THE QUEEN'S AWARDS FOR ENTERPRISE 2001

BBC

R&D White Paper

WHP 106

February 2005

How does one calculate and measure AGC time-constants?

R.H.M. Poole

Research & Development
BRITISH BROADCASTING CORPORATION

BBC Research and Development
White Paper WHP 106

How does one calculate and measure AGC time-constants?

R H M Poole

Abstract

This White Paper is the result of field trials carried out on some portable radio equipment. During the trials, it became evident that the automatic gain control (AGC) of the receiver was operating more slowly than desirable — a rapid change of incoming signal level caused the receiver output to 'bounce'. As a consequence, the output could temporarily either clip or sink into the noise.

In this case, alteration of the AGC time-constant proved to be an easy cure for the problem. However, there must be many other cases where an understanding of AGC systems would help in avoiding similar pitfalls. This White Paper describes, in simple terms, how to calculate and measure either the time-constant or its close relation the loop bandwidth. The Paper also introduces an *Excel* spreadsheet to simulate the effect of the AGC on any amplitude disturbance. Only a single-loop AGC system is considered, but it would be easy to extend the work to higher order systems.

White Papers are distributed freely on request.

Authorisation of the Chief Scientist is required for publication.

BBC Research and Development
White Paper WHP 106

How does one calculate and measure AGC time-constants?

1. Introduction

Recently, field trials were carried out on some portable radio equipment. During the trials, it became evident that the automatic gain control (AGC) of the receiver was operating more slowly than desirable — a rapid change of incoming signal level caused the receiver output to 'bounce'. As a consequence, the output could temporarily either clip or sink into the noise.

In this case, alteration of the AGC time-constant proved to be an easy cure for the problem. However, there must be many other cases where an understanding of AGC systems would help in avoiding similar pitfalls. This White Paper describes, in simple terms, how to calculate and measure either the time-constant or its close relation the loop bandwidth. The Paper also introduces an Excel spreadsheet to simulate the effect of the AGC on any amplitude disturbance. Although only a single-loop AGC system is considered, it would be easy to extend the work to higher order systems.

2. AGC Systems

A single-loop AGC system comprises three essential elements: an amplifier whose gain may be varied by means of a control voltage, a signal level detector, and a comparator/integrator. The output of the integrator is connected back to the control input of the amplifier, hence forming a closed loop:

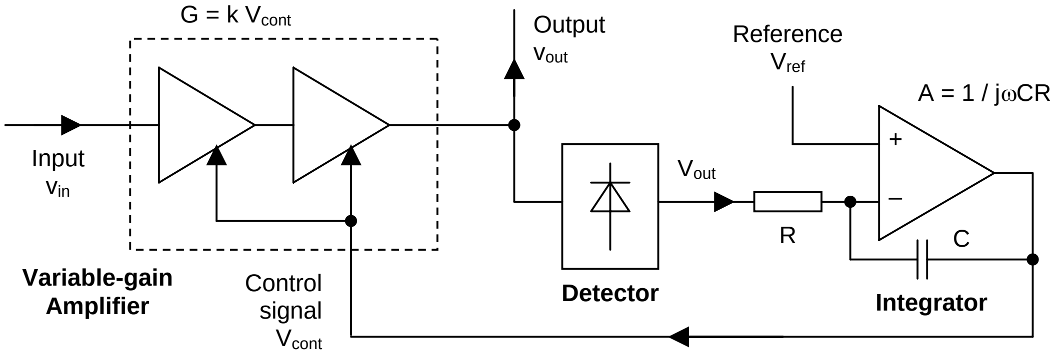


Figure 2.1: A Single-loop AGC System

When the system is in equilibrium, the detector output must always equal the AGC reference V_{ref}: any error will cause the control voltage V_{cont} to ramp up or down, so increasing or decreasing the amplifier gain as appropriate. Note the convention adopted in the diagram is for lower case 'v' to represent the amplitude of an RF signal, whereas the upper case 'V' refers to a direct voltage. It is assumed that the detector generates a direct voltage equal to the amplitude of the RF signal at the output, so that $V_{out} = v_{out}$.

