

Automatic Gain Control (AGC) in Receivers

Iulian Rosu, YO3DAC / VA3IUL – <http://www.qsl.net/va3iul/>

AGC was implemented in first radios for the reason of fading propagation (defined as slow variations in the amplitude of the received signals) which required continuing adjustments in the receiver's gain in order to maintain a relative constant output signal.

Such situation led to the design of circuits, which primary ideal function was to maintain a constant signal level at the output, regardless of the signal's variations at the input of the system.

Now AGC circuits can be found in any device or system where wide amplitude variations in the output signal could lead to a lost of information or to an unacceptable performance of the system.

Automatic Gain Control (AGC) circuits are employed in many systems where the amplitude of an incoming signal can vary over a wide dynamic range. The role of the AGC circuit is to provide a relatively constant output amplitude so that circuits following the AGC circuit require less dynamic range.

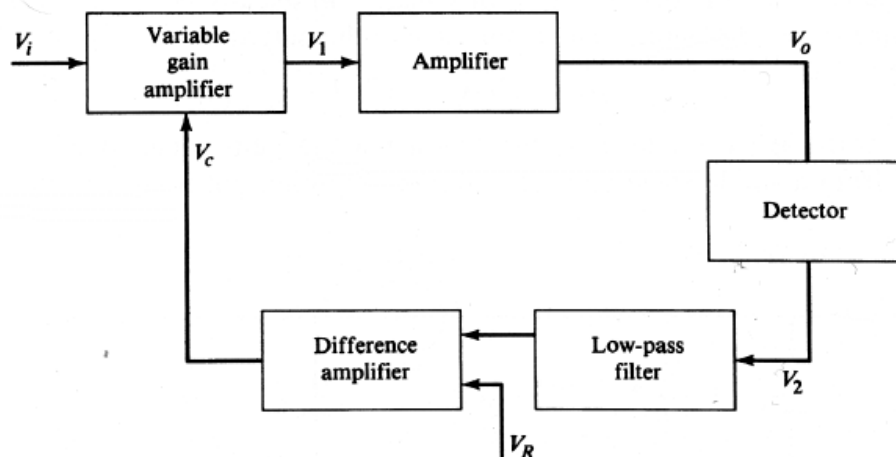
If the signal level changes are much slower than the information rate contained in the signal, then an AGC circuit can be used to provide a signal with a well defined average level to downstream circuits. In most system applications, the time to adjust the gain in response to an input amplitude change should remain constant, independent of the input amplitude level and hence gain setting of the amplifier.

The large dynamic range of signals that must be handled by most receivers requires gain adjustment to prevent overload or IM of the stages and to adjust the demodulator input level for optimum operation.

- A simple method of gain control would involve the use of a variable attenuator between the input and the first active stage. Such an attenuator, however, would decrease the signal level, but it would also reduce the S/N of any but the weakest acceptable signal.
- A typical AGC loop circuit is shown in the picture below. This is a feedback system comprising a forward controlled gain stage, feedback gain and a signal comparison stage that generates a differential error signal.

The AGC loop is analyzed in terms of its closed-loop gain (forward transfer function) and open-loop gain. There is temperature compensation at the detector diode stage to compensate for diode detector forward voltage variation with temperature.

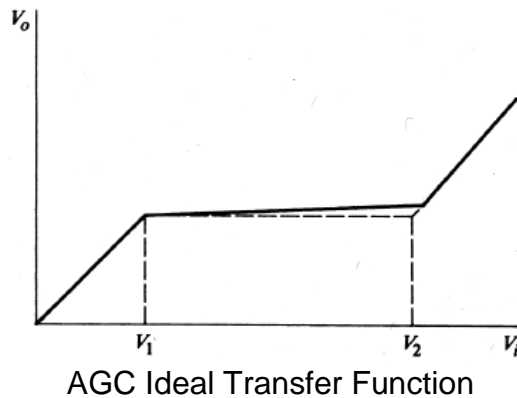
- Gain control is generally distributed over a number of stages, so that the gain in later stages (the IF amplifiers) is reduced first, and the gain in earlier stages (RF and first IF) is reduced only for signal levels sufficiently high to assure a large S/N.
- If the RF gain is small is enough switching in/out an attenuator at RF only for sufficiently high signal levels. Variable gain control for the later stages can operate from low signal levels. Variable-gain amplifiers are controlled electrically, and when attenuators are used in receivers, they are often operated electrically either by variable voltages for continuous attenuators or by electric switches (relays or diodes) for fixed or stepped attenuators.



