Concepts of Amplifier Efficiency

Earl McCune, CTO
Eridan Communications

Power is expensive to get, and it also is expensive to get rid of. Both are necessary in wireless transmitters. To minimize these costs we become inherently very interested in *Energy Efficiency in RF Transmitters*. This is the entire reason there is any interest at all in new transmitter designs. So what does this mean?

Whenever an amplifier provides signal power into a load resistance R_L, Ohm's Law says that the corresponding root-mean-square (rms) current that the signal has is related to the rms output power (P_{OUT}) according to

$$I_{Signal} = \sqrt{\frac{2P_{OUT}}{R_L}} \tag{1}$$

The current from (1) must flow through the load no matter what the amplifier structure is. So whatever the supply voltage is, this current must flow from it, as shown in Fig. 1.

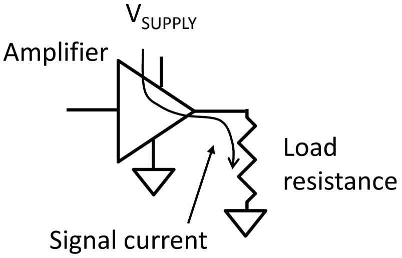


Fig. 1 Any transmitter must generate a signal current in the load resistance. This current must flow from the amplifier power supply, no matter how it gets modified through the amplifier, by the amplifier's DC to RF conversion process.

Power is dissipated (lost) when this current flows through any resistance, which particularly includes the amplifier's power transistor. How signal and supply power are related in any amplifier is shown with the flow graphic in Fig. 2.

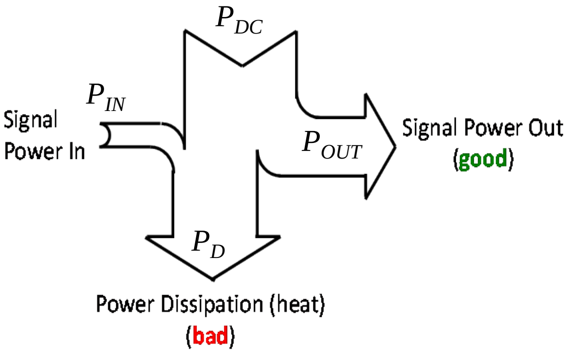


Fig. 2 Power flows in to an amplifier from the supply (P_{DC}), and from the input signal (P_{IN}). The desired signal output power (P_{OUT}) flows out. All 'leftover' power stays in the amplifier (P_D) and becomes heat. This heat must flow out of the amplifier and into a heatsink.

Ideally, we desire that the supply input power is all converted into the intended output power. Physics, however, forbids this to happen. We are left with the concept of amplifier efficiency, here more specifically called *energy efficiency* for good technical reasons. The relationship between these power flows that defines this energy efficiency is

$$\eta \equiv \frac{P_{OUT}}{P_{DC}+P_{IN}} \simeq 1-\frac{P_D}{P_{DC}} \quad \text{for small } P_{IN} \tag{2}$$

Plots from (2) where the supply power and dissipated power are evaluated across PA energy efficiency values are provided in Fig. 3. It appears to be generally accepted that the cost increase for a larger power supply, and for the heatsink, is acceptable for energy efficiency exceeding 40%. At lower efficiency values, the size and cost of both the power supply and amplifier heatsink increase, with no increase in useful output power for the desired communication. This is not good.

Yet, this is exactly the situation present today in most communication networks. Using linear amplifiers, for the cellular system operating with 3G signals the power amplifier (PA) energy efficiency is around 35%. For LTE operation the PA energy efficiency drops below 10%. Modern wireless local-area networks (WLAN) operate closer to 7%. And millimeter-wave (mmW) transmitters using these modern signals operate below 2% energy efficiency – way off the chart in Fig. 3. These low energy efficiencies are completely predictable, and are direct consequences of the signal modulations adopted for these standards.

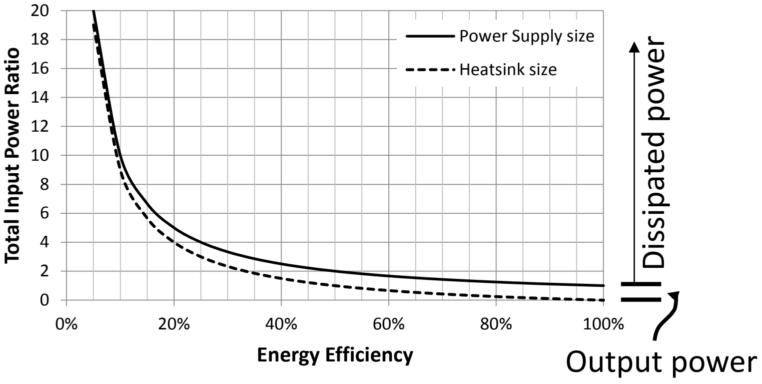


Fig. 3 Comparing supply input power and dissipated power output shows that both increase dramatically when the PA energy efficiency gets low. This