Although the system appears simple enough, a rigorous mathematical treatment is not straightforward. The main difficulty is that the control mechanism in the voltage-controlled amplifier involves a multiplication rather than an addition. What is more, the multiplication factor k is — or should be — a function of the gain G. Where only the loop bandwidth is needed, a simplified approach is possible.

Firstly assume that the system is handling DC and not an RF signal. It is reasonable to do this because DC coupling has been used throughout. An input of $(1 + m_{in})$ then corresponds to a steady RF signal of amplitude 1 which is modulated by a sinusoid of amplitude m_{in} — the steady signal has a disturbance which the AGC will attempt to remove. Secondly, assume that the output is $(1 + m_{out})$; in other words, the gain of the system to the wanted signal is unity, but m_{in} has been reduced to m_{out} by the action of the AGC. In order for the gain to be unity, take the control voltage to be $V_{(1)}$.

The final assumption that needs to be made is the effect of a small control voltage variation, m_{cont}, on the output. Since the amplifier effectively multiplies the input signal by the control voltage, an input of amplitude 1 gives rise to an output of amplitude $k\ m_{cont}$. (k is the multiplication factor of the gain control mechanism.)

The simplified circuit appears as

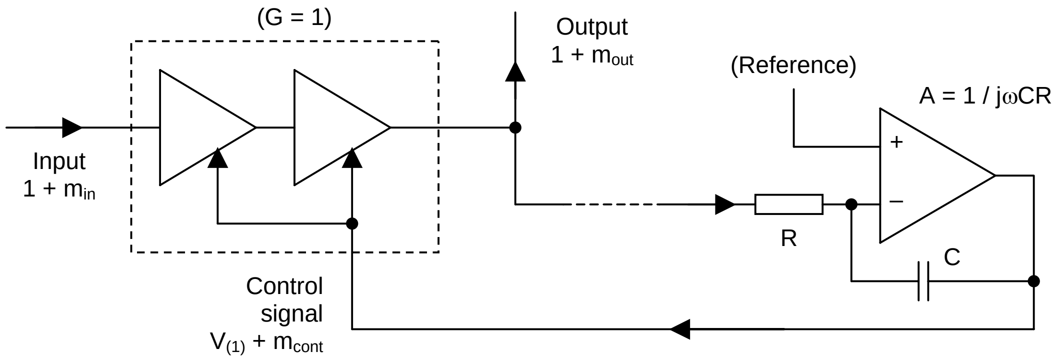


Figure 2.2: The Single-loop AGC System with Varying Input

The analysis is now straightforward. The output is the sum of the input and the disturbance generated by the control voltage:

$$(1 + m_{out}) = (1 + m_{in}) + k\ m_{cont}, \qquad \text{or} \qquad m_{out} = m_{in} + k\ m_{cont}.$$

Also, m_{cont} equals m_{out} after filtering by the loop integrator:

$$m_{cont} = m_{out} \times (1/j\omega CR) .$$

Substituting the expression for m_{cont} into the first equation gives:

$$m_{out} = m_{in} + k\ m_{out} \times (1/j\omega CR) .$$

Hence, $$m_{out}(1 - k/j\omega CR) = m_{in}, \qquad \text{or} \qquad m_{out}/m_{in} = 1 / (1 - k/j\omega CR) .$$

In other words, the AGC loop has a conventional first-order response, and will reduce an incoming disturbance by a ratio of $1 / (1 - k/j\omega CR)$, as illustrated in the plot overleaf. The 'relative frequency' equals the ratio ω/ω_0, where ω_0 is k/CR.

When the angular frequency ω is large compared with ω_0, the loop has little effect: the 'gain' $|m_{out}/m_{in}|$ is close to unity, or 0 dB. When ω is small compared with ω_0, $|m_{out}/m_{in}|$ falls by a factor of 2, or 6 dB, per halving of ω. In other words, halving the frequency of an incoming disturbance doubles the ability of the AGC loop to remove it.

