# Physics of Amplifier Efficiency

Earl McCune, CTO
Eridan Communications

The sinusoidal waveform used in radio communications is not an arbitrary choice, but is a consequence from Maxwell's Equations of electromagnetism. Looking at this solution we see that polar coordinates correspond to the physically natural form of the signal equation. Ohm's Law, itself also a consequence from Maxwell's Equations, shows how power dissipation happens in transmitters. Knowing how power dissipation reduces overall energy efficiency provides guidance on how to change designs to improve overall transmitter efficiency.

It is important to use models, both physical and mathematical, that not only describe well what the performance of these transmitters is, but also to be descriptive of the physical operations. This joint requirement on the models used here is used consistently.

## Maxwell's Equations

All electronics, radio included, follows from electromagnetism described by Maxwell's Equations (usually as reformulated by Oliver Heaviside) [1]. Maxwell's Equations are particularly important here because the solution to this linearly polarized propagating electromagnetic wave equation

$$\frac{\partial^2 B}{\partial z^2} - \mu\varepsilon \frac{\partial^2 E}{\partial t^2} = 0 \tag{1}$$

is a complex exponential of the form

$$E_x(t,z) = E_0 e^{j\Psi(t,z)} \quad ; \quad \Psi(t,z) = \omega\left(t - z/v\right) + \phi(t) \quad ; \quad v = 1/\sqrt{\mu\varepsilon} \tag{2}$$

showing that the magnitude of the electric field is perpendicular to the direction of wave propagation (z) [1, p. 247]. Considering only the time varying aspects of (2) by dropping the spatial propagation term and keeping only the real part of the complex exponential we get the equation

$$s(t) = A\cos\left(\Psi(t)\right) \tag{3}$$

which is expanded into the familiar signal equation by incorporating modulation (time variations) in the amplitude (A) , frequency ($\omega$), and phase-shift ($\phi$) parameters:

$$s(t) = A(t)\cos\left(\omega(t)t + \phi(t)\right) \; ; \; \Psi(t) = \omega(t)t + \phi(t) \quad . \tag{4}$$

This brief review confirms that the sinusoidal signal we use is no accident, but is required by electromagnetic theory. The signal (4) is *naturally of polar form*, not quadrature form. This means that our amplifier hardware provides output signals of particular magnitude, and generally do not care what the signal phase is. Physics shows itself to be polar, not quadrature, in operation.

The quadrature signal equation in wide use

$$s(t) = I(t)\cos\left(\omega(t)\right) + Q(t)\sin\left(\omega(t)\right) \tag{5}$$

is therefore *not* physically descriptive of how amplifier circuits actually work. Model (5) is simply a projection of (4) onto Cartesian axes (ortho-normal basis functions) and is very convenient for mathematical analysis. Using a comparison from history, while the solar system models by both Ptolemy and Copernicus (the latter beautifully generalized by Newton) succeeded in predicting the observed position of the planets, only one of these models is useful for interplanetary travel. We must be careful to read physical significance into our models only to the extent such physical significance is justified. For radio communications, the quadrature model (5) is Ptolemaic in that it beautifully describes all our observations while not describing at all what physically is happening. In this regard polar models are Copernican, in that they also describe what is happening as well as provide physical intuition on what the electrons are doing.

## Ohm's Law

Ohm's Law also follows from Maxwell's Equations [1]. Here though we are particularly interested in power output and power dissipation. Power output is good, and power dissipation within the PA is bad.

Ohm's Law states that power is developed when current flows through a resistance. When that resistance is the PA load, we desire this output power so it is important to have this current (and its corresponding in-phase voltage). But when that resistance is the PA transistor it is important to not have much in-phase current. This very quest has motivated PA research now for more than a century.

Constant power dissipation contours have hyperbolic shapes on a voltage current (VI) plane, as shown in Fig. 2-1. These curves show that power dissipation is low whenever the branch conditions (voltage across and current through) are located close to either axis. The greater the distance to both axes becomes, the greater the dissipated power.