AGC Block Diagram

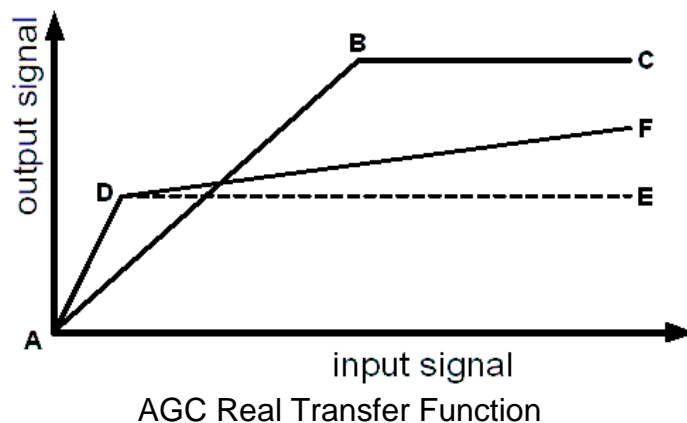
The input signal is amplified by a Variable Gain Amplifier (VGA), whose gain is controlled by an external signal VC. The output from the VGA can be further amplified by a second stage to generate an adequate level of V_o . Some of the output signal's parameters, such as amplitude, carrier frequency, index of modulation or frequency, are sensed by the detector; any undesired component is filtered out and the remaining signal is compared with a reference signal. The result of the comparison is used to generate the control voltage (V_c) and adjust the gain of the VGA.

- If the control time constants are determined primarily by the detector circuit and the additional amplifier has a wider bandwidth than the detector, then the attack and decay times will be shortened by the amount of the post-amplification
- An AGC circuit in the receiver provides a substantially constant signal level to the demodulator independent of the input signal level.



For low input signals the AGC is disabled and the output is a linear function of the input, when the output reaches a threshold value (V_1) the AGC becomes operative and maintains a constant output level until it reaches a second threshold value (V_2). At this point, the AGC becomes inoperative again; this is usually done in order to prevent stability problems at high levels of gain.

- If the gain loop is much greater than 1, the steady state change in the input is greatly reduced.



The line A, B, C represents a system that has no AGC applied. The output increases linearly with the input signal until point B is reached, when some element in the signal chain overloads and becomes non-linear.

Generally from point B to C the output signal is distorted and, unless the input signal is reduced, the system is unusable.

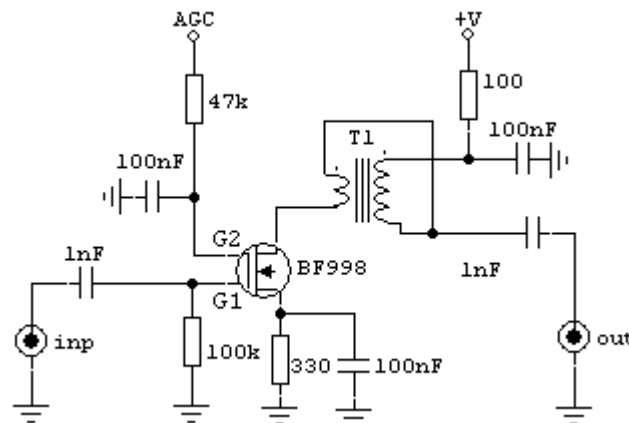
Increasing this value increases the slope of the line A to B and reduces the input signal level at which the signal distorts.

The line A, D, E represents a system that has AGC applied. The slope A, D is greater than unity and indicates that the AGC has gain prior to the AGC detector. The transition from a linear to constant output at D is known as the AGC ‘knee’ or threshold. From D to E the output level does not increase in response to an increase in input signal.

How flat the section of the line D to E is depends on the overall AGC loop gain and is called the “AGC Slope”.

- The AGC knee should not be set at too low input level.
- Above the AGC knee the AGC characteristics should be not flat, and must have a slope between 6dB and 10dB.
- The AGC system is basically a feedback amplifier and has a closed-loop gain characteristic, which is essentially low-pass.
- AGC systems that include a reference voltage inside the control loop are referred as delayed AGC.
- Many direct conversion receivers do not use AGC. The advantages of not using AGC are simplicity and the purity of the received signal, in that weak received signals sound weak and strong signals sound loud.
- An incorrectly designed AGC system can introduce considerable distortion to an otherwise clean signal.
- Within the noise and distortion limits of the VGA, the AGC loop is set up to provide both signal amplification and compression to maintain the output signal within the specified limits.
- If you need lower second-harmonic distortion, consider lowering the maximum AGC input level at the expense of dynamic input range.
- If the receiver’s LNA stage is to be AGC controlled, it is almost always critical to maintain its noise figure within reasonable limits. Unfortunately, any AGC action will naturally decrease the receiver’s NF, but by adding a *delay diode* in series with the AGC’s bias line, the start of gain control to the front-end amplifiers can be postponed slightly. This will permit the LNA to maintain its NF and gain, and therefore the noise performance of the entire receiver, until absolutely necessary. Even if there will not be an AGC connection to the LNA, or to any other RF amplifier, a delay diode is still a good idea for the first IF gain-controlled amplifier of the IF strip, since this will help in some small way in maintaining a superior system noise figure.
- All AGC-controlled amplifiers should be decoupled from each other by a small value of series resistor (such as 100 Ω), along with a capacitor to ground, at each VGA’s DC gain control port. This will slow any undesired interactions from taking place between all of the gain-controlled stages.