The phase shift is also of interest. At high frequencies, where the loop has little effect, the phase shift is zero, but at low frequencies any disturbances not absorbed by the loop are advanced in phase by 90°. More is said about this in Section 7.

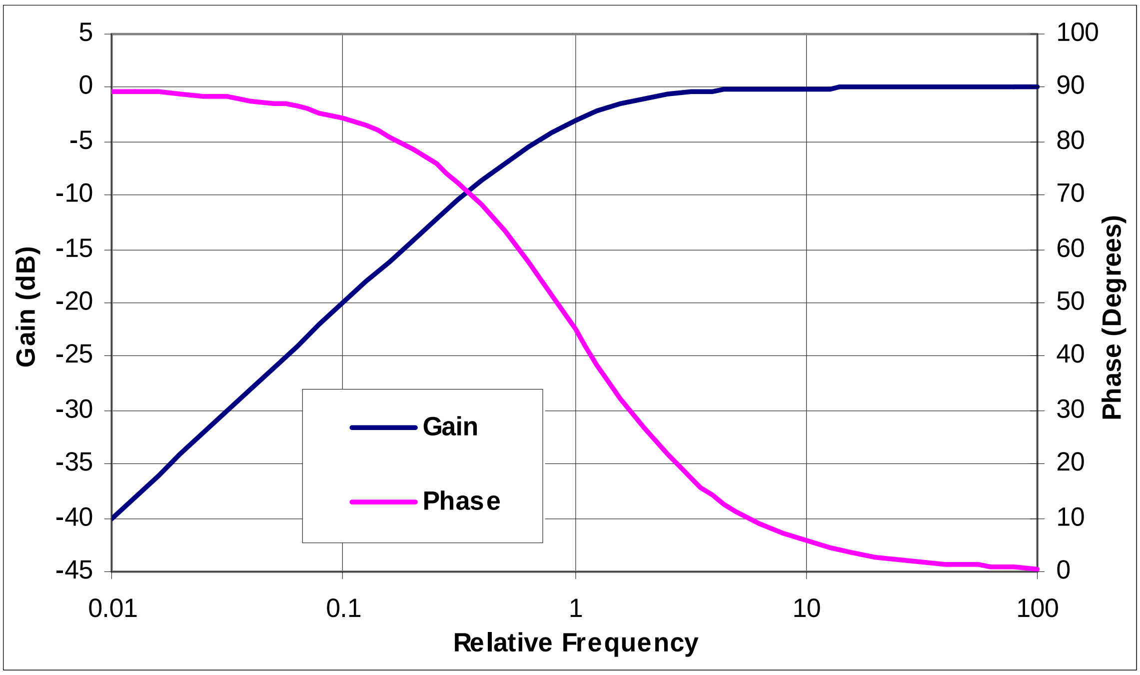


Figure 2.3: The Response of the AGC System to Input Variations

3. A Practical AGC System

Some life can be put into the above analysis by taking a real example. The loop integrator components of the previously mentioned radio receiver are 220 kΩ and 100 nF, and k is given by the manufacturers of variable-gain amplifier as equivalent to 34 dB per volt. However, some care is needed in interpreting k before including it in the formula for the loop time-constant. Note that the calculations implicitly assumed small quantities for m_{out}, m_{in} and k, and an output level of 1 V.

k is calculated as follows. The actual output level (or reference voltage) used in the receiver is 150 mV. Suppose that the disturbance on the control voltage m_{cont} is 1 mV. The resulting gain change is 34/1000, or .034 dB, corresponding to a ratio of 1.004. 1.004 × 150 mV equals 150.6 mV, and so the disturbance present on the output, k m_{cont} , is 0.6 mV. In other words, k is 0.6 — perhaps a surprisingly low value for so many dB per volt.

The loop time-constant CR/k is therefore $(100 \times 10^{-9}) \times (220 \times 10^{3}) / 0.6$, or .037. The corresponding frequency ($\omega_0/2\pi$) or loop bandwidth is 1/ (2π × .037), or 4.3 Hz.