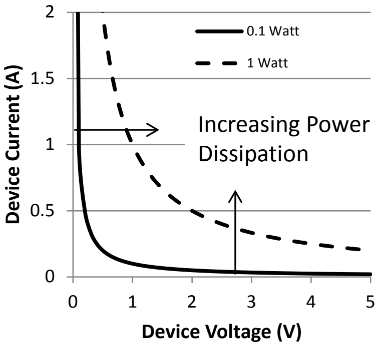


Fig. 2-1. Constant power in a circuit branch follows the hyperbolic curve P=IV

We also want all the power to be useful to (meaning: dissipated in) the load, so the load must be resistive only. If the source impedance $Z_s$ has a non-zero reactive component $X_s$ ($Z_s = R_s + jX_s$) then we apply a "matching" network to provide a conjugate match $Z_L = R_s - jX_s$ to 'tune out' the source reactance. Notice that in this process we have identical resistances, but opposite sign reactances of equal magnitude. This is the very definition of resonance, where reactances 'cancel' leaving only the resistances. Conjugate matching, at a particular frequency, is nothing more than building an appropriate resonator.

Power dissipation in the load is good, but power dissipation in the amplifier is not. Ohm's Law is similarly clear here: to eliminate power dissipation there must not be any in-phase presence of voltage and current in any circuit branch. In particular, this means in the RF transistor, with typical characteristic curves seen in Fig. 2. All linear amplifier theory depends on simultaneous currents and voltages, so energy efficiency is *fundamentally* a problem with linear amplifiers. This particular point continues to drive power amplifier development as this technology enters its second century – how can a in-phase voltage and current be developed in the load while avoiding this same situation within the amplifier? The answer is: by operating the device within the lowest power dissipation regions of its characteristic curves – as near as possible to the current or voltage axes.  Much more is explored about this in Chapters 4, 5, and 6.

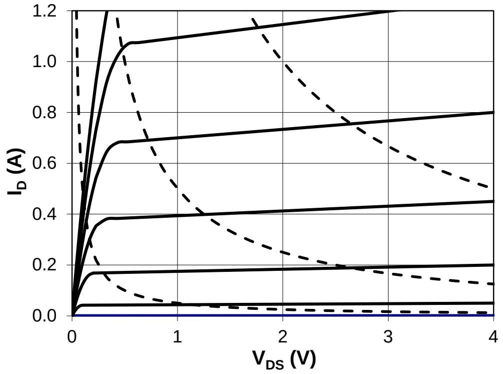


Fig. 2.  Power dissipation in a transistor increases as the operating point moves away from either axis of the characteristic curves: a) characteristic curves for a power FET, with curves of constant power dissipation (at 0.05, 0.5 and 2 watts) shown in dashed lines.

## Supply and Bias Definitions

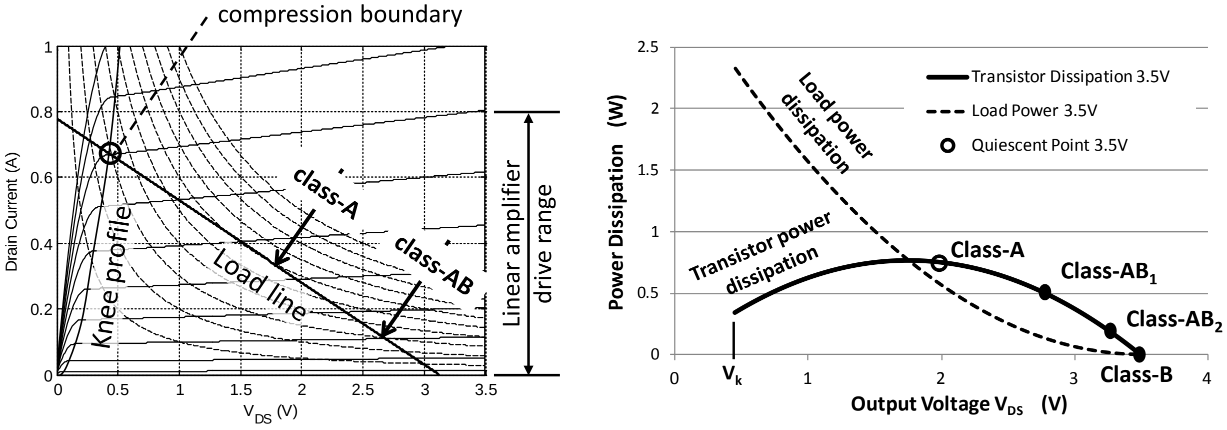
I clearly remember the beginning of my study in electronics, and my difficulty in understanding what was meant by the word *bias*.  Sometimes it was used to mean the entire operating situation of a circuit (e.g. ...with the bias established the amplifier gain is...).  Other readings commented - ...after biasing the grid to -5V then the supply can be applied....  For present purposes it is vital to have clear definitions for the terms *supply* and *bias,* because we encounter many situations where one, the other, or both are varied or dynamic.  After much study the ambiguity between the terms supply and bias is resolved by following these definitions:

