Charging

A further look at the CR coupling

by Cathode Ray

In reviewing basic theory since 1911 for the 60th birthday issue of *Wireless World* I mentioned that during the second World War I was shocked to find radar instructors teaching that when (say) a positive-going input signal was applied to a CR coupling the output also went positive *because of the charging of C.* In actual fact (as I went on to say) any charging or discharging of C appears only as *distortion* of the signal at the output. I included also the words 'of course', by way of apology to readers for wasting their time by explaining where the quoted teaching was wrong. Wasn't it too obvious in these enlightened days?

Apparently not, for I soon got a letter to say that it was I, not the instructors, who was wrong. Touched though I was by this loyalty to a fine body of men, I felt that this evidence that my own experience of them was not unique called for some more detailed exposition of the point in question, in case the fallacy lingered on in a bigger way than I had suspected. I admit that some trainees might have misunderstood what their instructors taught about this. I will go farther and declare that many trainees did misunderstand what their instructors taught about this and about many other things. So not all that they taught in 1941 should be judged by what their trainees thought they said. And even if some of them *were* wrong on this point of circuit theory, we won the war so what the hell?

No one is likely to argue that uncertainty on the part of some radar mechs about the precise mode of functioning of interstage couplings in pulse amplifiers was responsible for a major loss of effectiveness in Britain's wartime radar defences, but I will and do hold that anybody who wants to be clever with electronic circuits ought not to have a fundamental misconception about how capacitors function in such circuits. So let's make sure.

The vital fact to be remembered is that the potential difference between the plates of a capacitor cannot change instantaneously, but only as a gradual process due to current flowing in or out.

This follows from the basic equation for capacitance, as important for it as 'Ohm's law' for resistance:

$$V=\frac{Q}{C} \qquad (1)$$

in which C is any capacitance (in farads), Q the electric charge stored in it (in coulombs) and V the p.d. between its plates (in volts). We usually think in terms of current (amps) rather than coulombs, so we also have to remember that

$$Q=It \qquad (2)$$

which means that the charge Q in equation (1) is equal to the amount of current I (in amps) that has been flowing into C, and t is the time in seconds during which it has been flowing. (To make things simple we are assuming I is constant.) Putting (1) and (2) together, therefore, we see that the voltage across a capacitor cannot change unless the capacitor receives a proportionate charge, and that takes time. If time were not allowed, t would be zero, so for any charge at all I would have to be infinitely large, which is impossible.

Fig. 1 shows the classic capacitor-charging experiment. Before the switch is closed the capacitor C is uncharged, so in the basic equation (1) $Q=0$, so $V=0$. The moment the switch is closed the voltage E is applied across C and R in series. No time has elapsed since it was closed, so $t=0$, so $Q=0$, so $V=0$ still. So the whole of E appears across R. That means that a current (call it I) is flowing through R, and 'Ohm's law' tells us it is equal to E/R. That same current is flowing into C, charging it. After one second, $t=1$, so equation (2) tells us that $Q=I$. And we already know that $I=E/R$, so $Q=E/R$, so $V=E/CR$. The capacitor voltage is rising at the rate of E/CR volts per second.

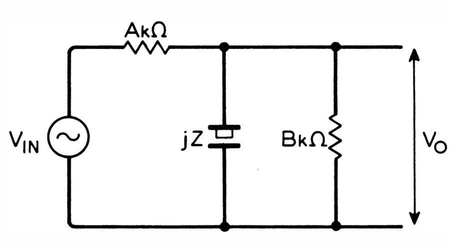


Fig. 1 The familiar circuit used to study the charging of a capacitor.

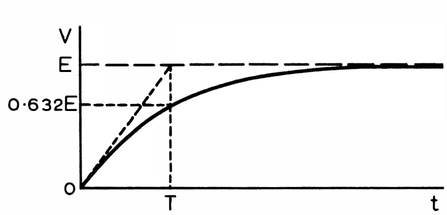


Fig. 2 The familiar (exponential) charging curve; a graph of voltage against time.

