THE DIGITAL AMPLITUDE MODULATOR

Timothy P. Hulick, Ph.D., W9QQ
886 Brandon Lane
Schwenksville, PA 19473
e-mail: dxyiwta@aol.com

Abstract. *This paper is the first of two presenting the Digital Amplitude Modulator and the Digital Linear Amplifier. One leads to the other with the Digital Amplitude Modulator the precursor and the first theorized.*

Linear amplitude modulation in very high power full carrier systems still requires rather large solid state or vacuum tube devices in Medium Wave (MW) and Short Wave (SW), telephone and VHF and UHF television transmission broadcast systems. The analog format for these systems for MW and SW is simple full carrier 100% AM with no phase modulation. For vestigial sideband television, and any of the X-QAMs, X-VSBs or any other digital format requiring both AM and PM to describe the envelope vector of the modulated signal, simultaneous AM and PM are required.

This paper presents a modern approach to digitally synthesized analog amplitude modulation boasting very high power efficiency, and linearity which is only a function of quantization error. This is the first of two papers on Digital Amplitude Modulation and Digital Linear Amplification where all non-orthogonal multiple carrier formats, digital or analog, are transmitted by RF or microwave by means of analog wave-form modulation.

Saturated class C amplifiers are used as binarily related carrier power sources each connected to one input of a set of 3 dB power combiners arranged such that the output of the (n-1)th combiner feeds the other input of the nth combiner. Linear amplitude modulation takes place in the array of power combiners as amplifiers feeding the array are gated on or off according to the bit content of the digitized baseband modulation signal as sampled by means of an analog to digital converter. Modulation is absolutely linear with no flat topping, differential gain or phase or incidental carrier phase modulation. The only source of distortion is the quantization error. The spurious free dynamic range is dictated by the A/D converter used.

The amplifiers need not be linear and may be operated near their theoretical class C efficiency limits. The upper carrier frequency limit is only a function of

the devices used, but generally the technique applies to frequencies at which the power combiners are small.

Introduction. Full carrier amplitude modulation is the oldest form of baseband transmission over a carrier signal by radio. The baseband modulation signal usually takes a CW voltage carrier from an unmodulated static level to near cutoff when the modulating signal is at its highest negative voltage value with respect to the carrier polarity, and to twice its static value when it is at its highest positive value with respect to the carrier polarity. One hundred percent modulation of a carrier in this way results in an instantaneous carrier voltage which swings up to two times the unmodulated level down to zero and back again if the peak to peak voltage of the modulating signal is equal in magnitude to the carrier voltage. It is impossible to properly modulate the carrier to any levels beyond twice carrier voltage level nor down past zero without phase inversion of the carrier and saturation of the active devices.

There are many ways to realize full carrier amplitude modulation. Perhaps the earliest is “high level modulation” of the power output stage operating in saturated class C. This is simply done by inserting the modulation voltage in series with the amplifier output circuit DC supply voltage. Modulator power capability must be equal to half carrier power in order to cause the carrier to instantaneously reach twice its unmodulated value and zero. Since the instantaneous carrier voltage reaches twice the unmodulated value, the peak envelope power of a 100% modulated full carrier system is four times the carrier power. The advantage of high level modulation is that the power RF stage is operated in saturated class C which can have a power efficiency in the eighties of percent.

It is also possible to use “low level modulation” where a low power stage in a transmitter is modulated as if it were the high level stage, but then amplified with linear amplifiers to the desired power level. A high degree of linearity is required after modulation so that the modulation envelope is not distorted on its way to becoming high power.

There are serious disadvantages to traditional high or low level amplitude modulation methods not the least of which is linear amplitude distortion in the modulator itself. Active devices used to amplify amplitude modulated signals have nonlinear transfer functions which may result in unacceptable differential gain and phase performance, compression at the top, cross-over distortion near cut-off, incidental carrier phase modulation and cross modulation all resulting in in-band or odd-order intermodulation products. Intermodulation products which are in-band are added signal components which were not present in the original set of modulating signals. Various methods and circuits have been devised to minimize the intermodulation products by pre-correcting or pre-distorting the low level modulated stage

in a way that counters the distortion caused by follow-on linear amplifiers with a theoretical net affect of zero distortion, but success is usually limited to about10dB of correction. Clearly, in this age of wireless communication demanding a much higher degree of linearity than ever before, other methods of providing high level linear amplification must be sought...and are possible! The purpose of this presentation is to introduce a very different approach to very linear amplitude modulation leading to a revolutionary approach to power linear amplification in the paper entitled, “The Digital Linear Amplifier.”

The 3 dB Coupler.