- **(power) supply**: a power source applied to the *controlled* device port (collector for bipolar transistors, drain for field effect transistors (FET), plate for vacuum tubes (or valves))
- **bias**: a power source applied to the *controlling* device port (base for bipolar transistors, gate for FETs, grid for vacuum tubes) to set quiescent (no input AC signal) currents at the controlled port

There are publications that use the word bias to refer to the varying power supply, or indeed referring to any power supply.  In the spirit of “we should not change the use of a term unless there is an unavoidably good reason”, the appropriate term is really supply with an appropriate adjective, such as ‘agile supply’ or dynamic supply’.  As noted in the title of this book, I favor using the term ‘dynamic power supply’ to refer to voltages and currents at the controlled port of the RF power transistors.

While the dynamic power supply operates, we may or may not also vary the bias at the controlling port of the RF power transistor.  It is easiest and lowest cost to set the RF transistor bias once and then leave it alone.  It is now known that additional performance is achieved by allowing the bias to also have some dynamic properties.  Fortunately with these definitions, we are free to discuss supply and bias separately and unambiguously.

The operation of conventional linear amplifiers is summarized in Fig. 3a, where traditional transistor characteristic curves and load lines are shown.  New here are the addition of constant power dissipation curves on top of the amplifier design information.  The linear (class-A) bias point along the load line, mid-way between the cut-off and compression boundaries, is very near to the highest value power dissipation curve.  A linear amplifier therefore maximizes power dissipation, which is opposite of what is needed for high energy efficiency. Any effort to improve amplifier efficiency is therefore equivalently an exercise in tolerating circuit nonlinearity, which includes class-AB and class-B operation shown in Fig. 3b. This challenge is more than a century old, and is fundamental.



(a) (b)

Fig. 3. Conventional linear amplifier characteristics: a) transistor characteristic curves with a load line including transistor knee profile and power dissipation curves; b) power dissipation along the load line in the transistor and the load resistance.

Additionally, any transistor operates as the intended controlled current source only when the voltage at its controlled port exceeds a minimum value called the knee voltage ($V_k$), also known as the device compliance voltage. The presence of this knee voltage reduces the available output signal magnitude from the amplifier, and its corresponding power available to the amplifier load resistance also falls. This further reduces the available amplifier efficiency. The impact of the knee voltage to amplifier available output power and energy efficiency is presented in Fig. 4.