Such a loop bandwidth is excessively low for the likely signal conditions. On the other hand, making the bandwidth too large would distort the signal by removing its valid amplitude variations. For this application, 100 Hz was chosen as a reasonable compromise, and was obtained by changing R from 220 kΩ to 10 kΩ.

A check of the loop bandwidth, as described in the next section, showed it to vary between 100 Hz and 200 Hz. There are two likely reasons for the variations. Firstly, the control voltage sensitivity of the amplifier is not constant. Evidence for this is the signal strength indication, which is derived from the control voltage, and which can be in error by as much as 5 dB. Secondly, the calculations do not take into account the additional AGC for the receiver front-end. The front-end AGC only comes into effect at high signal levels; when it does so, it increases the effective k.

4. Checking the AGC Loop Bandwidth

Theorising is dangerous, and it is always wise to make sure that a modification, such as the one described above, has had the desired consequences. Fortunately, measuring the AGC loop bandwidth is easy to do. All that is needed are a signal generator with an AM facility and a spectrum analyser that can be set to zero span:

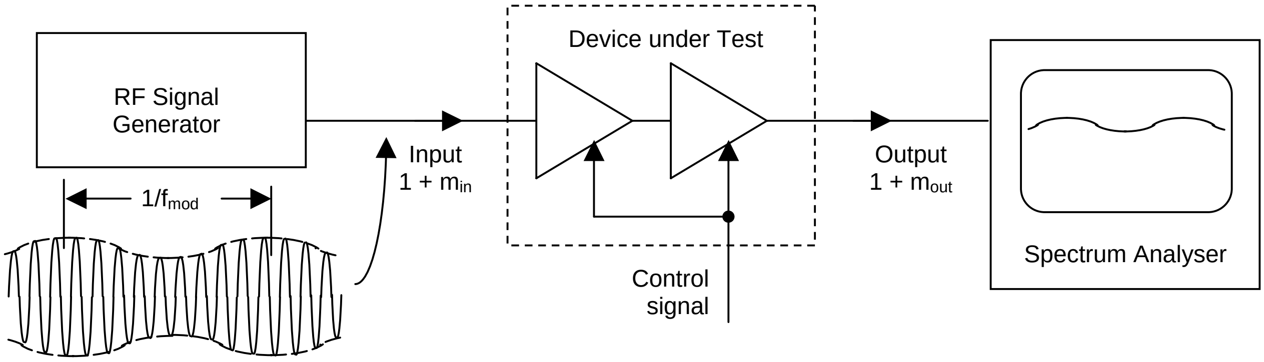


Figure 4.1: Set-up for Checking the AGC Loop Bandwidth

The signal generator is set to the desired carrier frequency and power, and a small amount of AM introduced — perhaps 10%. The output of the device under test is fed to the spectrum analyser, which is set to the same carrier frequency and zero span. If the vertical scale is set to 'linear', and the triggering controls adjusted appropriately, the modulation envelope should be visible on the display. Obviously, the resolution bandwidth must be sufficient to include the modulation sidebands.

Firstly, the modulation frequency is set to a value well outside the loop bandwidth — say by a factor of 10 — and the amplitude of the modulation envelope measured on the spectrum analyser. Secondly, the modulation frequency is reduced until the observed amplitude falls by 3 dB ($\div\sqrt{2}$). That frequency corresponds to the loop bandwidth.

If the frequency is reduced further, the amplitude should fall by 6 dB ($\div 2$) for each halving of frequency within the loop bandwidth. The results should be checked at different input levels, as the control's dB per volt (k) might not be constant.

5. Checking the AGC Loop Transient Response

A simple AGC loop is a nearly pure first-order system and should be well behaved. This feature is a welcome contrast to the standard synthesiser control loop, which is second-order and hence only conditionally stable. However, there are always opportunities for unwanted additional time-constants to creep in and cause trouble. Some sort of transient test is useful to seek these out.

An easy test is to switch the modulation on the test signal to square wave, as illustrated overleaf. The frequency of the square wave should be well within the loop bandwidth.

