it down. A more sophisticated designer will go on to worry about grid-plate capacitance, the usual limiting factor in designing narrow-band r.f. amplifiers. And then—but we are, for now anyway, audio designers. What are we to do?

The basic problem of the audio designer is to get gain without distortion. As you all know, you can trade gain for distortion by using negative feedback and you can also trade distortion for gain by using positive feedback. Not surprisingly, the bigger the deal the bigger the headache. Ideally each stage of an amplifier could then be used to maximum effect, but this really does involve quite a design job.

Tube Data

What we need in making our choice of tube type is some fairly simple criterion. In the old days the tube makers gave us a short set of tube data, curves and a couple of circuits all on a single small sheet of paper: now no self-respecting manufacturer would send out data sheets weighing less than the tube itself. What we need to do, therefore, is to find some way of summarizing this information into a form suitable for easy comparison between tubes.

Most of the vital statistics of a tube seem to be included, for our purposes, in a single curve. This is the curve of transconductance against bias. From this curve we can derive a whole mass of other information, and we can also just plot the curves for several different tubes on the same sheet of paper without getting into too much of a muddle. For some reason the textbook writers have never taken to this rather simple approach so that it is not nearly as widely known as it deserves.

First of all, let us assume that we are working with pentodes. Then the characteristics of the tube are, for all practical purposes, independent of plate load. We write down the plate current I_p as a function of grid voltage:

$$\mathrm{I_p} = \mathrm{I_o} + a\mathrm{e}_g + b\mathrm{e}_g{}^2 + g\mathrm{e}_g{}^3 + \ldots (1)$$

In this expression e_g is measured from the normal working point and represents the input signal. This corresponds to the practical arrangement in which the cathode is biased positive by the drop in the cathode resistor, the bias is held constant by a large decoupling capacitor and the grid is returned through a high resistance to ground. I_p is just the plate current with no signal applied. To find the transconductance we differentiate, giving

$$a\mathrm{I_p}/a\mathrm{e_g} = \mathrm{g_m} = a + 2b\mathrm{eg} + 3\mathrm{ye_g}{}^2 + \ldots (2)$$

Of course we normally consider the transconductance to be the limiting value for very small signals, so that actually a is the transconductance given in tube data.

In *Eq. (2)* we have written down

$$\frac{dI_p}{de_g} = g_m$$

This is just the same as writing down

$$\mathrm{Ip} = \int \mathrm{g}_m a_g$$

provided we put in the proper limits. We know that if $e_p = e_{cg}$, the cut-off voltage, we must have $I_p = o$ and if $e_g = o$ the plate current is I_o. Then

$$\mathrm{I_o} = \int_{e_{gc}}^{0} g_m a e_g$$

As I hope you remember, this is just the area under the $g_m - e_g$ curve, the area shown shaded in *Fig. 1.* Therefore we can tell the price we must pay in plate current for any particular trans-conductance, and thus for any particular gain. The area under the curve, if it is of the form shown in Fig. 1, is most easily found by calculating the area of the triangle C'AP and then either counting squares or making a rough estimate of the area of the small wedge on the left. Remember, when doing this, that the average tube tolerances are quite large and do not, I beg you, try to work to 1 per cent, or even 5 per cent. It just doesn't mean a thing!