An ideal coaxial transmission line is used to move energy from one point to another in a lossless way. The coaxial line consists of an inner conductor and a surrounding outer conductor, usually grounded for an “unbalanced” transmission line. The ratio of the inside diameter of the outer conductor to the outside diameter of the inner conductor along with the dielectric constant of the material between the conductors determines the characteristic impedance of the line. If there is no leakage between the conductors and if the resistance loss of the conductors is insignificant, the characteristic impedance of the line is *real* with no *imaginary* component. The characteristic impedance of the line is that which will cause no reflections from a traveling wave to either end of the line if the end to which the wave is traveling is terminated in the same impedance. This is logical since a termination of the same impedance simply simulates an infinite length of more line. A line that is infinitely long can never create a reflection because the traveling wave never arrives at the end. That being stated, suppose that a second inner conductor, similar to the center conductor is inserted between the two conductors and parallel to them such that it is sharing the same electromagnetic field. A traveling wave between the original conductors will induce a wave in the inserted conductor so that if it is terminated in a resistive load, power will be coupled to the load. Properly dimensioned, it is possible to space the distance between the two inner conductors while they are $\lambda/4$ in length such that half of the power from the main center conductor couples to the inserted conductor and remains phase coherent with it. Since the coupler is a reciprocal device, (there is no reason to think that it isn’t), coherent signal power going into both conductors will sum at the end of one of them. This form is shown in Figure 1. In the former case, the line is a signal splitter resulting in two coherent signals of equal power. In the latter case, the line is a signal combiner. The 3 dB hybrid used as a combiner will be shown to be the heart of the theory and practice of the modulator in the Digital Amplitude Modulator.

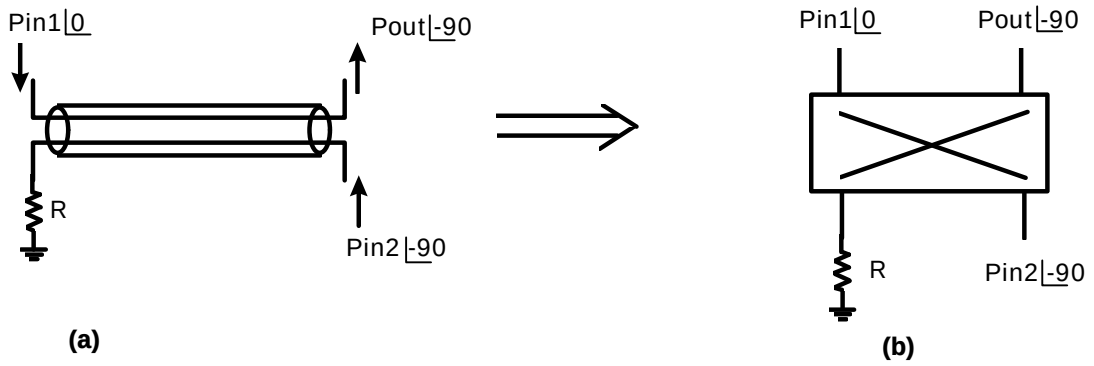

Fig.1. For a 3 dB coupler, when Pin1 has the same value as Pin2 and the voltage vectors are 90 degrees out of phase, as shown, all power adds to Pout and no power is lost in R. Schematically the hybrid coupler is shown in (b).

Referring to Figure 1, A wave of power P_{in1}, traveling from left to right enters the combiner at a relative phase of zero degrees (arbitrarily chosen to be the phase reference) and undergoes a change to -90° through the quarter wavelength path distance to P_{out}. A second wave of *equal power* already set to -90° entering at P_{in2} will add completely with P_{in1} at P_{out} since it is coherent with P_{in1} at the same cross section point along the line. Power inputs to this four port device do not have to be equal in magnitude, however, nor do they have to be 90° out of phase with each other, but in the derivation of the digital amplitude modulator, phase will always be assumed to be of the proper directional shift of a constant 90° to cause power addition as a function only of the relative amplitudes of the two signals being summed. In the general case:

$$P_{out} = \left(\sqrt{\frac{P_{in1}}{2}} + \sqrt{\frac{P_{in2}}{2}} \right)^2 \qquad (1)$$

and

$$P_R = \left(\sqrt{\frac{P_{in1}}{2}} - \sqrt{\frac{P_{in2}}{2}} \right)^2 \qquad (2)$$

For example, let $P_{in1} = P_{in2} = 2$. It is easy to show that (1) gives a P_{out} of 4 and by (2), the power dissipated in the resistor, P_R , sometimes called the reject load, is 0.