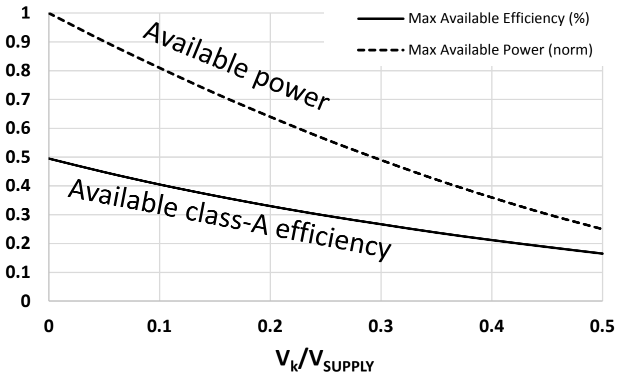
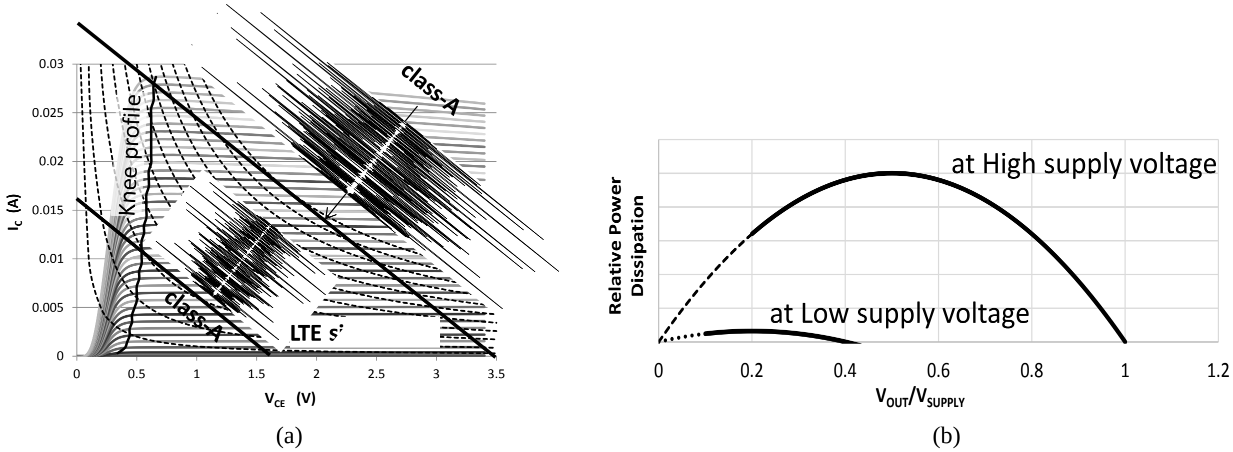


Fig. 4. Losses in output power and available energy efficiency when the transistor knee profile is accounted for

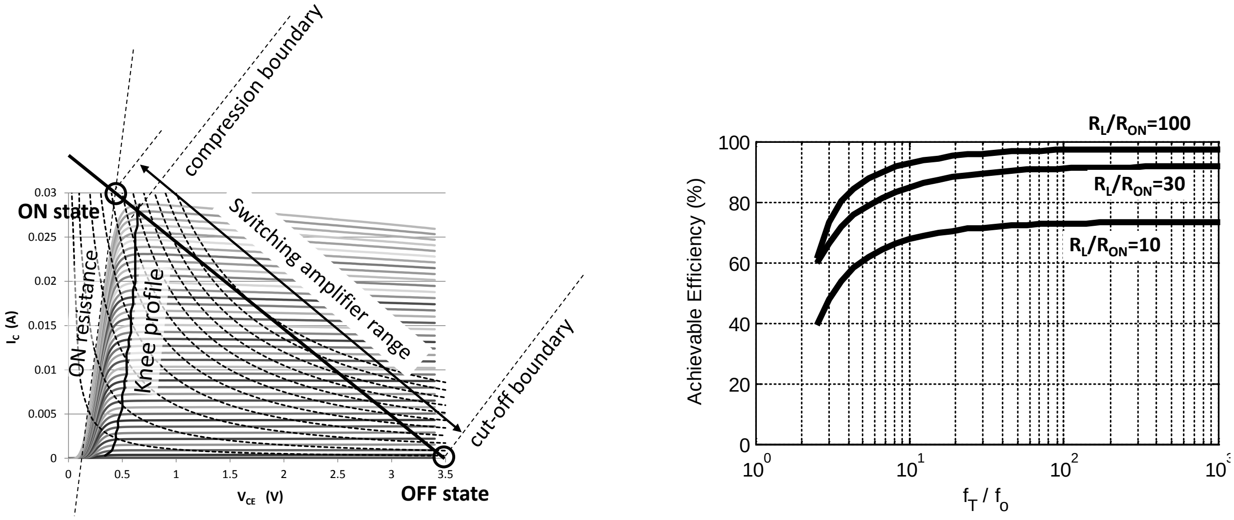
There are two major methods available to improve PA energy efficiency, which is identical to having the PA transistor operate at, or along, power dissipation curves of low value. One option is presented in Fig. 5, where the value of the PA supply voltage is reduced when the required output signal power is low. This shifts the load line while the PA still operates linearly, seen in Fig. 5a using full power and reduced power scales of a LTE uplink signal. The corresponding reduction in dissipated power is substantial, as seen in Fig. 5b. This is the average power tracking (APT) technique, a first step toward the envelope tracking technique [2].