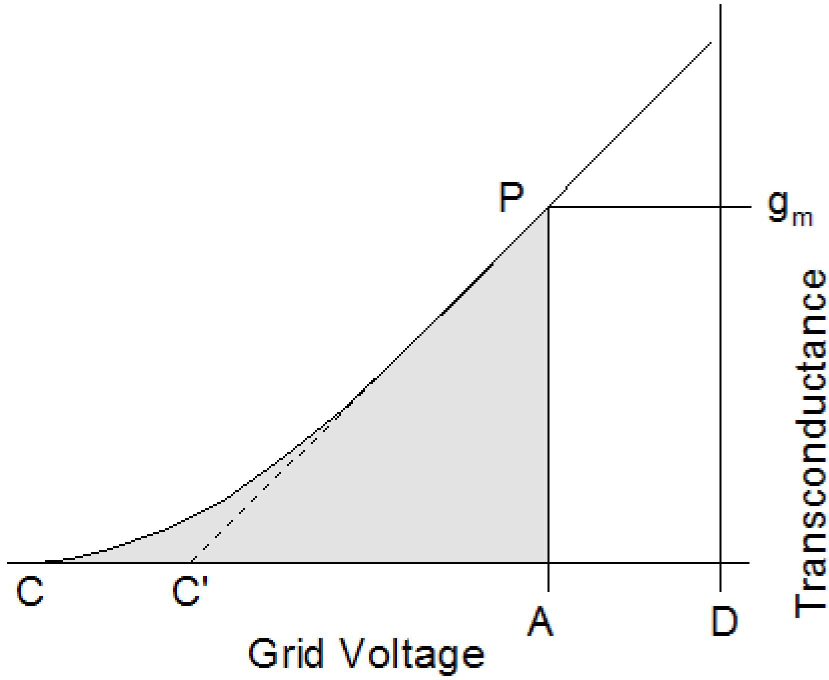


Fig. 1. A typical curve of transconductance against grid voltage. At P, the working point, the grid voltage is token to be zero, AO being the voltage of the cathode above ground. C is the cutoff point. The area shaded is the standing plate current.

What else can we find out from this graph? Well, let us look again at Eq. (1), and assume that $e_g = e \cos wt$. It is, by the way, always a good thing to use *cos wt* rather than *sin wt* in harmonic calculations, because then there are no minus signs to make the expressions more awkward. Since $e_g = e \cos wt$, we have

$$e_g{}^2 = e^2 \cos^2 wt = \frac{1}{2} e^2 (1 + \cos 2wt)$$

and

$$e_g{}^3 = e^3 \cos^3 wt = \frac{1}{4} e^3 \text{-} (3 \cos wt + \cos 3wt)$$

Equation (1) therefore becomes

$$I_p = I_o + \frac{1}{2} b e^2 + \left(a + \frac{3}{4} g e^2 \right) e \cos wt + \frac{1}{2} b e^2 \cos 2wt + \frac{1}{4} g e^2 \cos 3wt + \ldots \quad (1a)$$

This equation is correct as long as we are justified in neglecting anything above the third power of e_g. Already, as you see, the steady component is affected by the γ term and the fundamental is affected by the *ß* term. These interactions are actually intermodulation effects in which the signal mixes with its own harmonies to produce other harmonics. The more terms we take, the more likely the editor is to say he doesn't like mathematics!

If we simplify *Eq. (1a)* rather more by assuming that γ is zero, which means taking only second harmonic into account, and then turn to *Eq. (2)*, we have

$$g_m = a + 2 b e_g \qquad (2a)$$

Now let us differentiate this, giving

$$\frac{dg_m}{de_g} = 2b$$

This means that *ß* is a measure of the slope of the transconductance characteristic. The ratio of second harmonic to fundamental in the plate current is

$$\frac{1}{2} be^2 / a = eb / 2a$$

Let us look at *Fig. 2*, which is really only a part of *Fig.1* redrawn. The signal drives the grid alternately to the right and the left of *A* with a total range of *2e* (i.e. the peak-to-peak voltage). At *P* the transconductance is *a*, but the range over which the transconductance varies is 2_g, The slope of the transconductance curve is thus *2g/2e= g/e*. But this, we have just seen, is *2ß*. Hence we have *g = 2 ße* and the ratio of second harmonic to fundamental becomes *g/4a*.