In an ideal world, the modulation envelope of the output signal would be a differentiated square wave, with the tops of the square wave sagging exponentially. The time-constant of the sag would correspond to that of the AGC loop. In practice, bandwidth limitations within the system slow the leading and trailing edges, as shown in the diagram. Any oscillation or 'ringing' on the waveform should be regarded with suspicion, and the cause removed.

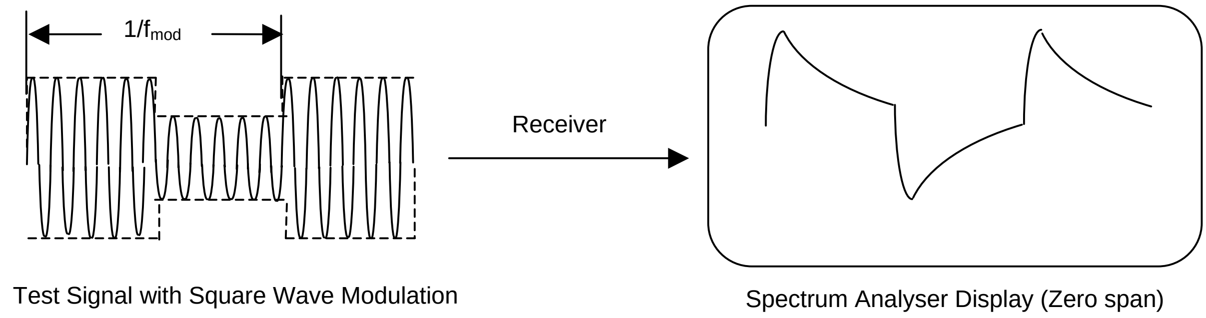


Figure 5.1: Checking the AGC Loop Transient Response

Another good test is to modulate the AGC reference voltage instead. In that case, the square-wave modulation should appear on the output signal, but with rise and fall times determined by the AGC time-constant. Once again, any oscillation should be investigated.

6. Non-Linearity within the AGC Loop

In the earlier analysis, it was hinted that a full mathematical treatment of AGC systems is difficult because of their inherently non-linear nature. It is worth having a quick qualitative look at the causes and effects of non-linearity.

Firstly, the 'ideal' control characteristic for the variable-gain amplifier is exponential; that is,

$G = \exp(aV_{cont} + b)$, where a and b are constants, G is the gain, and V_{cont} is the control voltage.

To see why this characteristic is ideal, differentiate this expression with respect to V_{cont}:

$\delta G/\delta V_{cont} = a \exp(aV_{cont} + b)$, where a and b are constants.

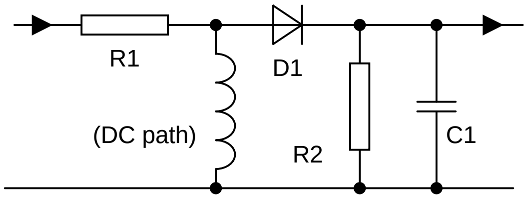
Hence $\delta G/\delta V_{cont} = a\,G$,

or $\delta G/G = a\,\delta V_{cont}$.

In other words, the same change in control voltage always results in the same fractional change in gain, whatever the absolute control voltage or gain.

However, the result of the 'ideal' exponential characteristic is that a sinusoidal component of the control voltage will not cause a sinusoidal modulation of the RF signal. The AGC loop will reduce the unwanted amplitude modulation (m_{in}) present on the incoming signal, but the residual modulation (m_{out}) will be distorted. During measurement of the loop bandwidth, misleading results are avoided by keeping the modulation small.

Secondly, the detector is unlikely to be linear, especially if a diode is used. A typical arrangement is as follows:

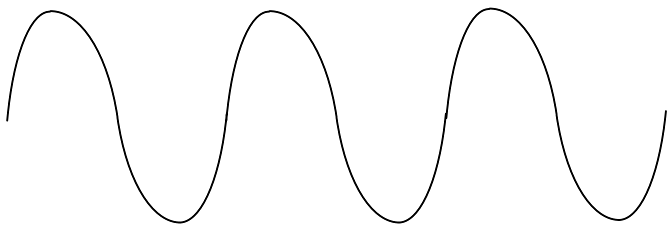


R1 is generally much smaller than R2. An increasing signal level at the input to the detector results in the output DC following faithfully. But, if the input signal falls rapidly, D1 can no longer conduct until C1 has discharged sufficiently through R2. Hence the detector does not respond symmetrically to rising and falling signal levels.