(a) (b)

Fig. 5. Improving energy efficiency with variable supply voltage: a) intersecting lower value power dissipation curves with a lower supply voltage for smaller output signals (LTE uplink is shown); b) corresponding power dissipation profiles at these two supply voltages

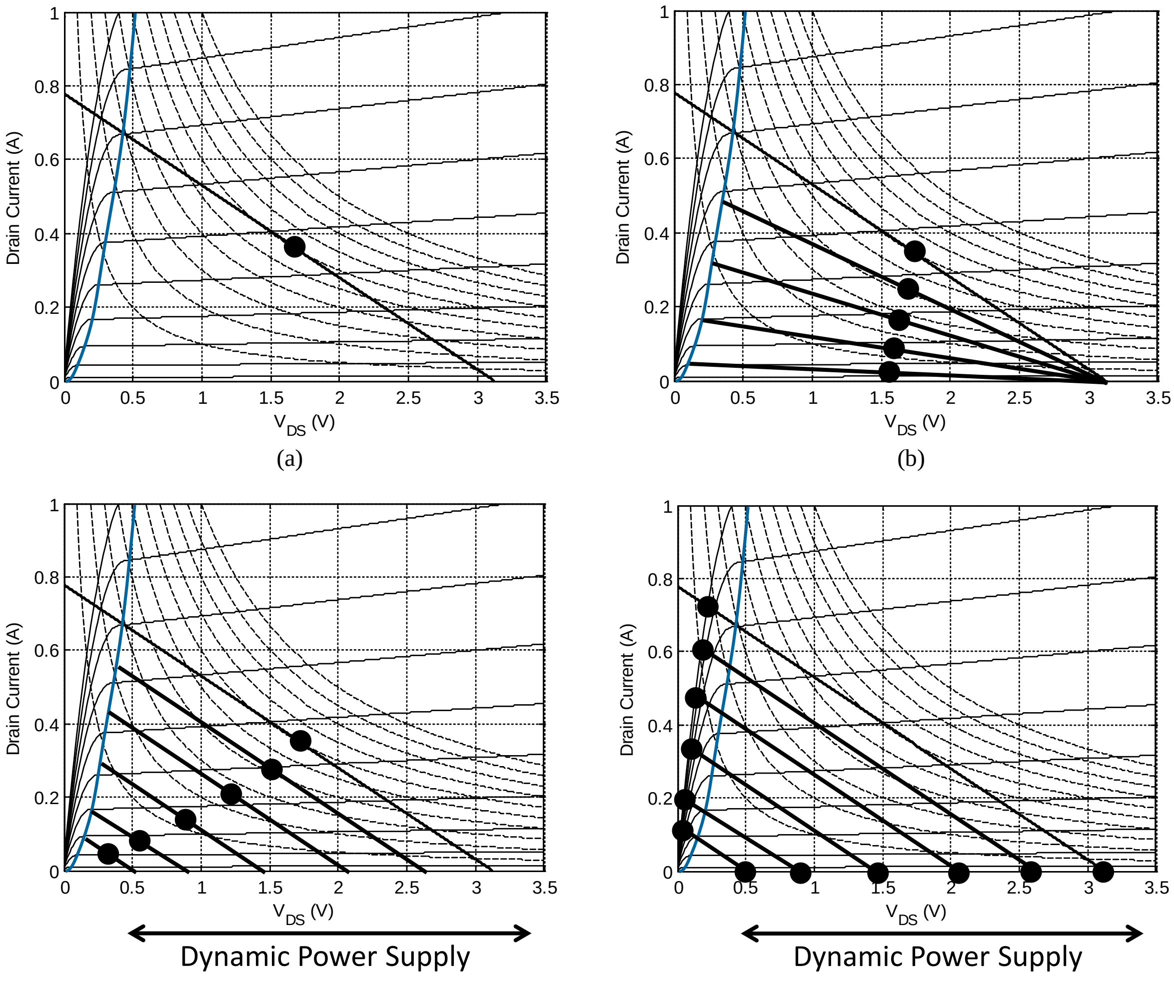
A second technique is to operate the PA only at the endpoints of its load line. This intersects the minimum possible value power dissipation curves, which maximizes the available PA energy efficiency. But it also eliminates all PA circuit linearity [3]. This is called a switch-mode power amplifier (SMPA) with operating characteristics shown in Fig. 6a. The output signal has a larger range because there is no restriction to operating the transistor as a current source. This means that more output power is available from the same transistor operating as a SMPA than as a linear PA. To successfully achieve SMPA operation and its associated efficiency it is necessary that the transitions between transistor ON and OFF states be fast compared to the signal frequency cycle time. This is somewhat related to the transistor transition frequency ($f_T$) specification as shown in Fig. 6b [4]. These results show that it is needed that the transistor $f_T$ to exceed the operating frequency by at least 20x. It is best when the transistor exceeds the operating frequency by at least 50x. In other words, in order to have well performing switching PA operation, the operating frequency should be less than 2% of the transistor $f_T$.