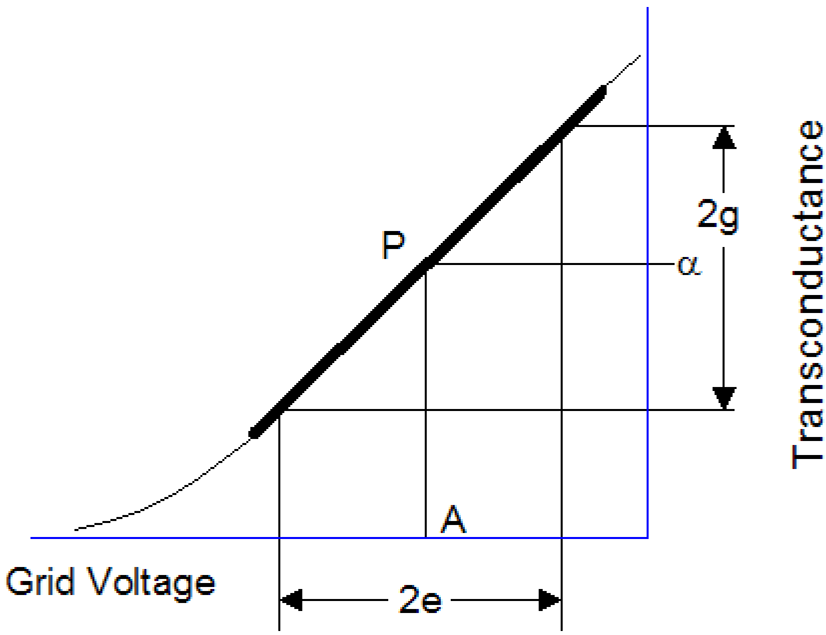


Fig. 2. A signal drives the grid to a peak distance of e volts on either side of A, the reference point. The transconductance varies from α - g to α + g.

It is, of course, very easy to find this just by looking at the tube characteristic.

We want to have the largest gain, that is the largest value of a, for the smallest value of distortion, which in this case means the smallest value of *g/4a*. If we take gain/distortion as a figure of merit we have $\alpha^2/4g$, or since we are always comparing tubes and need not keep dividing everything by 4, we have α^2/g as a gain/distortion figure of merit.

Practical Example

At this point we should, I suppose, look at some typical tubes. Skimming through a tube handbook I have picked out three tubes and have replotted their characteristics on the same scale, together, in *Fig. 3*. Choosing a working point at a bias of -1.6 volts and assuming a swing of ± 0.4 volts, we have for these three:

	Tube	a(transconductance)	2g (see Fig. 2)	a^2/g
1	6AU6	3.5	1.4	17.5
2	6BA6	4.0	1.0	32