Figure 6.1: A Diode Detector Circuit

Normal practice is to arrange for the time-constant C1.R2 to be short compared to the period of the modulation. However, the fast attack / slow decay characteristic of this sort of circuit can be useful. If C1.R2 is made large and the AGC loop integrator replaced with a straight amplifier, the AGC circuit will maintain the peak signal level constant, not its mean.

The receiver was tested with small amounts of amplitude modulation — typically 10%. When the level was increased to 50% or so, the modulation envelope at the output was noticeably distorted:



The effect of non-linearity is to sharpen the rising edges of the modulation envelope at the expense of the falling edges. Hence the waveform appears skewed.

Figure 6.2: The Distorted Modulation Envelope

7. *Excel* Helps out

The above AGC system described is simple to simulate with the help of an *Excel* spreadsheet; only five columns are needed:

Phase (Deg)	Phase (Rad)	Input (1 + m(in))	Output (1 + m(out))	AGC m(cont)
0	0	1	1	0
5	0.087266	1.043578	1.043578	0
10	0.174533	1.086824	1.078589	-0.00761

Figure 7.1: The *Excel* Spreadsheet

The phase of the sinusoidal disturbance m_{in} (shown as m(in) in the spreadsheet) is stepped in increments of 5°, and the complete input signal $1 + m_{in}$ appears in the third column. $1 + m_{out}$ — the output signal — is calculated in the fourth column, and is simply the input signal multiplied by the exponent of the AGC voltage m_{cont}. The only clever bit is the calculation of the AGC voltage itself. Since the AGC voltage is dependent on the output signal, which in turn depends on the AGC voltage, something has to be done to prevent a 'howl around'.

The trick is to derive the AGC from the *previous* sample of the output. m_{cont} is taken to be the difference between the previous sample of the output and the AGC reference (normally 1), which is then added to the AGC voltage that has already accumulated. The spreadsheet includes a scale factor for the ratio of modulation frequency to integrator time-constant, and another for the fraction of the period between samples. It is possible to start the AGC voltage at a predetermined value, although normally this would be set to zero.

The results for $k/\omega CR = 2$ and a sinusoid of amplitude 0.5 are shown overleaf. Two factors are of interest. Firstly, the modulation envelope is distorted because the ripple on the control voltage has an exponential effect on the gain. Secondly, the modulation on the output is advanced in phase by nearly 90° compared to that of the input. The reason is that the loop integrator adds 90° delay to the control voltage, but this delayed signal must be in phase with the input signal when the loop is functioning. Hence the output must be in advance by 90° for this to be true. (This requires some thought!)