(a) (b)

Fig. 6. Switching power amplifier operation and requirements: a) output signal swing increase from the maximum available from linear operation; b) achieving switching operation requires the transistor transition frequency ($f_T$) to exceed the operating frequency by 20x or more

Between these two extremes of amplifier transistor operation, there are many options with different trade-offs between circuit linearity (with minimum efficiency) and switching operation (having no circuit linearity). Two of these middle cases are presented in Fig. 7 along with the aforementioned extreme cases. The baseline linear amplifier is shown in Fig. 7a with a simplified version of that in Fig. 3a. The use of load impedance modulation with a fixed supply voltage is shown in Fig. 7b, where the improvement in efficiency is evident from the load lines intersecting contours of lower power dissipation as load impedance increases. Envelope tracking (ET) is shown in Fig. 7c, where the PA remains linear at all times and the action of the dynamic supply variation intersects contours of lower power dissipation as the signal envelope falls. Direct polar (DP) modulation [5], requiring SMPA operation, is presented in Fig. 7d. The difference between ET and DP operation is very significant.



(a) (b)

(c) (d)

Fig 7. PA strategies for improved energy efficiency: a) standard linear (baseline); b) Load modulation; c) envelope tracking; d) direct polar modulation

Comparing the transmitter efficiency enhancement techniques of ET and DP to that of conventional linear amplifiers in Fig. 8 shows the superiority of switching operation when energy efficiency, and available output power, must be maximized.

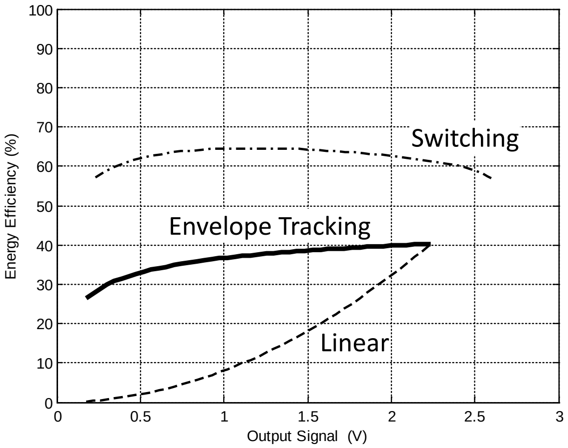


Fig. 8. Available efficiency increases using envelope tracking (linear) and polar (switching) power amplifiers and dynamic power supplies.