Being part of the adopted standards, these signal modulations are going to stay. But we also insist that their associated transmitters be energy efficient. The only solution is to accept, and tolerate, amplifier circuit operation that is either not linear, or is time varying, or both. This has been understood since 1915. And throughout the past century this knowledge has been followed with varying amounts of vigor [1].

The key is "simple", as shown in (2). To improve efficiency, "all" you need to do is to reduce the amplifier power dissipation. To increase this challenge, the laws of physics do not allow power dissipation to be reduced arbitrarily from circuit linearity. This is where the tradeoff comes from between good circuit linearity or high energy efficiency. Maximizing one necessarily minimizes the other. Getting into the middle ground has motivated many developments in transmitter designs, including the Doherty architecture [2], the outphasing structures of Chierex [3] and many others, the envelope-elimination and restoration (EER) architecture of Kahn [4], and polar modulation from Heising [5]. More recently another foray into this middle ground called envelope tracking (ET) [6] is getting research interest. These last two both fall under the heading of dynamic power supply transmitters. Eridan uses these techniques

PA dissipated power is the product of the current in the load times the voltage difference between the supply voltage to the amplifier and the output signal voltage. When the voltage supplied to the amplifier is a constant value, and by far the most common design practice, the situation in Fig. 2a results. Power dissipation in the amplifier is maximum when the output signal voltage is ½ of the supply voltage. When the output signal voltage is higher, even though the current value is larger the voltage drop is less and the power dissipation is lower. Similarly, when the output signal voltage is small, even though the voltage drop is now large the current in the load is smaller and again the power dissipation is lower.

To keep this internal power loss in the amplifier small, the voltage drop must remain small at all times because the signal current cannot change for the same output power. Various techniques to achieve this goal are shown in the progression of charts in Fig. 1-2. When the signal peak voltage is known to not go all the way to the available supply voltage, the actual supply voltage to the amplifier can be reduced as shown in Fig.

1-2b. This is called the average power tracking (APT) technique. The voltage drop inside the amplifier is reduced further when the supply voltage to the PA (the dashed line) now varies along with the output signal envelope. This situation is shown in Fig. 1-2c. How closely the dynamic power supply follows the signal envelope, the *offset voltage*, has a *huge* impact on the operating properties of the amplifier and the transmitter it is in. Details set by the specific application this transmitter supports dictate design limits on this offset voltage based on the required properties the transmitter must have. In general, this offset voltage must be very carefully controlled.

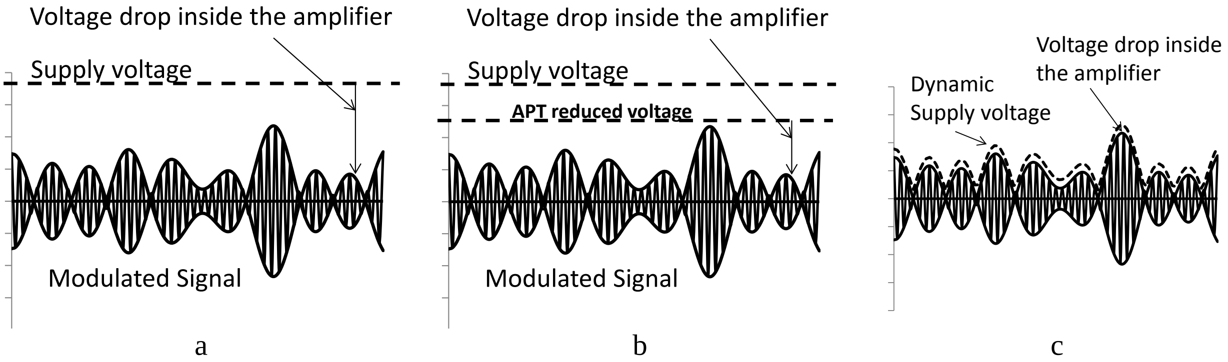


Fig. 4   Power loss in an amplifier depends on the voltage drop inside the amplifier: a) voltage drop across the PA transistor with a fixed supply; b) voltage drop inside the amplifier with a reduced supply voltage from APT operation; c) voltage drop across the PA transistor with a dynamic supply that follows the signal envelope

By adopting a dynamic supply voltage there are two new problems that need to be solved. One is how to get the waveform needed for the tracking supply, which in modern times is most often directly calculated by the digital signal processor that is calculating the signal itself. Once this waveform exists, it must be properly aligned in time with the signal when they both come together at the final power amplifier.