But not quite. By the end of the first second the voltage across R is no longer E; it is $E-V$. So the charging current is less than it was, so the rate of charging is less. The nearer the capacitor voltage gets to E, the less voltage is left to drive current through R and the slower the charging continues. This is shown by the familiar rate-of-charging curve, Fig. 2. Theoretically the capacitor never quite gets charged to the full voltage applied, E, but the deficiency soon becomes negligible.

To continue this lesson in elementary theory we draw the dotted line sloping upwards in Fig. 2 to show how the capacitor would charge if the starting rate could somehow be maintained. The instant at which C would be charged to the applied voltage E is indicated by the point at which the sloping line reaches the E level. Dropping a vertical dotted line from there to the time scale shows (or would do if the scale were graduated in seconds) how long this would take. As our scale is not graduated we will call the answer T.

From what we already know we can find a general formula for T. Combining equations (1) and (2) by substituting It for Q in (1) we get

$$V=\frac{It}{C}$$

At the end of our imaginary uniform-rate charge, $V=E$, $t=T$, and $I=E/R$. So

$$E=\frac{ET}{CR}$$

and for that to be true

$$T=CR$$

I'm quite sure that all the radar instructors included this result in their repertoire, whether or not they proved it in the above or any other way. T, the time a capacitance C would take to charge to the applied voltage through a resistance R if the starting rate could be maintained, is called the time constant of the series combination of C and R. If they are in farads and ohms (or more conveniently in microfarads and megohms) T will be in seconds.

Because it refers to a mode of charging that doesn't exist in normal practice you might consider all this a waste of time. But as we noted earlier one cannot say how long a capacitor takes to charge in the real practical Fig. 1 way, because theoretically it always takes an infinitely long time, and that is not a very helpful piece of information. The only thing left, then, is to decide on how charged is 'charged'; 99%, say?

The mere suggestion may bring before you a vision of endless committee meetings all over the world trying to agree on a percentage to use as an international standard. Happily, there is no need for this. It turns out that the actual charging curve in Fig. 2 has a fixed shape, so that the time taken to charge to any given percentage of 'full' is an easily calculated factor of T, which is so simply equal to CR. The simplest possible factor is of course 1, and it happens that CR is the time taken to charge to 63.2% of 'full', as shown in Fig. 2. That looks like rather short measure. 99% requires an odd factor, so I suggest a choice of either $4CR$ (for 98.17%) or $5CR$ (for 99.33%).

The radar instructors probably mentioned the name of the curve of this particular shape (the exponential curve) but they may well have decided it was unnecessary (for the purpose of fitting people to keep radar equipment working) to go into the mathematics of the thing. I too am saying it is unnecessary for our present purpose, and anyone who really wants to know can find it in almost any of the textbooks on electricity (with or without magnetism). The only vital point to carry away just now is that some idea of how long in seconds C μF takes to charge through R MΩ is given by multiplying C by R, and that charging is practically complete in 4 or 5 times CR.

Now we have got the basic principles straight we can apply them to a circuit of the type which might have given rise to the lecture on CR time constants. It is a circuit in which a square wave developed in the output of one stage has to be passed on to the input of another stage for amplifying, blanking, gating or whatever. Fig. 3(a) shows the relevant part of such a system. Valves are shown, because they were used in wartime radar and because in many cases the input of the second stage had such a high resistance that R was not appreciably shunted by it. Fig. 3(b) is a transistor equivalent for the benefit of those to whom valves are devices that used to be used, too long ago to be worth trying to understand. But an allowance will have to be made for the shunting of R.

The square input waveform is shown in Fig. 3(a), and the object is to reproduce it, with as little distortion as possible, at the

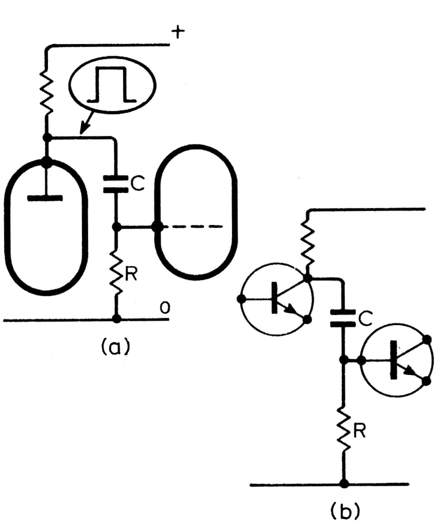


Fig. 3 The part of a circuit in which the theory developed in Figs. 1 and 2 is useful: (a) the valve version considered, and (b) its transistor equivalent.