3	6BS7	2.0	0.55	14.5

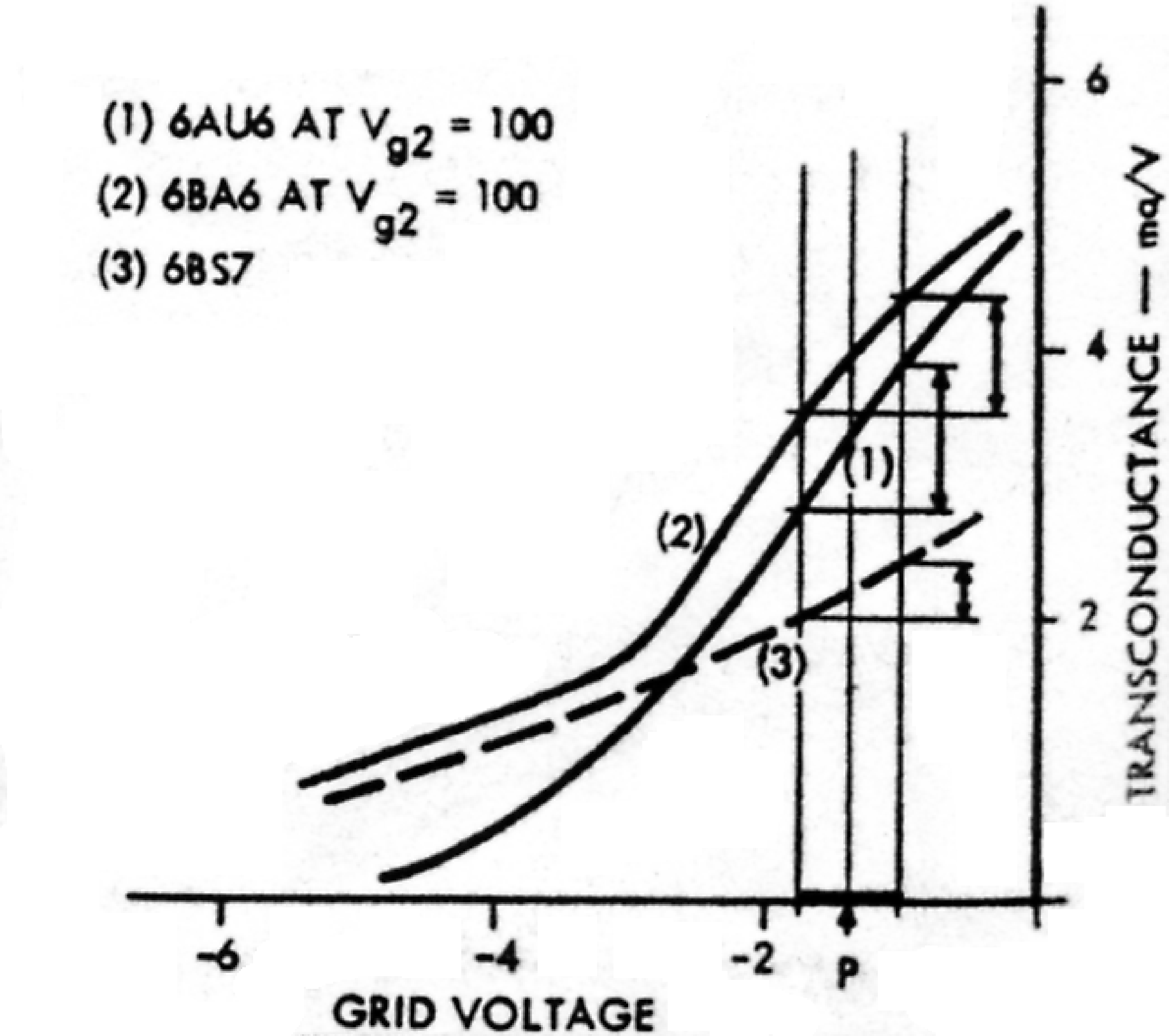


Fig. 3. Comparison of three tube types operating at - 1.6 v. bias with a grid swing of ± 0.4 volts.

It would appear, then, that the 6BA6 has the highest figure of merit, and that the low-g_m tube, the 6BS7, although it produces less distortion than the 6AU6, does not pay off because it is better to accept the distortion for the sake of the extra gain.

One thing will be noted, however. The 6BA6 curve is everywhere above the 6AU6 curve and in fact the current taken by the 6BA6 is about twice that taken by the 6AU6. We really should carry out another check of this sort, based on equal plate currents. Then we look at our particular problem and decide whether economy in current is important. Very often it is, even if only because it becomes easier to decouple the early stages from the power-supply disturbance originating in the output stage. In portable equipment every milliampere adds ounces to the weight of the smoothing components.

It is worth while also working out what the distortion actually is : for our three tubes under the conditions given we have:

g/4α x 100 = 2nd harmonic distortion in per cent

1. 6AU6	**(0.7/14) x 100 =**	**5.0**
2. 6BA6	**(0.5/16) x 100 =**	**3.12**
3. 6BS7	**(0.275/8) x 100 =**	**3.44**

This shows us that we are favoring the 6BA6 for both distortion and gain. It also gives us some numbers to compare with the distortion in other parts of the amplifier, so that we can decide if this is our critical point or if we should be spending more time worrying about another stage.

Triode Considerations

We have now seen how to assess a pentode stage for its second harmonic distortion. Suppose, however, we are triode users. The answer then is, I'm afraid, it all depends on the tube maker. I have found one who provides me with curves of transconductance vs. grid voltage for three different plateloads, for the 12AT7, anyway. Sometimes you only get curves of μ and r_p. When that happens you have to do rather more work.

The gain of a triode stage is

$$\mu R_L/(r_p+R_L)$$

so that the effective transconductance is $\mu / (r_p + R_L)$. It is tedious but not exhausting to tabulate µ and r_p for different values of e_g, then add R_L to each r_p and then work out this effective transconductance. To bring triodes into our general net we must do this. I think it is worthwhile, just because one unified approach does save quite a lot of thought and effort in the long run.