Now let $P_{in1} = 2$ and $P_{in2} = 1$. Using (1) shows that $P_{out} = 2.914$ and by (2), $P_R = 0.$. When the two input power levels are equal, none is lost to the

according to (1) and (2) so that the output power is not the simple sum of the input power levels. It is precisely this power summing behavior that

A Cascade of Combiners. Shown in Figure 2 is a cascade of combiners

$_1$) is

connected to one of the inputs of the next. The output of the second is

one of the inputs of a fourth, and so on. For this example, only four combiners are used. $_0$ represents the "carry power" from the continuance of an infinite number of imaginary combiners to the left.

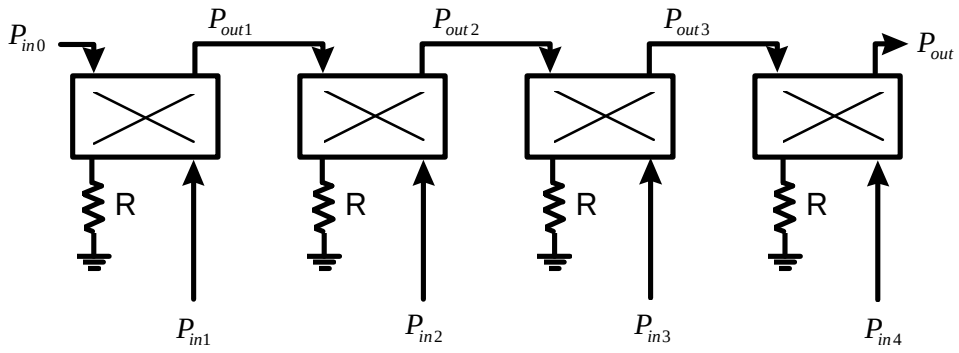

Fig.2. Four 3 dB power combiners are arranged so that the output of one feeds the next. P_{in0} is considered to be from an infinite number of combiners to the left. All of the $P_{out1,4}$ follow (1) while all reject load powers follow (2).

are equal, the output power is twice either input level. Inotherwords, an output power is double that of one of its inputs. Following the cascade of $P_{in2} = 2P_{in1}$. In general, let $P_{in(n)} = 2P_{in(n-1)}$.

For a real example, arbitrarily let $P_{in0} = 1$. Let $P_{in1} = 1$. Let $P_{in2} = 2$, let $P_{in3} = 4$ and let $P_{in4} = 8$ wattage units. Figure 3 shows the same combiner array as in Figure 2, but with real relative watts indicated.

No power is wasted in the reject loads because both inputs to each combiner are always equal in power level and of the proper 90° phase shift.

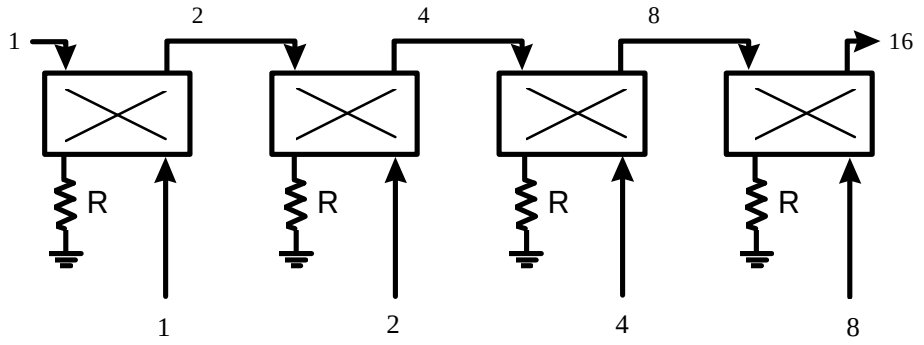

Fig.3. Four cascaded combiners are fed with power sources which are double in value as the signal progression moves from left to right. It becomes obvious that the power sources are binarily related in level in order for all power to add to the output without wasting any in the reject loads.

The “carry power” coming from the imaginary infinite array of combiners to the left is represented by 1 watt in the example shown in Figure 3. There is no control over it since the carry power as a dynamic source doesn’t exist. It is always either present or absent, 1 watt or 0 or any other constant value that makes sense. It is arbitrarily set to $P_{in0} = 0$.