It is very important to separate the concepts of instantaneous voltage drop in the amplifier from the peak value of the power supply available to the amplifier. Ohm's Law says that while keeping the instantaneous voltage drop small to gain high efficiency, it is equally important to operate any real amplifier from the highest practical voltage available to further improve its efficiency. Radio transmitters are power-based signal processors, and like the electric utility, the transmission efficiency is greatest when starting from a high voltage. This is the opposite characteristic of the energy-based signal processing performed by digital circuitry (particularly CMOS) where energy efficiency is improved at lower voltages. These two circuit types optimize at opposite ends of the voltage scale.

Linearity and Linearization

In more modern times wireless signals are used to communicate a lot of information. These signals are increasingly intolerant of distortion in the transmitter circuitry (and anywhere else, for that matter). This increases the demands on linear performance of the amplifier circuitry. Nothing comes for free – improve the circuit linearity and the power dissipation also increases. As mentioned before, any attempt to reduce the single transistor power dissipation for the same output signal power necessarily results in reduced linearity from the circuit.

Customers usually don't care about the laws of physics. They insist on having good linearity and high energy efficiency. To meet this demand we in the product engineering community need to more carefully understand the actual need, which is *output signal accuracy*. Circuit linearity is not necessary to get output signal accuracy, but it is harder to do if circuit linearity is not available. The term used for getting output signal accuracy without depending solely on circuit linearity is called *linearization*.

Fortunately it is possible to achieve very accurate output signal properties in the complete absence of circuit linearity. This does appear to solve the linearity / efficiency trade-off, but as always there is a cost that may or may not be acceptable. Mainly the costs here are 1) higher complexity in the necessary implementation along with 2) a compatibility issue with several of the high data rate signals used in present communications systems.

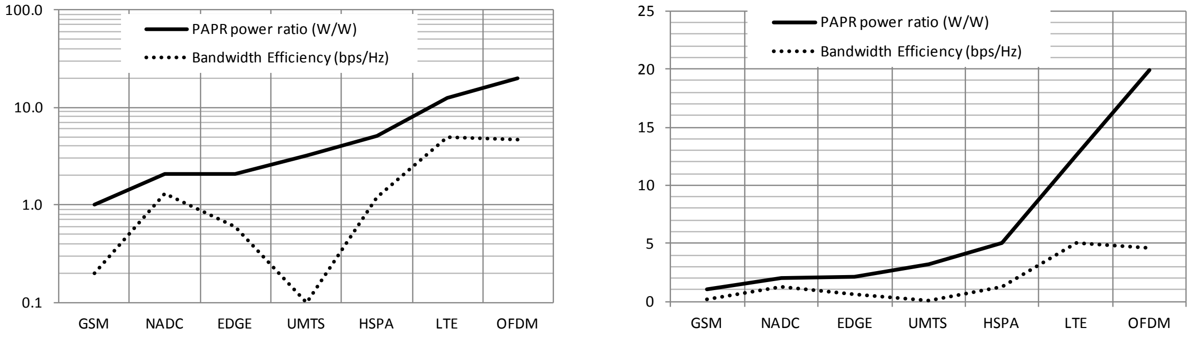
Linearity in circuit performance is not necessarily easy to get either. Particularly in the CMOS geometries below 100 nanometers, the individual transistors get progressively faster and progressively less linear. Architectures that use this dynamic supply voltage for improved energy efficiency that also can use the varying supply to improve output signal accuracy become attractive. The amplifier operating mode must change to have the dynamic power supply shift from being a source of output signal distortion to being a linearizer.

Reliability Improvement

When power dissipation goes down, temperature goes down. When operating temperature goes down, it is well known that circuit reliability goes up exponentially. Therefore the most important parameter that predicts long term reliability of a component is its operating temperature, and the parameter driving the operating temperature is the component power dissipation. Anything that can be done to reduce component power dissipation will improve its reliability. Incorporating a dynamic power supply (DPS) into the transmitter power amplifier (PA) directly reduces power dissipation in the PA.