Transistor amplifiers require special circuits or devices for amplifier stage gain control. One circuit arrangement for this application uses one gate of a dual-gate FET as the gain control device while the signal is applied to the second gate.



Dual-gate MOSFET amplifier controlled by AGC

AGC Detectors Types

The AGC detector operates essentially in a linear mode, the DC output being proportional to the RF input voltage.

- Envelope Detector (Rectifier)

The output voltage of the envelope detector is proportional to the magnitude of the instantaneous RF input voltage. Assuming that sufficient low pass filtering is applied at its output to eliminate RF ripple, this detector produces a voltage proportional to the envelope amplitude of the RF signal. Assuming that the loop's bandwidth is made sufficiently small as to avoid significant gain pumping, the effect of the loop using an envelope detector is to stabilize the average rectified voltage of the signal. The resulting power is therefore dependent on the RF signal's envelope waveform.

- Square-law Detector

This type of detector has an instantaneous output which is proportional to the square of the instantaneous RF input voltage, so the output is proportional to input power. This behavior, when incorporated into an AGC loop of sensible bandwidth, makes the loop's equilibrium average output power independent of the input waveform. As with the envelope detector, the output can never go negative, resulting in the loop having a similar tendency towards slew rate limited behavior when reacting to abrupt decreases in input amplitude. The response to large abrupt increases in input amplitude can be even more striking, however, because the square-law detector characteristic exaggerates the effect of the input increase. The extent to which this happens depends on the clipping level of either the VGA or the detector, whichever appears at a lower level.

- True-RMS Detector

This detector comprises a square-law detector followed by a low pass filter followed by a square-root function. The low pass filter performs the "mean" operation associated with the root-mean-square (RMS) function, and it should have a sufficiently long time constant to smooth the output variations of the squaring detector that would otherwise arise from the legitimate modulation of the signal.

Because of the square-root element in this detector, the average output is proportional to the signal voltage, not power, so the loop's response to small abrupt decreases or increases of signal level should essentially be the same as that for an envelope detector, provided that the added filter pole within the RMS detector is correctly compensated for elsewhere in the loop. The fact that the added pole is located in a region of the signal path that is square law brings forth the possibility of the large-step response being different from that of the simple envelope detector, which can indeed be seen in the figure. Note that the RMS detector has a slightly slower recovery from a large downward amplitude step than does the standard envelope detector, but a slightly faster recovery (and a bit of overshoot) from a step up in input amplitude. In common with the square-law detector, the true-RMS detector will make the AGC loop's equilibrium point independent of the RF signal waveform.

It should be noted that the presence of the long-time-constant low pass filter in this detector may have a marked influence on loop dynamics; indeed, this filter may even provide the dominant pole in some designs. This time constant must therefore be coordinated with the remainder of the loop design.

- LOG Detector

This type of detector produces an output proportional to the logarithm of the RF input voltage. Because this behavior is complementary to that of the linear-in-dB VGA in the loop, the resulting loop dynamics are those of a linear system, assuming that signal level fluctuations during transients remain within the measurement range of the log detector. Subject to that assumption, the AGC loop's response to abrupt large changes in input level will not be slew-rate limited, and will often be faster to recover from amplitude decreases.

As with the envelope detector, the equilibrium point of an AGC loop using the log detector will depend on the RF input waveform.

Comparison of Responses with Different Detectors

The results show that the loop's large-step transient response is markedly dependent on the type of detector. At one extreme, the log detector gives the fastest response to large abrupt decreases in input level because the logarithmic curve has a very steep slope for low inputs, which exaggerates the loop's response. However, the log detector has a shallow slope for high input levels, resulting in a diminished response rate to sudden increases in signal level. At the other extreme, the square-law detector's small slope near zero input level gives it a very sluggish response to large decreases in input amplitude. Conversely, the square-law detector exaggerates the response to large signals, giving the fastest response to increasing signals. The envelope and RMS detectors, having intermediate characteristics, give response speeds in between.