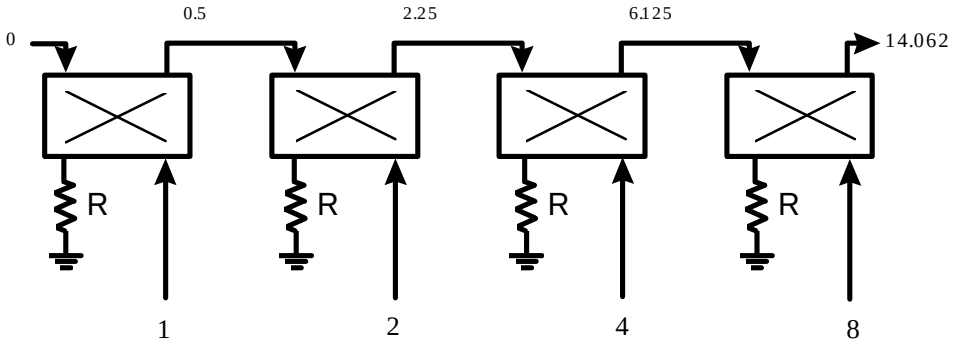

Fig.4. With the “carry power” from the left fixed at 0, the power lost in all reject loads is 15.000-14.062=0.08 which is only 6.25%.

The 1,2,4 and 8 watt sources need not all be present at the same time, but when they are, theoretically no power is lost to the reject loads. Some of the power sources may be selectively shut off. It is possible to calculate the power that will be at the output port of the cascade. In Figure 5, the first and the third power sources are set to zero.

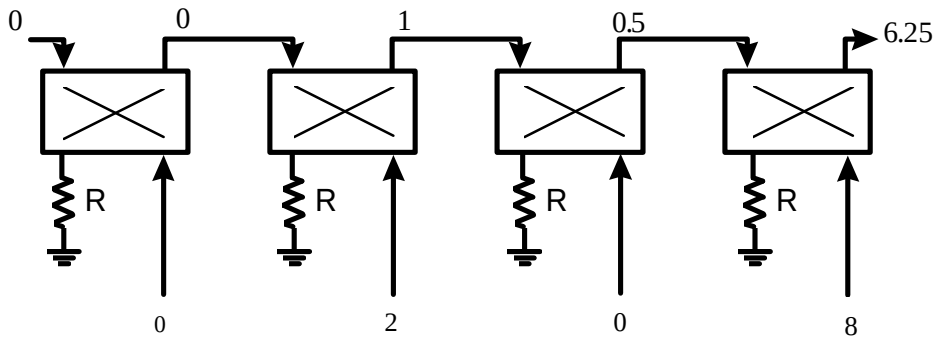

Fig.5. With the first and third power sources shut off, using (1) the power at the output becomes 6.25. Since the total input power is 2+8=10, the power lost to the reject loads is 3.75 relative watts.

It is expected that some power will be lost in the reject loads since power levels into an individual combiner are no longer equal. The power at the output of the cascade has dropped dramatically from the 14.062 relative watts in Figure 4. In fact, the maximum that the output power can ever be is 14.062 when all sources are ON and the lowest is zero when all sources are OFF.

The Digital Amplitude Modulator. From the least significant power position (least significant bit or LSB) to the left in Figure 5 to the highest significant power position (most significant bit or MSB) to the right, a parallel binary word may be derived to be 0101. (To be true to the proper order of writing numbers, the MSB should be to the left so that 0101 becomes 1010, but this is only convention and the order is chosen to remain 0101 since it conforms to the flow of power addition in the figure). This means that the 1 watt source is OFF, the 2 watt source is ON, the 4 watt source is OFF and the 8 watt source is ON.

Any of 16 possible state combinations between zero output power and the maximum of 14.062 watts may be selected according to a parallel four-bit binary word between 0000 and 1111.

Sampling a bandlimited modulation signal such as voice or video is possible with today's analog to digital converters. They are available from many sources with sampling rates at 100 MBS for ten bit converters becoming commonplace. Shannon's sampling theory requires that the sampling rate be at least twice that of the highest frequency component sampled. Staying above this Nyquist criteria ensures that no sampled spectrum folds back into the baseband signal spectrum causing distortion by "aliasing." For television video with a baseband width of 4.18 MHz, a sampling frequency may easily be chosen and implemented which is three

to four times greater than this. For the moment it is sufficient to understand that the baseband signal must be bandlimited so that the minimum sampling frequency may be determined.

Assume an arbitrary baseband signal which is bandlimited to some finite frequency. Knowing this, it must be periodic, but nevertheless, arbitrary such as that shown in Figure 6. The highest value for the waveform is assumed to be 16 volts and the lowest is 0. The signal is fed into an analog to digital converter capable of sampling such a waveform such that when the analog signal reaches its peak of 16 volts, all output bits at the parallel output are logic ones. When the modulating signal is at its lowest value, zero, all output bits are logic zeros. This is a voltage waveform with all other in between values properly represented by the correct binary word as determined by the analog to digital converter. For example, if, at the moment of sampling, the analog waveform is at an 8 volt point, the digital word at the output of the A-D converter is 0001 (in keeping with the chosen convention). If it is at a 5 volt point, then the bits change to 1010. This straight binary representation of the analog sampled point may be used to control an RF drive switch to each corresponding power source feeding the combiner array. If a bit is a logic 1, the drive switch is closed. If it is a logic zero, it is open. The system is shown in Figure 7.