While we are talking triodes let us deal with the effect of a local feedback loop. This is long-hair language for leaving the cathode resistance without decoupling. (By the way, just what is the difference between the long-hair approach and the egg-head?). We know that the gain of the stage becomes

$$\mu RL/[RL + rp(\mu + 1)RK]$$

And we have already decided that $\mu/ (R_L + r_p)$ was the effective transconductance, which we will call g_m. So the gain becomes

$$R_L / \left(\frac{1}{g_m} + \frac{\mu+1}{\mu} R_k \right)$$

Now µ is always big enough for us to take $(\mu + 1) \mu =1$. After all we shall probably use 20 per cent tolerance components for R_K. We thus have a new transconductance, the effective feedback transconductance g_{mf}, which is given by

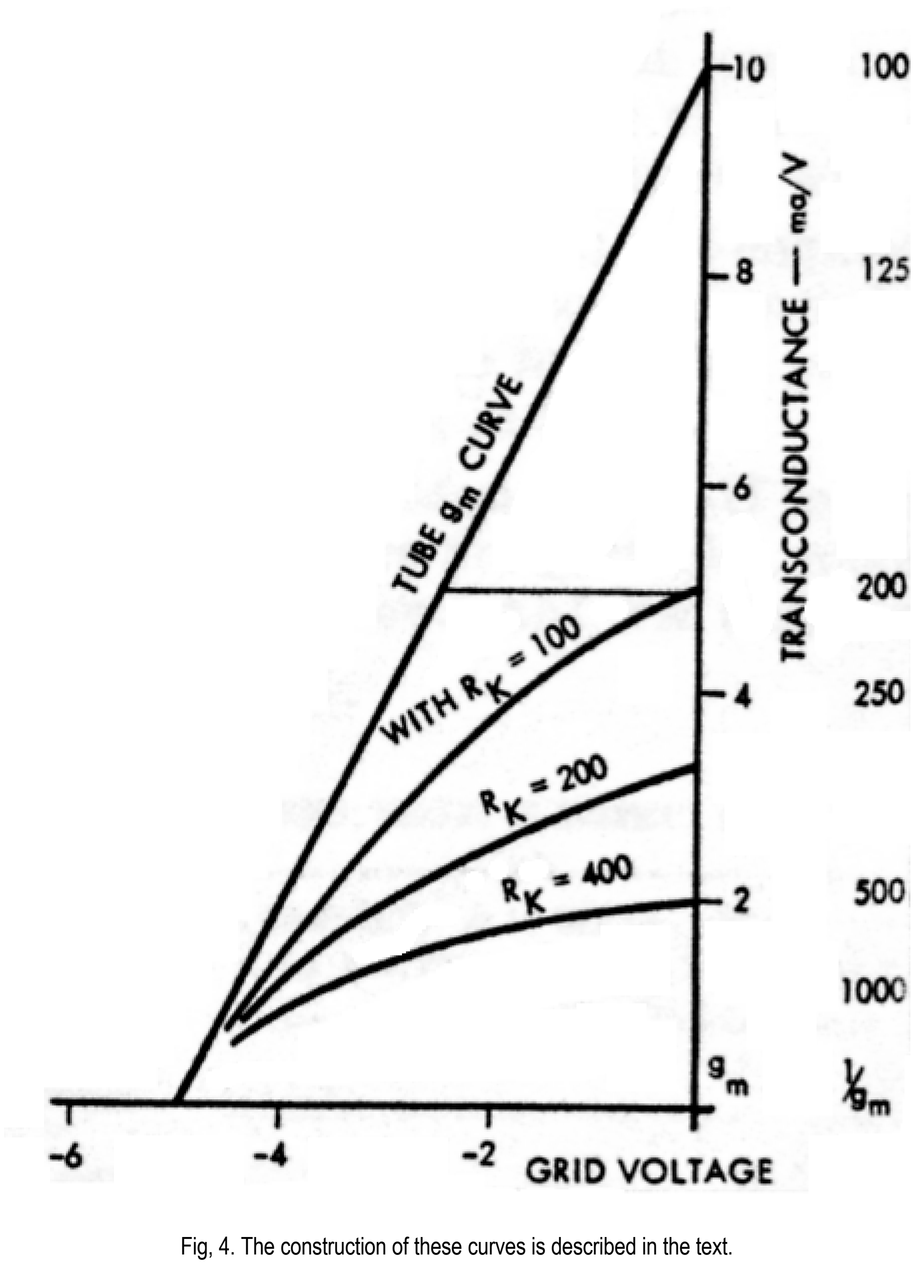
$$\frac{1}{g_{mf}} = \frac{1}{g_m} + R_k$$