Variable Gain Amplifier - VGA

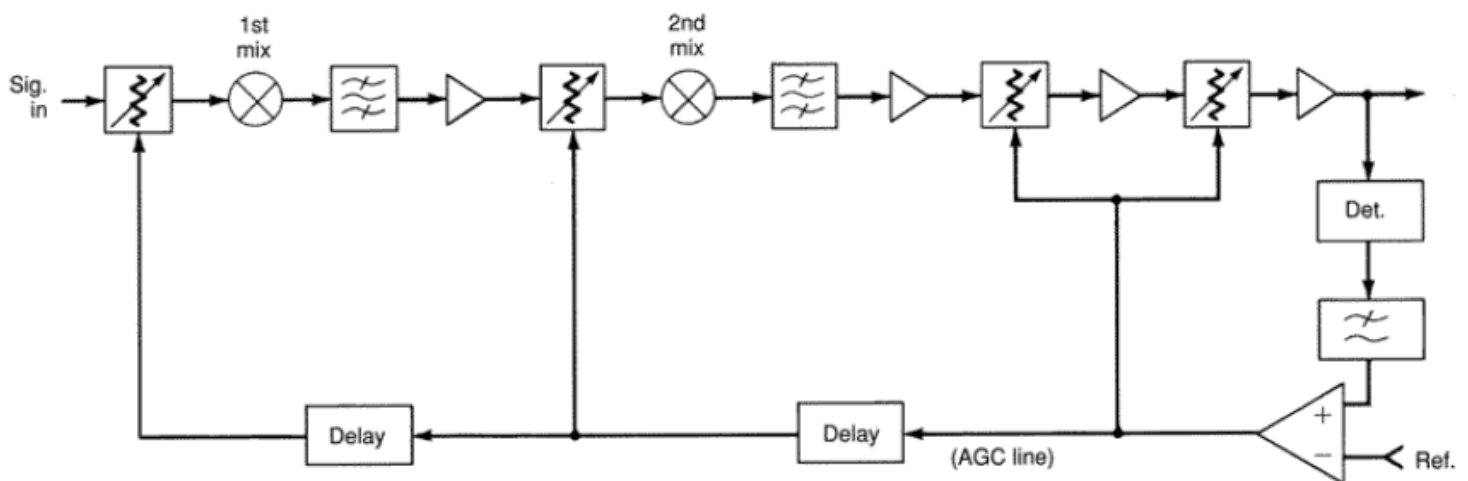
Variable Gain Amplifiers are most often used in feedback configurations as automatic-gain-control (AGC) amplifiers where the amplitude of the output signal is kept constant for all input signal levels.

- The ratio of maximum-to-minimum input signal amplitude that can be handled is called the dynamic range.

Most wide band variable-gain amplifiers can be derived from three high-performance amplifiers which all consist of a quadruple of transistors driven by an input pair. Considerations are restricted to balanced circuit configurations because of the resulting cancellation of even-order distortion.

- Input and output ports are differential, and input and output impedance are independent of amplifier gain.
- In general, the output distortion can be reduced by using lower signal levels. However, the noise generated in the amplifier then becomes important and limits the dynamic range. Distortion also limits the dynamic range of the VGA mainly due to third-order distortion. Ideally if the circuit is balanced, no second-order distortion is present.
- In a multistage VGA to achieve a minimum NF, the gain of the first stage should be as high as possible. In other words, when the signal is large, the gain of the second and third stage should be reduced before the gain of first stages is reduced. However, for large input signals, the noise requirement of the VGA is actually relaxed. When all stages are set to maximum gain, the noise figure of the whole VGA is minimized, however the IP3 is degraded to minimum value as well.
- For a large input signal, the total gain of the VGA will be decreased to a smaller value. Since the signal power is large, the noise contribution from the VGA is not important any more, the gain of all stages can decrease at the same time. The IIP3 of the whole amplifier can also be improved when the gain in the first and second stage decreases.

AGC Receive System using Variable Attenuators



With DC bias-controlled amplifiers care must be taken to confirm that severe distortion does not occur when the gain is varied by the AGC, since the transistor can easily be biased into a nonlinear part of its operation, especially critical if the input signal is of a high amplitude. This issue is not nearly as much of a consideration when AGC amplifiers use voltage variable attenuators at their input, instead of employing AGC bias control, for this gain control function.

In this case the receive system includes voltage-controlled attenuators in the RF, first-IF, and second-IF sections.

A detector at the output develops a DC output voltage proportional to the IF or AF output level.

Intermediate frequency (IF) detection is normally used because it gives less time delay (more charging cycles per second) and has output even when the product detector audio output frequency is zero.

After low-pass filtering, the voltage is compared with a reference such that gain control action commences when the signal input has reached some threshold level.

If all of the attenuation were placed at the receiver's input, the output signal-to-noise ratio would rise to the threshold value and stay there (or rise only slightly) as the signal level is increased.

The delay blocks represent voltage delay circuits that enable the forward-placed IF and RF attenuators at signal levels which are ideally just in advance of succeeding stage overload or IM level specification limits.