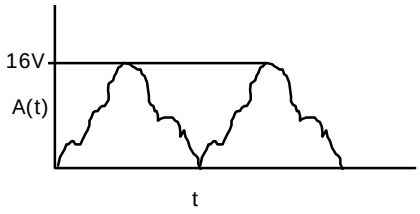

Fig. 6. An arbitrary bandimited modulation waveform, I.e., it is periodic with maximum possible amplitude of 16 voltage units.

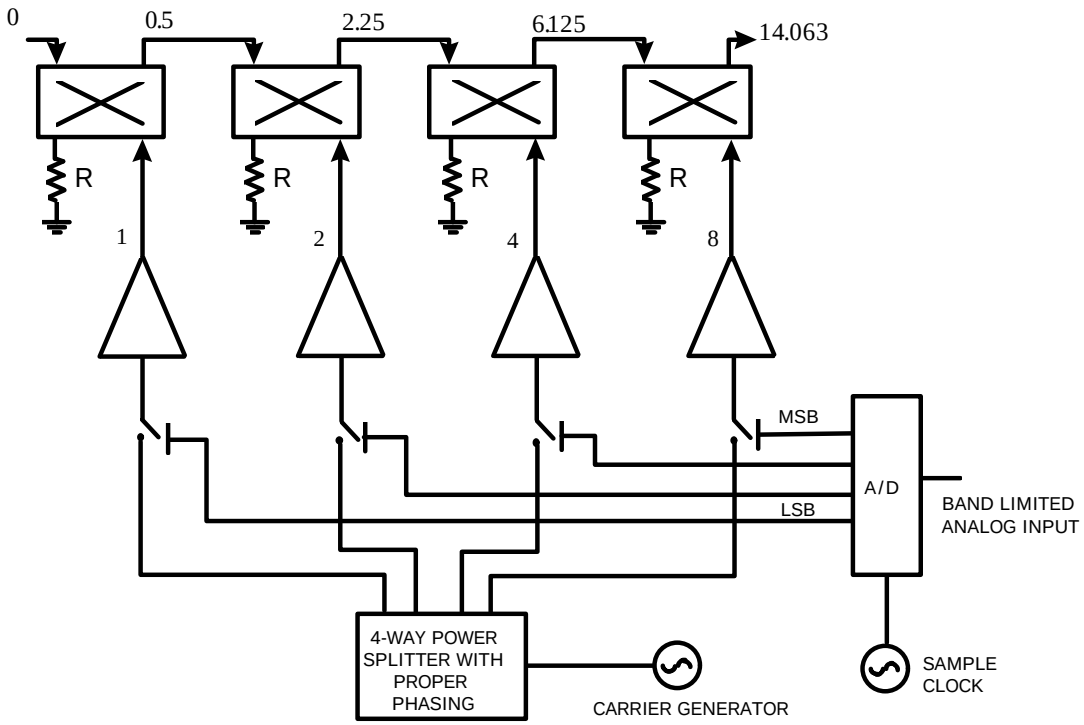

Fig. 7. The RF drives of four binarily related carrier power amplifiers are controlled to be ON or OFF by the logic state of the sampled digital baseband equivalent of the analog modulating signal in this four bit Digital Amplitude Modulator.

It has been shown that when the modulating signal is at its maximum value, i.e 16 volts in this example, that all bits in the four bit word are logic one (1111) which makes the power output of the combiner array equal to 14.063 watts. It has also been shown that when the modulating signal is at its minimum, that all bits in the four bit binary word are logic zero (0000) which makes the power output of the combiner array equal to 0.00 watt. It should be obvious that when the word is a mix of ones and zeros that the power output will be something in between these extremes. The remaining issue is "*whether or not there is a linear relationship between the sampled baseband binary value and the RF voltage output derived as the square-root of the output power.*"

By example, a voltage value of a sampled point with a value of 3 volts is assigned. The binary representation of 3 is 1100. This means that drive to the LSB amplifier and the next LSB amplifier are gated ON. The other two are OFF. This situation is shown in Figure 8.