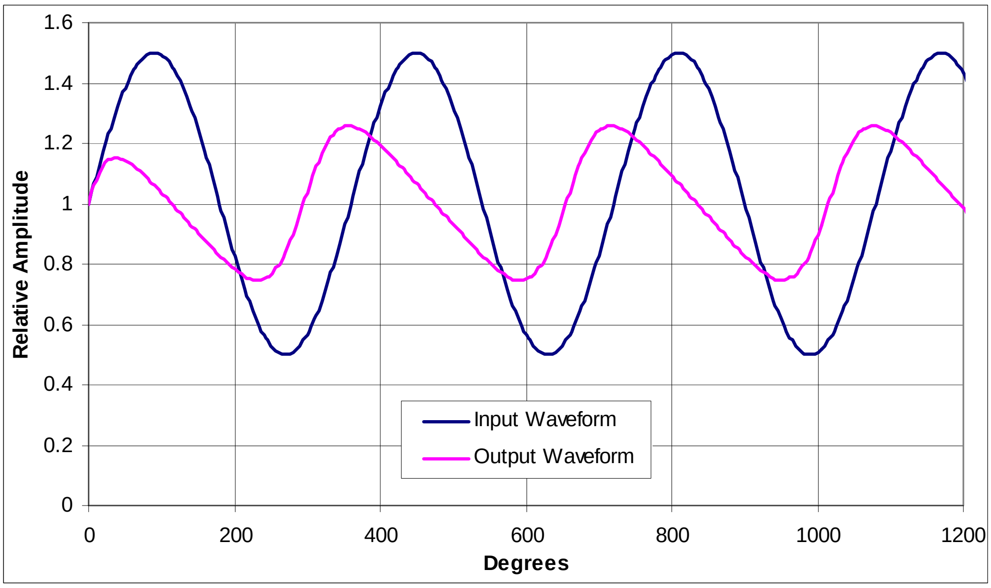


Figure 7.2: A Simulation Using the *Excel* Spreadsheet

The spreadsheet is readily adapted to include square wave (or any other) modulation. If the square wave has amplitude 0.5, and if k/ωCR = 1, the results are as shown below. Note that the effect of the exponential control characteristic is to enhance the positive-going transients at the expense of the negative-going transients. Fortunately, the plots accord well with what is observed in practice.

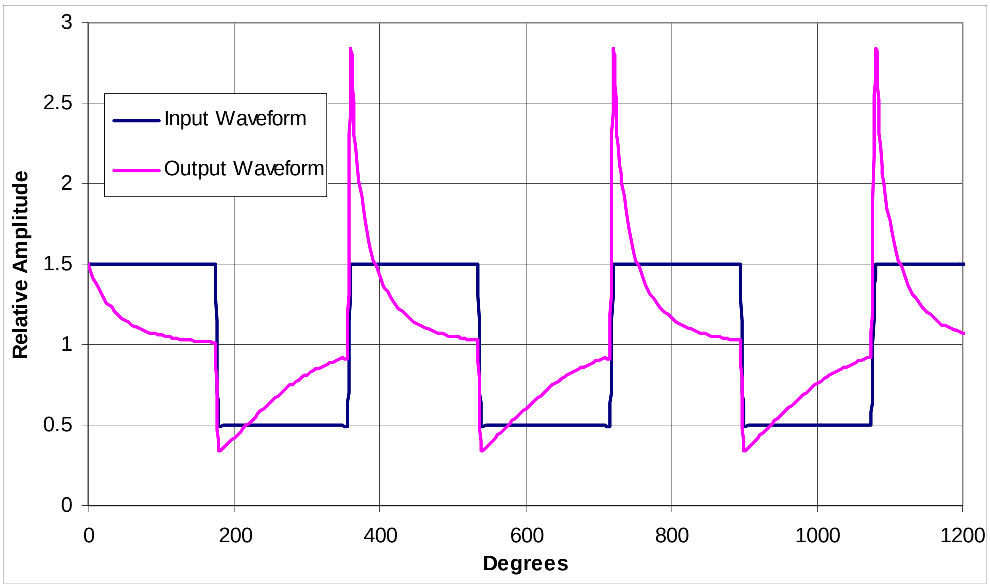


Figure 7.3: A Simulation of the AGC Loop Transient Response

Perhaps these results for first-order systems are not very exciting, as such systems are inherently well behaved. However, the simulation could easily be extended to higher-order systems where good behaviour cannot be taken for granted. For instance, a common arrangement includes a second gain-controlled stage close to the receiver input, the purpose of which is to prevent high signal levels from overloading the receiver. The control voltage for the second stage is derived, perhaps by means of an integrator, from the control voltage for the first.

8. Summary

This White Paper has shown how to calculate AGC time-constants, and has explained how to check these by measurement. It has also introduced an Excel spreadsheet for simulating the effect of the AGC system on the envelope of an RF signal.

Although only simple, first-order systems have been discussed, the work could readily be extended to cover more complex systems.

9. Tools

The Excel spreadsheet is available here:
http://www.bbc.co.uk/rd/pubs/whp/whp-pdf-files/whp106excel.xls