As the input attenuator begins to operate with rising signal strength, the output signal-to-noise ratio becomes constant with a value ranging from 30 to 50 dB in typical designs. The ultimate signal-to-noise ratio is usually limited by the synthesizer phase noise when the IF delays are properly placed.

AGC Response Time

The AGC system has a delay in its response to changes in input. This means that the AGC control voltage holds constant for a short time after a change in signal level and then follows the change to compensate for the level change. In practice, it is not desirable for an AGC to have too fast a reaction time. In such a case, any static pulse, ignition noise, or other impulsive interference with very fast rise time would be detected by the AGC detector and would desensitize the receiver for a hold time required to discharge the AGC filter capacitors.

For many years, communications receivers used attack times between 1 and 5 ms. For CW operation and in thunder storms, this proves too fast and hangs up the receiver. The fastest attack time that is possible depends upon the filtering of the detector, the response of the amplifiers, the IF selectivity, and the IF itself.

For SSB reception, some receivers derive the AGC from the audio signal. Rather than use a high-level IF amplifier, such receivers may use a product detector to convert the signal to audio at a level of about 10 mV. The resultant signal is then amplified and rectified to develop the AGC control voltage. If an IF derived AGC is used, the lowest practical frequency is about 30 kHz. If the lowest audio frequency generated were 50 Hz, the audio generated AGC attack time might be extended to 20 ms. Care would need to be taken not to use the audio-derived system for control of an RF carrier (at or near zero beat). On the whole, it appears that a baseband-derived AGC design should be avoided in high-performance receivers.

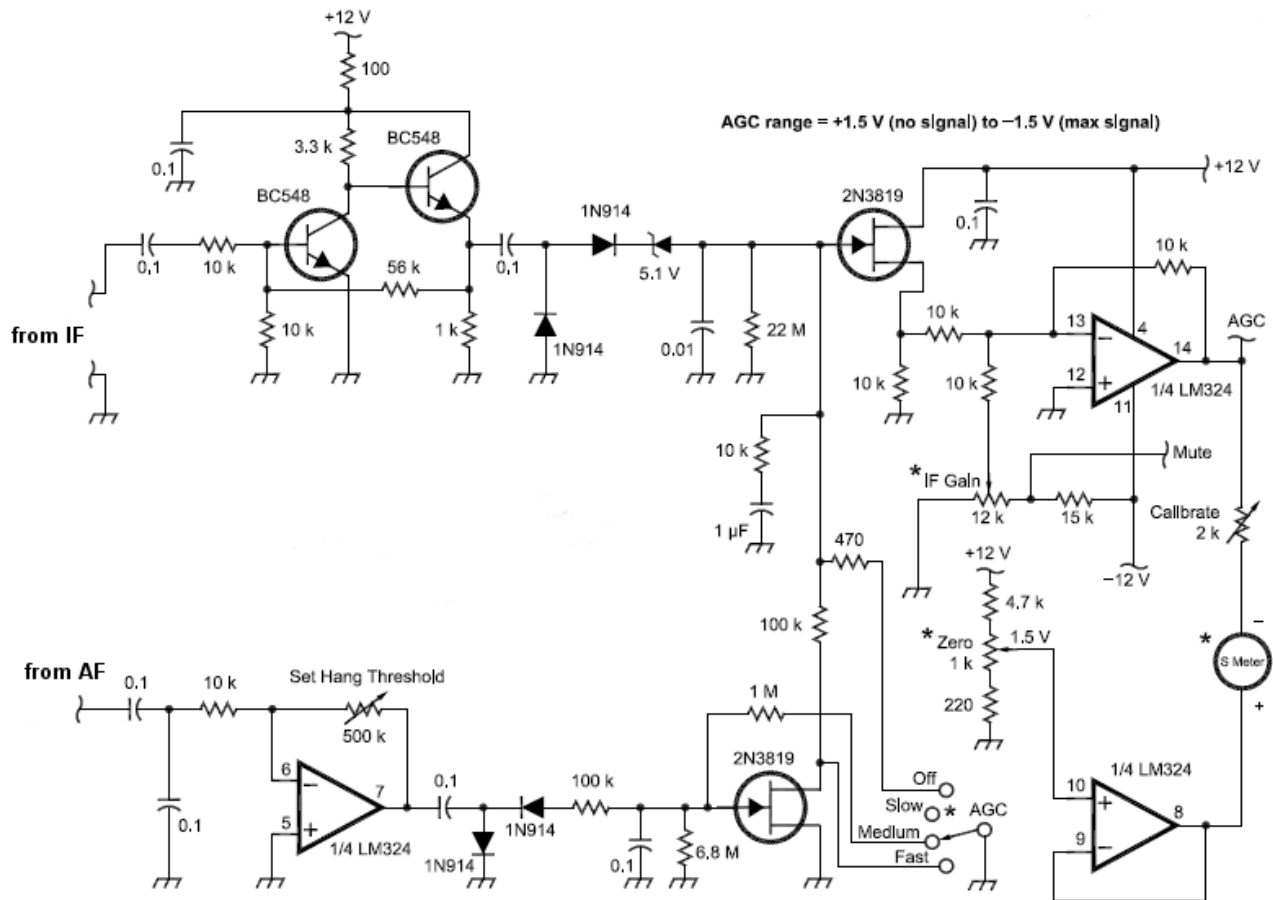
- Attack and Decay times are typically defined as the time it takes to get within a certain percentage of the final value after a signal appears or disappears. It turns out, however, that the loop gain, which determines the loop band-width, is dependent upon the actual gain reduction. Therefore, the attack and decay times should be defined for the highest-gain reduction or maximum input voltage.
- The reasons of AGC instabilities in the highest-gain reduction is because we have to deal with the phase shifts of the various amplifiers, and therefore instabilities can occur.
- In the case of an AM signal, the AGC cannot be made faster than the lowest modulation frequency. In a broadcast receiver, 50 Hz or 20 ms is too small a margin, and a 60- to 100-ms attack time should be preferred. If the AGC time constant is made too fast, the modulation frequency response will be changed and distortion can occur.
- AGC threshold determines the signal level above which the AGC starts to reduce the gain.