High Peak to Average Power signal types

Standardization committees have adopted signals in recent decades which have increasingly high peak to average power ratio (PAPR) properties, as shown in Fig. 5 for uplink (mobile to infrastructure) signals. The bandwidth efficiency of these signals does not track well with the PAPR; indeed, there are several signals where the bandwidth efficiency decreases while the PAPR increases. This is particularly true for the 3G signal used in the universal mobile telephone service (UMTS), where the spread spectrum chip code needed for code division multiple access (CDMA) operation expands the signal bandwidth with no change in the information data rate. It is widely assumed that in order to achieve high values of bandwidth efficiency the signal necessarily must have a high PAPR value. This correlation is actually very weak, as the data in Fig. 5 shows.



a b

Fig. 5 PAPR (solid line) and bandwidth efficiency (dotted line) for a progression of uplink signal modulation types in wide use: a) logarithmic scale; b) linear scale

Energy Efficiency

When the primary energy source is a battery, or particularly some type of energy harvesting mechanism, the electron source is best considered as finite. The wireless communication feature exists to communicate, and it is viable to consider the radio with regard to its effectiveness of using these electrons for the needed communication. Implementation architectures and circuits must consider this electron utility factor, the ratio of electrons drawn from the primary energy source to those that actually result in the needed communication signal, to select the most effective option.

Looking only at the transmitter PA, the electron utility factor is exactly equal to the amplifier efficiency factor as shown in Fig. 6. This is obvious, but still useful to describe in this electron utility format because it is electron utility that governs the real design target: battery life.

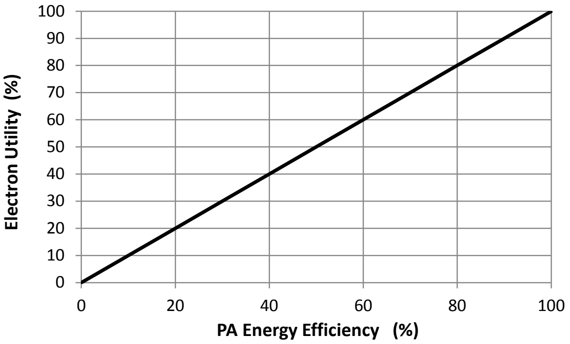


Fig. 6 Electron utility ratio of a transmitter depends on the PA operating energy efficiency: when the energy source is finite (like a battery) signals and circuits that enable higher electron utility are more viable and valuable.

Efficiency improvement vs. Signal PAPR

The effectiveness of dynamic power supply architectures in providing the needed output efficiencies differ among the available DPS architectures as shown in Fig. 4b and 4c. The simplest DPS architecture is called average power tracking (APT), which actually is a misnomer because what really is happening is peak power tracking. In this DPS architecture the voltage applied to the PA is set to be slightly above the signal peak envelope voltage as the output power is varied, in an attempt to always operate the PA linearly at its maximum available output efficiency for the signal being used.

At the other extreme of DPS architectures is one that operates the power amplifier in accordance with Fig. 4c but with the voltage drop inside the amplifier set to zero. This provides the highest possible overall energy efficiency the transmitter can have. This meets the first step in efficient transmitter design, which is to make the PA itself maximally efficient. This is done by eliminating all amplifier circuit linearity.

It then becomes a task of the adopted architecture to make this efficiency ‘visible’ to the local energy source (power supply). This is the task of DPS design: to have sufficiently high efficiency to minimally detract from the energy efficiency realized by the PA. At the highest output power APT is actually less efficient than just the linear amplifier itself. This is because the DPS implementing the APT’s reduced supply voltage has some power dissipation of its own, which is not present when the DPS is not there. The major APT benefit is realized at lower output power levels where the overall efficiency drops much more slowly, which now corresponds more closely with the efficiency curve available from the PA.

References

[1] F. Terman, *Radio Engineers Handbook*, McGraw-Hill, New York, 1943

[2] W.H. Doherty, “A New High Efficiency Power Amplifier for Modulated Waves,” *Proc. of the I.R.E.*, vol. 24, No. 9, Sept. 1936, pp. 1163 - 1182

[3] H. Chireix, "High Power Outphasing Modulation," *Proc. of the I.R.E.*, vol. 23, No. 11, Nov. 1935, pp. 1370-1392

[4] L.R. Kahn, “Single-Sideband Transmission by Envelope Elimination and Restoration,” *Proc. of the I.R.E.*, vol. 40, No. 7, July 1952, pp. 803 - 806

[5] R. A. Heising, “System for the Transmission of Intelligence,” *US Patent 1,137,315*, issued April 27, 1915

[6] J. Staudinger, et al., "800 MHz Power Amplifier Using Envelope Following Technique," *Proc. 1999 IEEE Radio and Wireless Conf.* (RAWCON'99), Denver, CO, Aug. 1 - 4, 1999

[7] E. McCune, *Practical Digital Wireless Signals*, Cambridge University Press, 2010, Appendix G