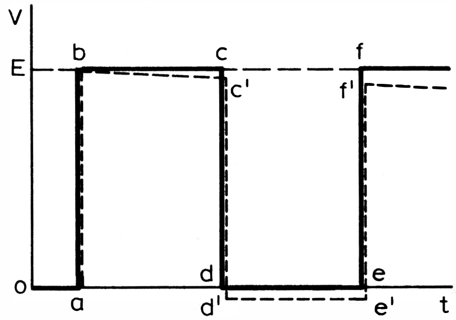


Fig. 4 The solid-line square wave is the input to CR (shown in Fig. 3), less any continuous voltage bias, and the dotted line is the output at the junction of C and R.

input to the next stage—i.e., the junction of C and R. Of course if direct coupling is used C and R are not needed and distortion does not arise, but with valve circuits especially it is usually necessary to maintain a fixed p.d. between the two stages by means of C, to keep the electrode working voltages right. When considering signal voltages this fixed p.d. can be ignored. So in the signal-voltage/time graph (Fig. 4) we can assume both the input voltage (applied across C and R) and the resulting output voltage (across R) start from zero level.

Up to the point on the time scale marked a the input signal voltage remains at zero, and so does the output, so there is no voltage across the capacitor, so (as equation (1) tells us) it must be totally uncharged. But at a the input suddenly goes E volts positive. (Of course it can't do this absolutely instantaneously, but let us suppose that compared with the time ad the rise time is negligible.) This is the point at which I have heard instructors go on to say 'so C charges, making the output (which is the input to the next stage) positive'. But I have, I hope, by now convinced even the most instructor-loving reader that it just isn't possible for C to charge appreciably during the rise time, and the fact that the output follows the input and goes positive to the same extent is actually evidence of it. In other words, C does this part of its job by *not* charging. For as long as it stays uncharged, both sides of it are at the same potential and the output is an exact undistorted copy of the input waveform. The ideal, then, is for C never to be charged, at all.

Let us now consider the state of affairs from b to c. Because the input, E volts, is applied across C and R, and the voltage across C alone (at b) is zero, the whole E comes across R, causing a current to flow through it. Assuming (as we did) that the second valve takes no grid current, all the current has to go into C, beginning to charge it at a rate of E/CR volts per second. The voltage now rising across C is no longer available for R as output voltage. So the output voltage falls. How much it falls in the period bc depends on the time constant, CR. If the output is to be undistorted, it mustn't fall at all; which means that CR must be infinite. It can be made very nearly so by removing R altogether, leaving a gap. But then the grid potential would be at the mercy of stray circuit leakages. To ensure that it starts definitely from zero (or any other designed voltage) R must be used, but its resistance should be made not lower that is needed to anchor the grid to zero volts. Provided that C also is made large enough, the drop in output signal voltage, represented in Fig. 4 by cc', can be kept small, as shown. Incidentally, because the voltage across R is nearly constant, the rate of charge is nearly constant and bc' is nearly a straight line.

At c the input returns abruptly to zero volts *(d)*, and as the p.d. across C cannot change so quickly the grid side of C drops by the same voltage *(E)*. As it started from c', less than $+E$ volts, it now goes slightly negative, d'. This negative voltage, dd', to which C became charged during the period bc', is now applied to R, through which the charge leaks away during the period $d'e'$. Because the voltage is so small the rate of discharge is very small and $d'e'$ is practically horizontal. So when the input goes positive again, from e to f, the output at f' is practically the same as at c'. It therefore starts its decline during the next positive half-cycle from a lower voltage than it did in the first.

