RF Power Amplifier Classes of Operation

(Class A, AB, B, C, D, E, F, G and H)

By Larry E. Gugle K4RFE, RF Design, Manufacture, Test & Service Engineer (Retired)

Conduction Angle or Angle of Flow

Output stages of Power amplifier circuits are Class A, B, AB and C for 'Analog Designs', and Class D and E for 'Switching Designs' based on the 'Conduction Angle or *Angle of Flow* (Θ)', of the input signal through the (or each) output amplifying device, that is, the portion of the input signal cycle during which the amplifying device conducts. The image of the conduction angle is derived from amplifying a sinusoidal signal. (If the device is always on, Θ = 360°.) The angle of flow is closely related to the amplifier power efficiency.

The classes can be most easily understood using the diagrams in each section below. For the sake of illustration, a Bipolar Junction Transistor (BJT) is shown as the amplifying device, but in practice this could be a Metal Oxide Semiconductor Field Effect Transistor (MOSFET) or Electron (Vacuum) Tube device. In an analog amplifier (the most common kind), the signal is applied to the input terminal of the device (base, gate or grid), and this causes a proportional output drive current to flow out of the output terminal. The output drive current comes from the power supply.

Class A

100% of the input signal is used (conduction angle Θ = 360° or 2π); i.e., the active element remains conducting (works in its "linear" range) all of the time. Where efficiency is not a consideration, most small signal linear amplifiers are designed as class A. Class-A amplifiers are typically more linear and less complex than other types, but are very inefficient. This type of amplifier is most commonly used in small-signal stages or for low-power applications (such as driving headphones). Subclass A2 is sometimes used to refer to vacuum-tube class-A stages where the grid is allowed to be driven slightly positive on signal peaks, resulting in slightly more power than normal class A and Subclass A1 is where the grid is always negative, but incurring more distortion.

Amplifying devices operating in class A conduct over the whole of the input cycle. A *class-A amplifier* is distinguished by the *output stage* being biased into class A.

Advantages of class-A amplifiers

Class-A designs are simpler than other classes; for example class-AB and -B designs require two devices (push–pull output) to handle both halves of the waveform; class A can use a single device single-ended.

The amplifying element is biased so the device is always conducting to some extent, normally implying the quiescent (small-signal) collector current (for transistors; drain current for FETs or anode/plate current for vacuum tubes) is close to the most linear portion of its transconductance curve.

Because the device is never shut off completely there is no "turn on" time, little problem with charge storage, and generally better high frequency performance and feedback loop stability (and usually fewer high-order harmonics).

The point at which the device comes closest to being cut off is not close to zero signal, so the problem of crossover distortion associated with class-AB and -B designs is avoided.

Disadvantage of class-A amplifiers

They are very inefficient; a theoretical maximum of 50% is obtainable with inductive output coupling and only 25% with capacitive coupling, unless deliberate use of nonlinearities is made (such as in square-law output stages). In a power amplifier this not only wastes power and limits battery operation, it may place restrictions on the output devices that can be used (for example: ruling out some audio triodes if modern low-efficiency loudspeakers are to be used), and will increase costs. Inefficiency comes not just from the fact that the device is always conducting to some extent (that happens even with class AB, yet its efficiency can be close to that of class B); it is that the standing current is roughly half the maximum output current (although this can be less with square law output stage), together with the problem that a large part of the power supply voltage is developed across the output device at low signal levels (as with classes AB and B, but unlike output stages such as class D). If high output powers are needed from a class-A circuit, the power waste (and the accompanying heat) will become significant. For every watt delivered to the load, the amplifier itself will, *at best*, dissipate another watt. For large powers this means very large and expensive power supplies and heat sinking.

Class-A designs have largely been superseded by the more efficient designs for power amplifiers, though they remain popular with some hobbyists, mostly for their simplicity. Also, many audiophiles believe that class A gives the best sound quality (for their absence of crossover distortion and reduced odd-harmonic and high-order harmonic distortion) which provides a small market for expensive **high fidelity** class-A amps.

Single-ended and triode class-A amplifiers

Some aficionados who prefer class-A amplifiers also prefer the use of thermionic Electron (Vacuum) Tube designs instead of transistors, especially in Single-ended triode output configurations for several claimed reasons:

Single-ended output stages (be they tube or transistor) have an asymmetrical transfer function, meaning that even order harmonics in the created distortion tend not to be canceled (as they are in push–pull output stages); by using tubes or FETs most of the distortion is from the square law transfer characteristic and so second-order, which some consider to be "warmer" and more pleasant.

For those who prefer low distortion figures, the use of tubes with class A (generating little odd-harmonic distortion, as mentioned above) together with symmetrical circuits (such as push–pull output stages, or balanced low-level stages) results in the cancellation of most of the even distortion harmonics, hence the removal of most of the distortion.

Though good amplifier design can reduce harmonic distortion patterns to almost nothing, distortion is essential to the sound of electric guitar amplifiers, for example, and is held by recording engineers to offer more flattering microphones and to enhance "clinical-sounding" digital technology.