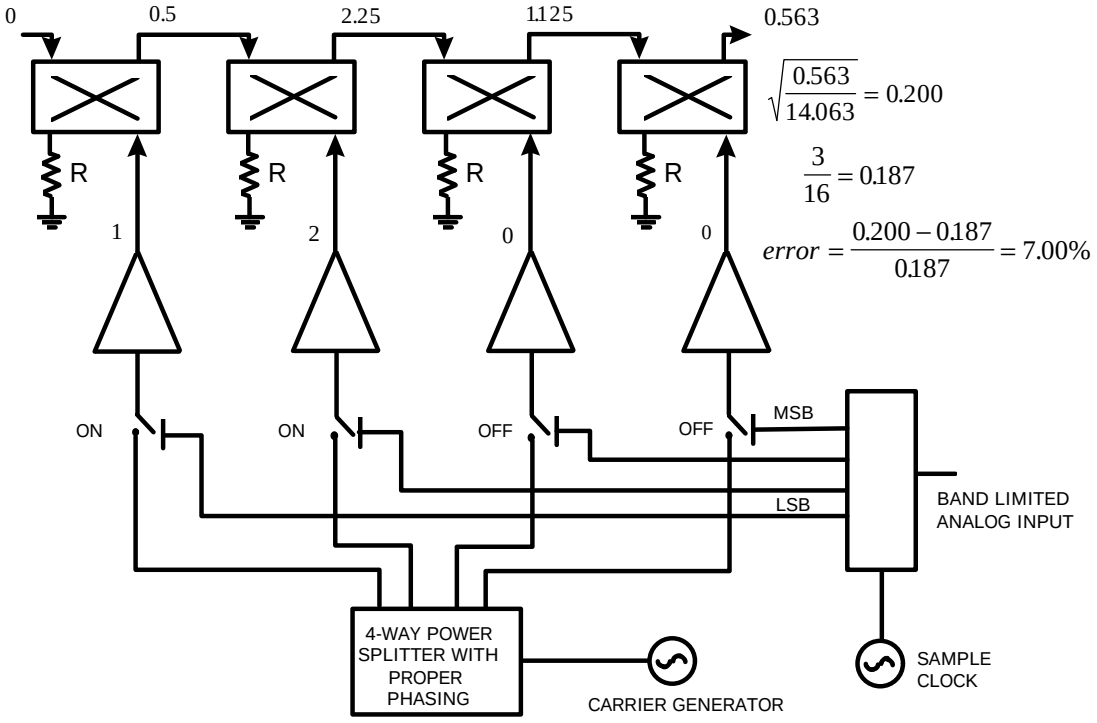

Fig. 8. For the sampled signal at 3 volts out of a possible 16, the analog input voltage is 3/16 or 0.187 of its maximum. The square root of the corresponding output power ratioed with the maximum possible, 16 watts, provides 0.200 of its maximum possible output envelope voltage. The error is 7.00% which agrees with that expected of a four bit system.

In a four bit system, the maximum error is expected to be the quantization error of the least significant bit or 1/16=6.25%. 7.00% is calculated in Figure 8 because of the arbitrary choice for the carry value. A carry value of 0.5 or some other value will minimize the error. The point to be made is that as the number of bits increases, the value of the error decreases. For, say, a ten bit system, the error will decrease to something like $1/2^{10}$=0.1% which leads us to conclude that the choice for the carry value becomes insignificant as the number of bits increases to a value which is useful and pure for the application. For this simple example, linearity seems to be preserved, at least for the sampled modulation wave form amplitude of 3 volts.

To further illustrate that modulation is linear, consider another example where the sampled modulating waveform is at 9/16ths of its maximum value of 16, or, 9 volts. This corresponds to a binary word of 1001. Figure 9 shows the condition of the drive switches with a carry power again set at zero.