Effect of d.c. barrier

So long then as the output half-cycles continue to be more positive than negative, the different rates of charge and discharge bring them gradually more nearly equal, as shown by the dotted waveform in Fig. 5. In the end, whatever the input waveform, the output will arrange itself so that the time×voltage area below the line is equal to that above the line. The line, of course, represents the level to which the output is anchored by R; in this case zero volts. This phenomenon, which we have been examining in detail, results inevitably from the fact that a capacitor is a barrier to d.c. So a signal that starts (as in Fig. 4) all above the line, or more one side of the line than the other, inevitably adjusts itself so that this d.c. component disappears and the output is wholly alternating. The less the time constant CR the faster it adjusts—and the more distortion it introduces.

If the signal frequency is very low, so that C has a long time in which to discharge during each half-cycle, a very long time constant is needed to avoid appreciably distorting a square wave. And the system takes a very long time to readjust to a change of input amplitude. This problem arises in oscilloscopes where capacitance couplings are used in the deflection ampli-

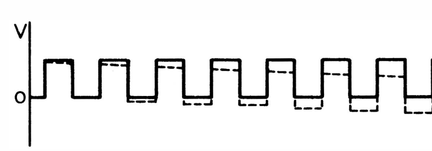


Fig. 5 How the voltage/time graph started in Fig. 4 continues.

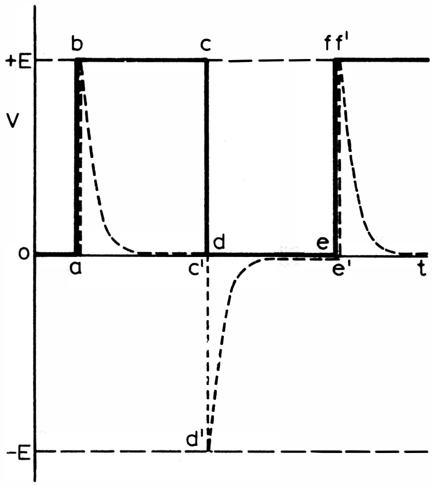


Fig. 6 Here for comparison with Fig. 4 is what happens when the time constant is only a fraction of one half cycle.

www.keith-snook.info/capacitor-soakage.html

fiers. It is so tedious waiting for them to settle down that nowadays designers almost always provide direct-coupled amplifiers.

The d.c.-losing effect can be prevented by suitably connecting a rectifier in the circuit, creating a 'd.c. restorer'—but that is another story.

The only other thing I think I need mention—and it will be familiar to radar trainees past and present—is that a CR coupling is often used not to pass on the original undistorted form but to introduce deliberate distortion. The commonest application is for changing square waves into brief pulses. For this purpose the time constant is made much less, so that instead of a gradual charge such as *bc'* in Fig. 4 the capacitor charges practically completely within the half-cycle, as in Fig. 6. When the end of the square-wave half-cycle comes *(cd)* the output going negativewards by the same amount (*c'd'*) yields equal negative and positive half-cycles from the start. The negative ones can then be removed by a rectifier and the positive ones clipped by another, to give a train of pulses.

Note that (whatever the instructor said) C charges from *b* to *c'* and discharges from *d'* to *e'*, in Fig. 4 and in Fig. 6.

I used to find that even fellows who could state Kirchhoff's voltage law quite correctly when asked for it seemed to forget all about it when considering the CR type of circuit. One form of the law says that the sum of the voltages across the components in a series circuit is equal to the voltage applied. Now in Figs. 4 and 6 the voltage applied is represented by the height above zero of the 'input' waveform: alternately *E* and O. The Voltage across R ('output') is represented by the height of the dotted line, so the voltage across C (due to its charge) must, by Kirchhoff's law, be the vertical difference between the two. Looking at the matter this way, one can be in no doubt about when and how much the capacitor is charging and discharging.

The essential thing is to grasp the message of Figs. 1 and 2. Then a correct view of the action of any CR circuit is (to coin a phrase) a piece of cake.