It is also possible to separate the phases of the voltage and current waveforms at the transistor output, with the goal of having one or the other waveform near a zero value at all times. This is what class-E amplifiers do. Transitions between the transistor ON and OFF states do not traverse the load line, but follow near to very low value power dissipation contours. This is demonstrated in Fig. 9.

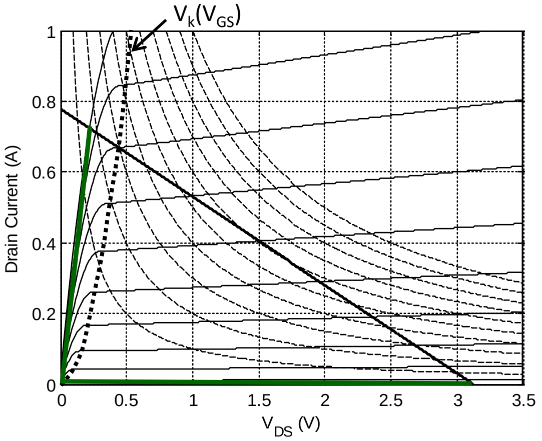
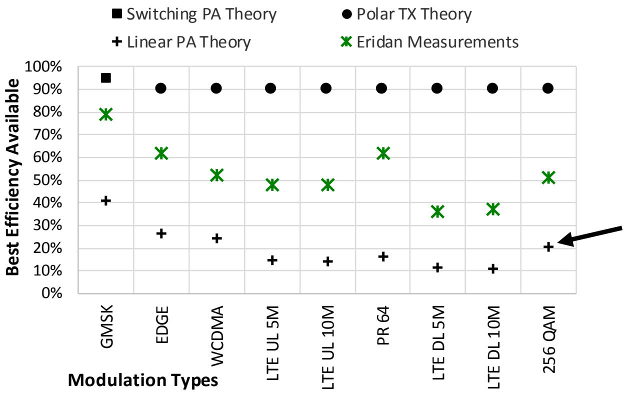


Fig. 9. Keeping the PA transistor current and voltage waveforms phased such that only one of them is non-zero at any particular time produces the transition profile shown here. This is the objective of class-E operation, along with other proposed PA architectures.

When linear PA designs are used, the maximum energy efficiency that can be obtained, even under ideal conditions, is shown by the (+) markers in Fig. 10. The trend for this energy efficiency ceiling is downward for newer signal types. Also shown in Fig. 10 are the markers for the theoretical energy efficiency possibilities for switching PA operation (■) and for polar modulation (●). Polar modulation is now successfully implementing all modulation types, including 256-QAM, OFDM, and LTE. The resulting energy efficiency is presently much greater than the ceiling which limits linear amplifiers.



Modulation-available energy efficiency limits using linear PA

Fig. 10. Theoretical energy efficiency limits by technology for 9 different signal modulation types, and the presently achieved energy efficiency values at Eridan using their polar modulation technology.

The primary goal then, which now is achieved, is to apply polar modulation to signals which possess envelope zero crossings so that it is completely compatible with the presently standardized signal modulations.

## References

[1] S. Ramo, J. Whinnery, T. Van Duzer, *Fields and Waves in Communication Electronics*, John Wiley & Sons, New York, 1965

[2] F. H. Raab, “Intermodulation distortion in Kahn-technique transmitters,” *IEEE Trans. Microwave Theory Tech.*, vol. 44, Dec. 1996, pp. 2273–2278

[3] L. R. Kahn, “Single sideband transmission by envelope elimination and restoration,” *Proc. IRE*, vol. 40, no. 7, pp. 803-806, July 1952.

[4] E. McCune, “Fundamentals of Switching RF Power Amplifiers,” IEEE Microwave and Wireless Components Letters, Dec. 2015

[5] E. McCune, “Operating Modes of Dynamic-Power-Supply Transmitter Amplifiers,” IEEE Trans. on Microwave Theory and Techniques, vol. 62, no. 11, Nov. 2014, pp. 2511-2517