A state variable type of AGC loop (e.g. fast, medium, slow) called Hang AGC is very often used to obtain rapid recovery of receiver gain after signal cessation, while also having very low envelope distortion due to audio on the gain control line. The loop filter is designed to charge rapidly to follow the rising signal

input and then to remain at that level after the signal drops. In effect, the receiver gain "hangs" for a preset amount of time.

For voice systems the attack time should be in the 2-ms region, with a hang time of about 0.3 sec, followed by a gradual recovery time of up to 1 sec. Adaptive circuits can be designed that give shorter hang and recovery time as the signal on-time decreases, thus minimizing impulse loading of the AGC system.

- AGC is inactive for a HANG time long period after a signal peak that determined the latest gain reduction.
- When the HANG period expires, AGC starts increasing the gain at AGC DECAY rate.
- Faster rate enables the receiver to recover gain in shorter time, but too fast a rate will reduce readability due to disproportionate increase of noise and interference volume.
- At the end of the hang period, the AGC volts drop to a low level very rapidly, so that the AGC gain is restored quickly. This is to ensure that should the pause be overly long, then the receiver is ready to adjust its gain ready for the next transmission which may be significantly weaker than the previous.



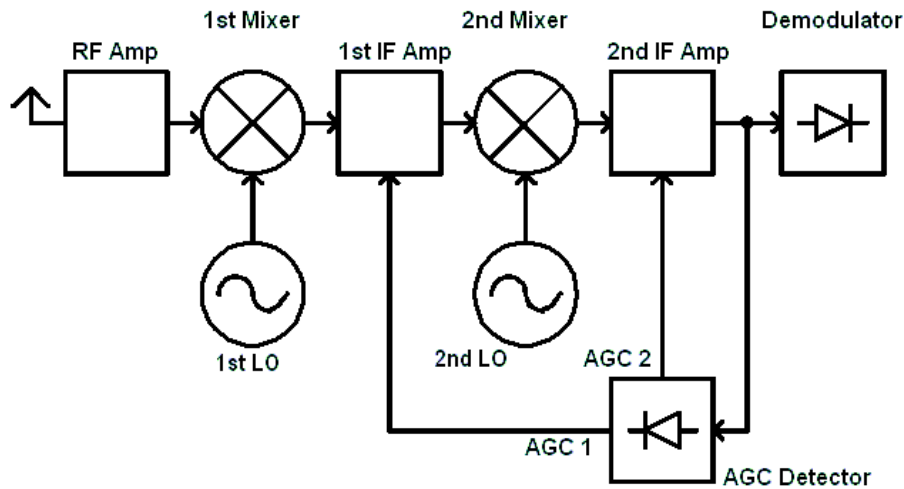
Schematic example of an AGC with Hang Threshold

Effect of Filters Delay

A selective filter introduces not only frequency selectivity but also delay. The filter, especially if it has a flat top and sharp cutoff characteristic (like a Chebyshev filter), may introduce substantial delay, ranging from a few microseconds to as much as 50 ms. Some mechanical filters, resonant at low frequencies, such as 30 kHz, have extremely steep skirts, and also can have delays of 50 to 100 ms. With such a delay, the AGC detector produces a gain-control voltage responding to the signal at a substantially earlier time. If the AGC attack time is smaller than the delay, such delays can cause AGC instabilities. It is important, therefore, to make sure that the AGC attack time is longer than possible delays or else to avoid delays in the system so as to use a short attack time.