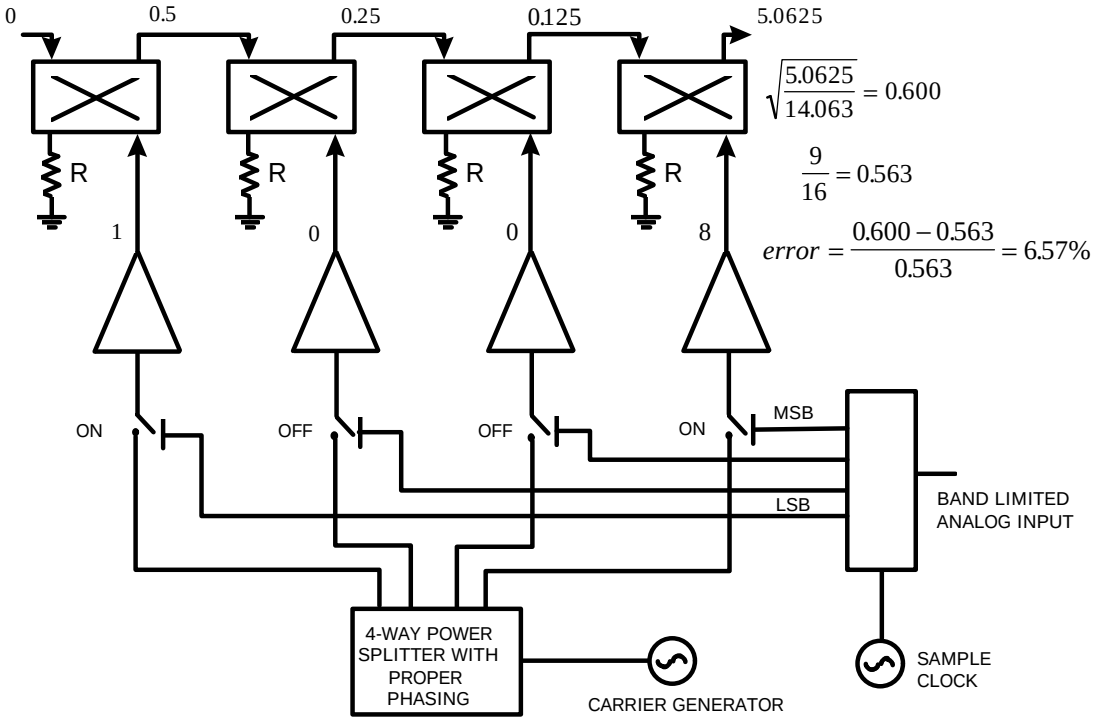

Fig. 9. For a sampled baseband waveform at 9/16th of its maximum, or 9 volts in the example, the four bit output power is calculated to be 5.062 watts out of a possible 14.063 maximum for an equivalent voltage ratio of 0.600. This corresponds very closely with the modulation waveform at 9/16ths of its peak at 0.5625 with an error of 6.57%.

For the modulating wave-form value of 9, the error in the RF output level is 6.57%, again, within the predicted value for a four bit system. It is shown that at least four out of sixteen possible modulating waveform values (0,3,9 and 16) that the square root of the output power which is the voltage magnitude of the modulated carrier follows the relative magnitude of the modulation voltage wave-form, at least crudely for a four bit array. Imagine the accuracy or linear integrity if four bits were raised to, say, ten bits which is well within the state of the art for 100 MHz sampling. In fact, linear distortion will be maintained to -6 dB per bit. For a 10 bit system, this turns out to be -60 dB!

The system (four bits) presented by example in Figure 9 is an oversimplification used only for illustration. A four bit system simply has too much quantization error to be useful, but a ten bit system is entirely different for artifact free television viewing and medical imaging offering 1024 amplitude levels instead of only 16. This may be verified by searching the websites of manufacturers of analog-to-digital converters for bit size and application. They include Harris (now Intersil), Analog Devices, Maxim and Burr-Brown.

To this point in the description of the Digital Amplitude Modulator, it should be obvious that except for losses in the power combiner array that power efficiency may be that of class C operation since the power sources do not pass a modulated signal and, therefore, need not be linear. Also no power is lost in the reject loads when the modulator is operating at the highest peak envelope power (all logic ones). Power is lost to the reject loads for all lesser power levels, but this only occurs when less power is required and power loss (efficiency) is a lesser issue.

Since power is necessarily lost to the combiner array reject loads for this modulator to be linear, a table of combiner power efficiency versus binary word value is presented in Table 1. The maximum output power availability of 14.063 watts is rounded off to 16 so that the numbers are binary friendly.

Binary Word (LSB to MSB)	Combiner Efficiency
0000	1/16
0001	2/16
0010	3/16
0011	4/16
0100	5/16
0101	6/16
0110	7/16
0111	8/16
1000	9/16
1001	10/16
1010	11/16
1011	12/16
1100	13/16
1101	14/16
1110	15/16
1111	16/16

Table 1 . Power Combiner Array power efficiency as a function of digital word value is presented in tabular form. It is necessary that power is wasted in the reject loads for the modulator to be linear, however, power lost goes to zero as peak envelope powe rgoes to the maximum. When power output is low, efficiency is low, but for low output power, power efficiency becomes an insignificant issue.

From Table 1 it is shown that power efficiency is the highest when the highest power is required. When envelope power is low at the bottom of a modulation wave-form, efficiency suffers, but it is not of any consequence at this level since so little relative power is being made.

The Real Ten Bit Digital Amplitude Modulator. As painful as it may seem, it is worth going through the numbers of a ten bit version of the four bit example Digital Amplitude Modulator simply to see where the linearity falls and the power efficiency is gained. Figure 10 shows the configuration of the ten bit Digital Amplitude Modulator with the binary representation of the 16 volt maximum sampled baseband signal set at, say, 11.375 volts. This

corresponds to the (11.375/16) X 1024 = 728th binary level out of a possible 1024 levels. This is equivalent to a ten bit binary word from LSB to MSB of 0001101101 which is the representation of the number 728 in a 1024 possible state system.