Let us look at *Fig. 4.* To make the work easier I have drawn a straight line g_m characteristic and I have marked the g_m, axis in the values of $1/g_m$. Thus at $e_g = 0$ the g_m is 10 mA/v, and $1/g_m = 100$ *ohms*. If we consider the effect of a cathode resistance of 100 ohms the value of $1/g_mf$ is clearly, at this point, $100 + 100 = 200$ ohms, so that $g_{mf} = 5.0$ *mA/v* and we can plot this point. By working out a reasonable number of values in this way you can quite easily sketch in the curves shown in *Fig. 4.* You can do this for a pentode on the ordinary transconductance curve, and for a triode on the effective transconductance curve described earlier in this article.

There is a rather interesting thing which appears if you draw out this set of curves very carefully, very large. When you work out the figure of merit for a swing from $e_g = o$ to $e_g =-2.5$ for the tube alone and for the tube with a 100-ohm cathode resistor, there is a small advantage in favor of the tube alone. It is less than 10 per cent better, which is not very much, of course. But the feedback does not quite reduce the distortion in the same proportion as it reduces the gain.

Just why this is so is a rather complicated question which I hope to discuss in another article.

Another point about the $R_K = 100$ and $R_K = 200$ curves of *Fig. 4* is that they are not, like the tube transconductance characteristic, straight lines.. Of course the practical tube transconductance itself would have some curvature, but adding the feedback seems to introduce some extra curvature. The result is, in fact, to add a third harmonic term. Physically this is produced by a mechanism of the following kind: we apply a cosine wave to the grid and produce some second harmonic in the cathode current. Because the cathode resistor is not decoupled the cathode voltage contains a second harmonic term. Between grid and cathode, then, we have a signal containing both fundamental and second harmonic. The tube now acts as a mixer, to produce terms of the $(2f + f)$ and $(2f - f)$ kind, of which the first is, obviously, the third harmonic.



Fig, 4. The construction of these curves is described in the text.