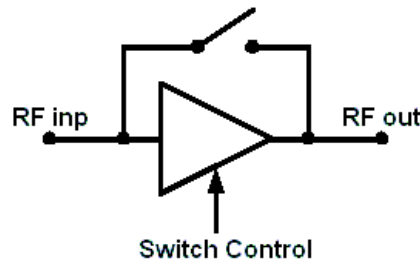
- The delay varies across the filter band, the most critical points being between the -3 and -10 dB points of the selectivity curve. At these points, extreme delays occur and the AGC is most vulnerable.

- To avoid this situation it is possible to use the high-delay crystal filter in the AM or SSB detector path, and use a broader filter with smaller delay in the AGC loop.



Automatic Gain Switch

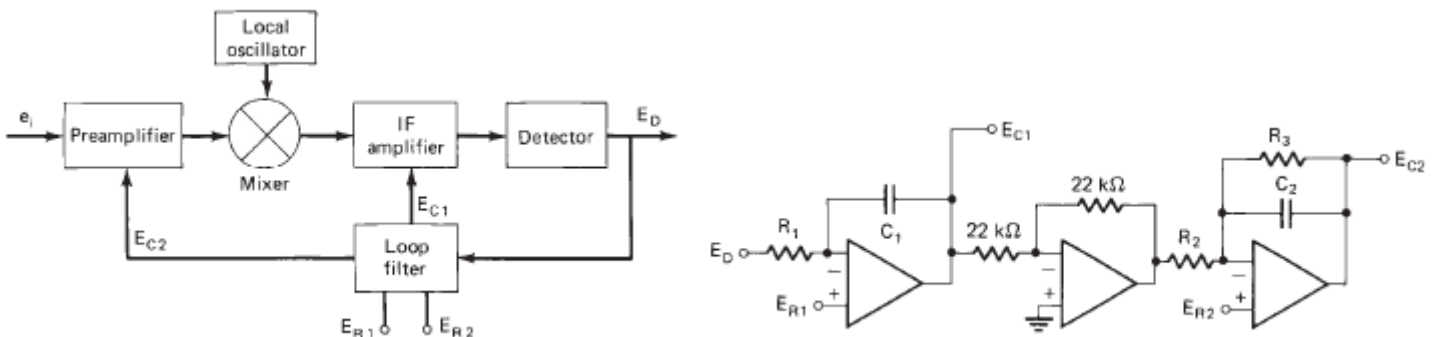
In order to further improve the dynamic range of the input amplifier, some form of automatic gain control is required. A continuous gain variation, controlled by the magnitude of the largest output signals of the amplifier, is not feasible, since the noise and distortion performance of optimally designed variable-gain stages is not consistent with our requirements. A better solution is to switch the gain to a lower value as soon as a certain input level is exceeded. Several methods can be used to switch the gain from one value to another. A single switch, which should have no detrimental effect on the noise and distortion performance, is preferred. Therefore, a solution is to switch the feedback capacitor values.



- An external JFET is a suitable device to operate as a floating switch. The gate of this device can be driven by a logic signal obtained by appropriate detection and comparison of the amplifier output signal.
- A comparator with hysteresis is needed because the input signal for this circuit drops when the switch is activated, and without hysteresis an unstable situation would result. With a gain variation of 12 dB, a hysteresis of 15 dB provides a safe margin of 3 dB where the gain remains low before the switch is deactivated.

Dual-Loop AGC

In applications where NF is very important, high-gain preamplifiers can be used preceding the first mixer. For operation over a wide dynamic range, it is necessary to apply AGC control to the preamplifier to prevent strong input signals from overloading the following stages.