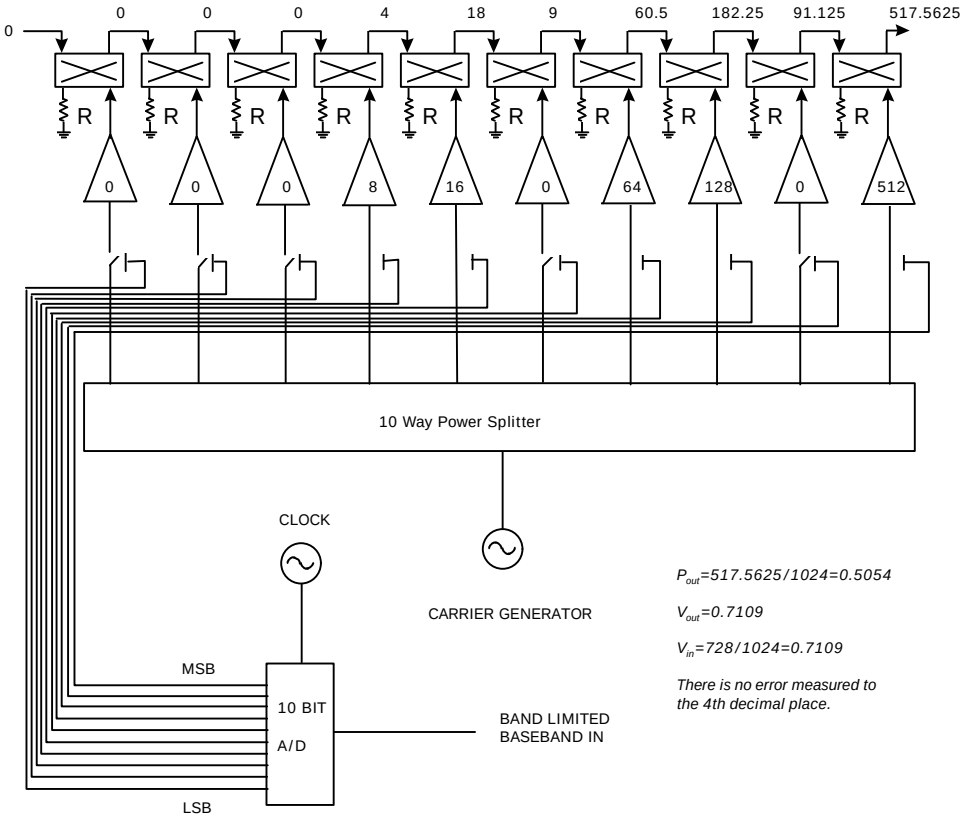

Fig.10 For a ten bit system, there is no measurable error to the fourth decimal place

Figure 10 shows that for a ten bit system, where there are 1024 discrete amplitude steps, the significance in the choice of the carry value from the left is insignificant whether it represents a logic one, zero or any value in between.

The straight binary representation of the sampled analog baseband modulation signal controls the presence or absence of relative numerically equivalent carrier power levels for digitally perfect amplitude modulation in the cascade of power combiners.

Review and Summary. It has been shown that full carrier linear amplitude modulation of an RF carrier takes place in a cascade of 3 dB power combiners when each combiner is fed with the correct binarily related power level relative to all the others. It is further required that drive to these amplifiers be gated on and off according to the straight binary

representation of the sampled analog modulating signal. Since the system is direct coupled from analog input to RF output, the carrier level may be set by a DC offset in the A/D converter such as in the case of broadcast television where the carrier is 25% down in voltage from the peak of sync value. Carrier level and modulation wave-form are arbitrary, but the modulation level cannot exceed 100%. The only source of distortion is the quantization error since there is no compression or crossover distortion. Certainly, quantization error may be made arbitrarily small by the choice of digital word length in keeping with the Nyquist criteria and the state-of-the-art.

Practical Considerations. Although the system diagram shown in Figure 10 is quite simple in concept, there are practical considerations which could complicate the design of a high power AM transmitter using this technique. For example, when using a high number of bits for the digital representation of the modulating signal, it is necessary to make sure that all RF levels are always at their correct values to within half the power of the least significant bit. In a ten bit system operating at, say, 1024 watts PEP, each bit level must hold its RF carrier value to within ½ watt! Ways to maintain binary power levels and other ways to simplify the concept shown in figure 10 to make a practical design are presented in a follow-on paper entitled, “The Digital Linear Amplifier.” The second paper presents the digital approach to a much more generalized application, the amplifier instead of the modulator. These simplification procedures apply to the Digital Linear Amplitude Modulator and the Digital Linear Amplifier.