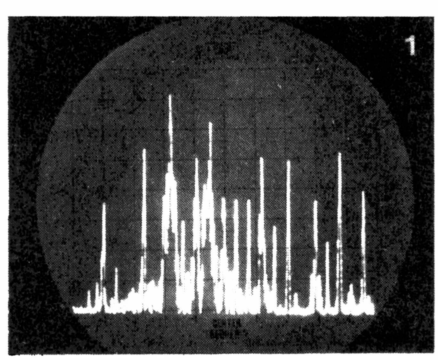
Single-sideband Experimental Broadcasts

For some years there have been discussions on the possibility of utilizing the medium-wave sound broadcasting band more effectively by means of single-sideband transmissions. At first sight it seems attractive in view of the fact that s.s.b. is now so well established in h.f. communications. But there are complications in reception, the main one being that the simple envelope detector found in conventional sound receivers inevitably leads to excessive distortion and must be replaced by a product detector, in which case, for tuning, a local oscillator of high stability, among other things, is required. In Britain the broadcasting authorities don't seem very enthusiastic about s.s.b. but in Germany there is considerable interest—measured by the fact that the Deutschlandfunk broadcasting organization has been putting out experimental s.s.b. transmissions from its station at Mainflingen, near Frankfurt.

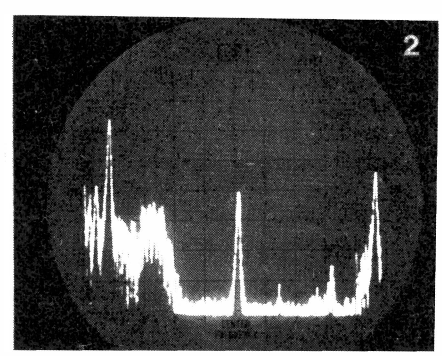
The broadcasts took place in the early hours of the morning, after close-down of normal broadcasting, on 1475 kHz. At least one group of British radio research people was willing to stay up in order to study and listen to the transmissions. This was a radio section of the Department of Electrical and Electronic Engineering at University College Swansea, headed by Dr. R. C. V. Macario (author of an article on an s.s.b. receiver module in the July issue). Some results of their monitoring are shown in the accompanying frequency spectra. Fig. 1 is a 200 kHz wide part of the m.f. spectrum showing the s.s.b. transmission at 1475 kHz, in relation to the permanent a.m. transmission from the Mainflingen broadcasting station on 1538 kHz and to Radio Luxembourg on 1439 kHz. More detail can be seen in Fig. 3, which is 50 kHz wide. The upper sideband of the s.s.b. transmission can be seen as an asymmetrical distribution of energy in contrast to the symmetrical distributions, like church spires, of the a.m. stations on each side of it. In Fig. 4 the frequency scale is 20 kHz wide and shows the upper sideband in even greater detail.

The carrier alone of the s.s.b. transmission was suppressed 20dB below the peak sideband levels, and is shown in Fig. 2, on a frequency scale 20 kHz wide.

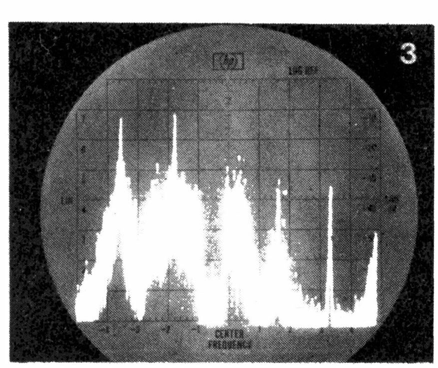
The spectra were displayed on a Hewlett-Packard spectrum analyser, model 8552A/8553L, with a stored display. A simple roof wire aerial was used. Recordings of the transmissions were made via various receiving systems, but it is interesting to note that direct conversion was possible since the lower sideband of the transmission was relatively free of interference.



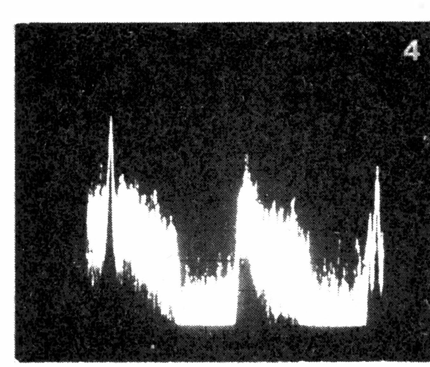
Luxembourg s.s.b. Mainflingen



s.s.b carrier only



u.s.b. 1475 kHz



u.s.b. 1475 kHz