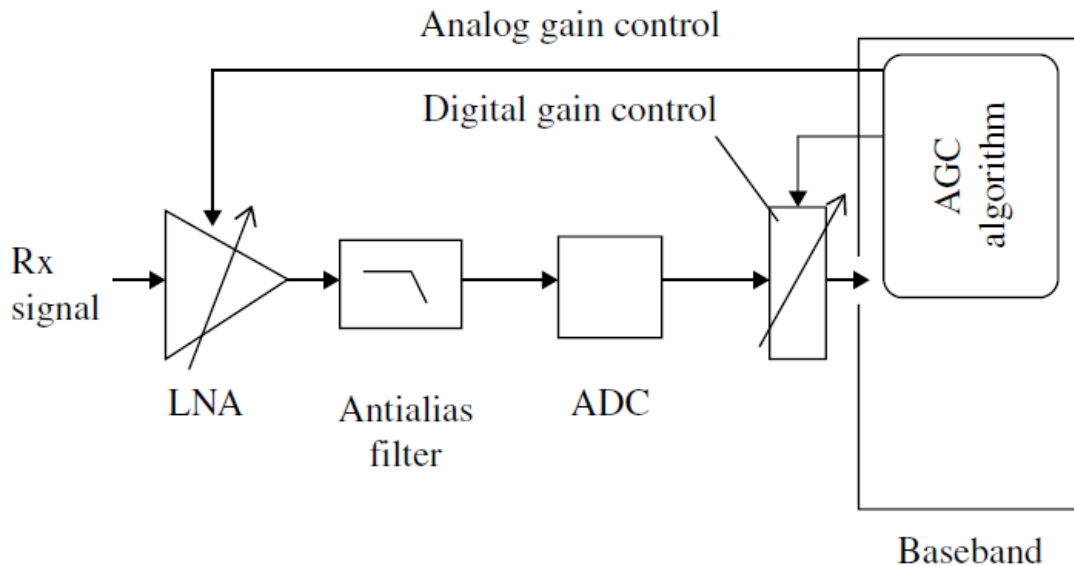


On the other hand, reduction of the preamplifier gain increases the NF of the system. Low NF is the reason for using the preamplifier stage in the first place. To solve these difficulties, it is necessary to delay application of AGC to the preamplifier until the gain in the later stages has been sufficiently reduced by the input level.

AGC and Digital Communications

AGC circuits are widely applied in digital communication systems, mobile phones and base stations, disk drive read channels, and many other digital systems.

- A receiver for data signals can have a very flat output requirement so that the following data detection circuits have a constant-amplitude input.
- All the ADC and DAC circuits has a dynamic range – if the signal strength varies beyond that, then there will be error (saturation) leading to signal clipping.
- Automatic gain control plays an important role in the mobile and base station receivers by adjusting the gain of the amplifier according to the received signal strength. The baseband AGC algorithm monitors the RSSI (received signal strength indicator) and averages that over a period. Then, based on that, it programs the required gain value to the LNA (or/and to the IF amplifier), so that the received signal level is controlled.



Today, in mobile phone receivers, mostly two levels of gain setting are used for the LNA stage. Analog gain is applied by a feedback loop on an amplifier circuit before the signal is converted into digital by ADC, whereas the digital gain is just multiplication by a factor of the digitized value as shown in the picture above.

- The dynamic range is determined by minimum carrier-to-noise ratio and blocking signal (interferer) level at the ADC input.
- Generally, the dynamic range of a receiver is often limited by the dynamic range of the ADC. AGC can be used to dramatically reduce the required dynamic range of the ADC.
- The RF system designers know that adding variable gain in the receiver chain is much less expensive than increasing the dynamic range of ADC.
- The receiver gain is controlled at several stages at the front end (at the IF if is the case) and also at the baseband module through a software algorithm to obtain a better result.

Example: In a typical GSM mobile phone receiver circuit premonitoring is used to maintain a constant output level. This premonitoring is done in three phases and determines the settling time for the RX AGC.

The GSM receiver is switched ON approximately 150 μ s before the burst begins and processor measures the received signal level and then adjusts the AGC-DAC in accordance with the measured signal level and / or switches ON/OFF the LNA with the front-end amplifier control line.

Usually, error free recovery of data from the input signal cannot occur until the AGC circuit has adjusted the amplitude of the incoming signal. Such amplitude acquisition usually occurs during a preamble where known data are transmitted.

The preamble duration must exceed the acquisition or settling time of the AGC loop, but its duration should be minimized for efficient use of the channel bandwidth. If the AGC circuit is designed such that the acquisition time is a function of the input amplitude, then the preamble is forced to be longer in duration than the slowest possible AGC circuit acquisition time. Consequently, to optimize system performance, the AGC loop settling time should be well defined and signal independent.

- In a digital communication receiver strong signals that fall outside the narrowband digital filter bandwidth, but inside the analog IF translator bandwidths, can overload or saturate the A/D converter. This results in the generation of in-band IMD products and can result in significant degradation of the desired signal. If large signal levels are detected at the A/D converter, the receiver gain may have to be re-distributed by reducing the pre-conversion analog gain and increasing the digital gain to maintain the desired signal output level. This will, however, reduce the desired signal-to-quantization noise ratio.

References:

1. *Communications Receivers – U. Rohde, Whitaker, Bateman*
2. *Complete Wireless Design – C. Sayre*
3. *Design and Operation of Automatic Gain Control Loops – Analog Devices*
4. *A discussion on the Automatic Gain Control (AGC) – Phil Harman, VK6APH*
5. *AGC Design – Adam Farson, AB4OJ/VA7OJ*
6. *Mobile Terminal Receiver Design - S.K.Das*