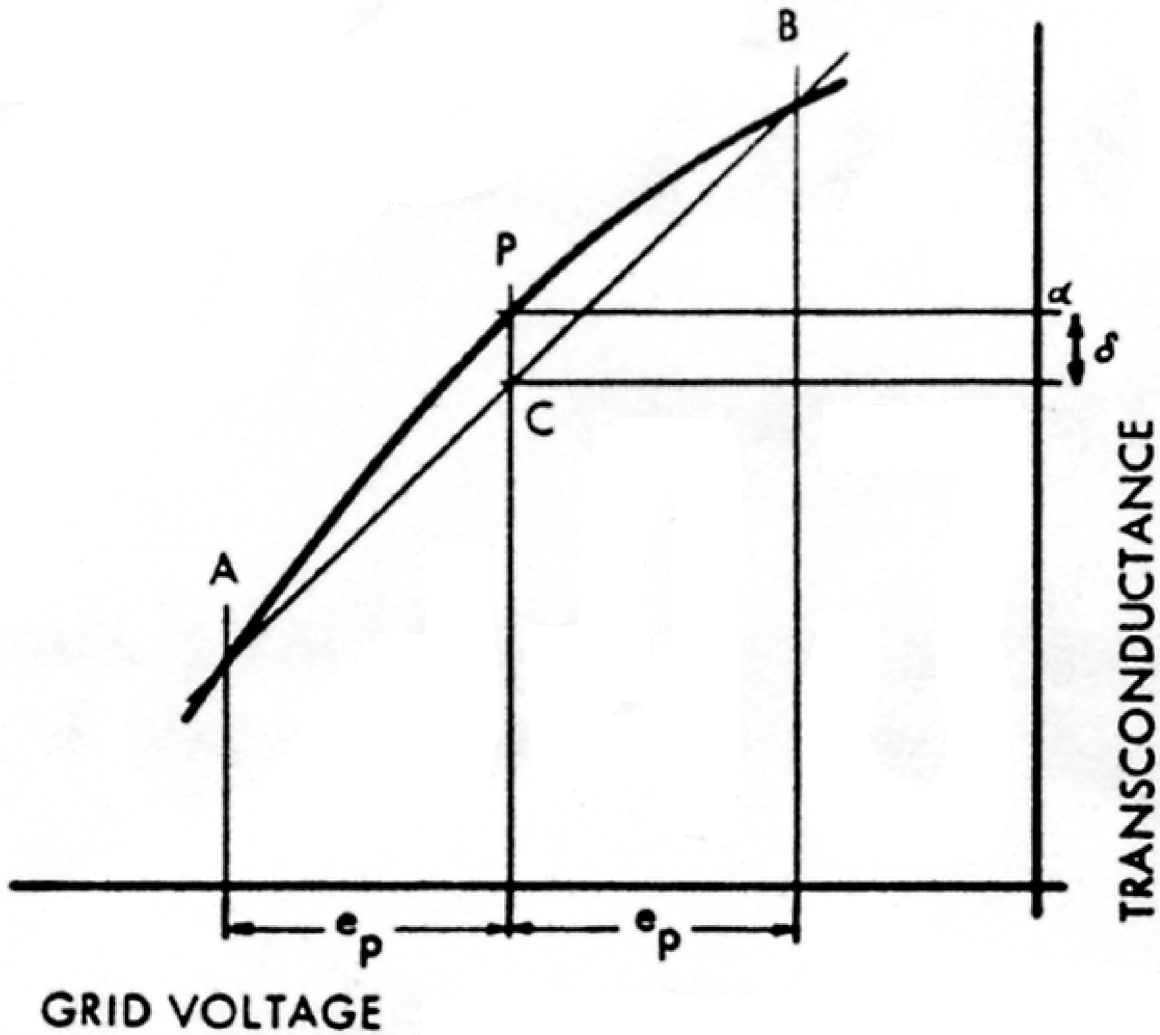


Fig. 5. From the deviation of the transconductance curve from a straight line —the chord AB—we can find the third harmonic distortion.

The transconductance curve contains information about this third harmonic. I am not going through the mathematics in detail, because it is rather lengthy and can be regarded as an exercise for the enthusiast. The practical result requires us to refer to the idealized characteristic of *Fig. 5*. Having chosen our working point *P* and the maximum signal amplitude $\pm e_p$ we draw the chord *AB* across the transconductance curve. At *P* the transconductance is *a*. We measure the distance *PC* which we call *d*. For Any input of less than e_p, say *e* volts, the third harmonic distortion is

$$\frac{3d}{a}\left(\frac{e}{e_p}\right)^2$$

I am not going to apply this to the curves of *Fig: 4,* because it takes pretty accurate drawing. Looking back to our second harmonic expression $g/4\alpha$ we see that if $d = g/12$ the second and third harmonics will be equal. You need to look pretty closely to check on the third harmonic.

I think we can now get a rough idea of when local feedback will really be profitable. Most tubes tend to have transconductance characteristics which sag below the straight line I have been drawing so glibly. A little cathode feedback produces a curvature in the opposite direction. By careful choice of the cathode resistor you can get a pretty good linear characteristic and then, going push-pull, balance out the second harmonic. I remember that this worked out very well indeed with the 5763, though I have no figures at hand now to show the improvement.

Some readers may feel that all this is just a paper exercise. Maybe so, if you can afford to buy the wrong tubes. But if you want to get the best performance out of something you design yourself it is worth-while to sit down for a few hours and think before you put your money on the counter.