Historically, Electron (Vacuum) Tube amplifiers often used a class-A power amplifier simply because Electron (Vacuum) Tubes are large and expensive; many class-A designs use only a single device.

Transistors are much cheaper, and so more elaborate designs that give greater efficiency but use more parts are still cost-effective. A classic application for a pair of class-A devices is the long-tailed pair, which is exceptionally linear, and forms the basis of many more complex circuits, including many audio amplifiers and almost all op-amps.

Class-A amplifiers are often used in output stages of high quality op-amps (although the accuracy of the bias in low cost op-amps such as the **741** may result in class A or class AB or class B, varying from device to device or with temperature). They are sometimes used as medium-power, low-efficiency, and high-cost audio amplifiers. The power consumption is unrelated to the output power. At idle (no input), the power consumption is essentially the same as at high output volume. The result is low efficiency and high heat dissipation.

Class B

50% of the input signal is used (Θ = 180° or π/4); i.e., the active element conducts half of the time and is more or less turned off for the other half (works in its "linear" range). In most class B, there are two output devices (or sets of output devices), each of which conducts alternately (push–pull) for exactly 180° (or half cycle) of the input signal; selective RF amplifiers can also be implemented using a single active element.

These amplifiers are subject to *crossover distortion* if the transition from one active element to the other is not perfect, as when two complementary transistors (i.e., one PNP, one NPN) are connected as two emitter followers with their base and emitter terminals in common, requiring the base voltage to slew across the region where both devices are turned off.

Class-B amplifiers only amplify half of the input wave cycle, thus creating a large amount of distortion, but their efficiency is greatly improved and is much better than class A. Class B has a maximum theoretical efficiency of π/4. (i.e. 78.5%) This is because the amplifying element is switched off altogether half of the time, and so cannot dissipate power. A single class-B element is rarely found in practice, though it has been used for driving the loudspeaker in the early IBM Personal Computers with beeps, and it can be used in RF power amplifier where the distortion levels are less important. However, class C is more commonly used for this.

A practical circuit using class-B elements is the push–pull stage, such as the very simplified complementary pair arrangement shown below. Here, complementary or quasi-complementary devices are each used for amplifying the opposite halves of the input signal, which is then recombined at the output. This arrangement gives excellent efficiency, but can suffer from the drawback that there is a small mismatch in the cross-over region – at the "joins" between the two halves of the signal, as one output device has to take over supplying power exactly as the other finishes. This is called crossover distortion. An improvement is to bias the devices so they are not completely off when they're not in use. This approach is called *class AB* operation.

Class-B push–pull amplifier

Class AB

In class-AB operation, each device operates the same way as in class B over half the waveform, but also conducts a small amount on the other half. As a result, the region where both devices simultaneously are nearly off (the "dead zone") is reduced. The result is that when the waveforms from the two devices are combined, the crossover is greatly minimised or eliminated altogether. The exact choice of quiescent current, the standing current through both devices when there is no signal, makes a large difference to the level of distortion (and to the risk of thermal runaway, that may damage the devices); often the bias voltage applied to set this quiescent current has to be adjusted with the temperature of the output. Often used as thermally tracking bias voltages is to include small value resistors in series with the emitters.

Here the two active elements conduct more than half of the time ($\Theta > 180° < 360°$) as a means to reduce the cross-over distortions of class-B amplifiers. In the example of the complementary emitter followers a bias network allows for more or less quiescent current thus providing an operating point somewhere between class A and class B. Sometimes a figure is added (e.g., AB_1 or AB_2).

1. class AB_1: for vacuum-tube stages where the Control Grid voltage is always negative with respect to the cathode and **indicates that Control Grid (CG) current does not flow during any part of the input cycle.**
2. class AB_2: for vacuum-tube stages where the Control Grid voltage may be slightly positive, and **indicates that Control Grid (CG) current does flow during some part of the input cycle,** adding more distortion, but giving slightly higher output power on signal peaks.
3. Solid-state class-AB amplifier circuits are one of the most popular amplifier topologies used today.

Class-B or Class-AB push–pull circuits are the most common design type found in audio power amplifiers. Class AB is widely considered a good compromise for audio amplifiers, since much of the time the music is quiet enough that the signal stays in the "class A" region, where it is amplified with good fidelity, and by definition if passing out of this region, is large enough that the distortion products typical of class B are relatively small. The crossover distortion can be reduced further by using negative feedback. Class-B and Class-AB amplifiers are sometimes used for RF linear amplifiers as well. Class-B amplifiers are also favored in battery-operated devices, such as transistor radios.

Class-AB sacrifices some efficiency over class B in favor of linearity, thus is less efficient (below 78.5% for full-amplitude sinewaves in transistor amplifiers, typically; much less is common in Class-AB vacuum-tube amplifiers). It is typically much more efficient than class A.

Class C

Less than 50% of the input signal is used (conduction angle Θ < 180°). The advantage is potentially high efficiency, but a disadvantage is high distortion.

Class-C amplifiers conduct less than 50% of the input signal and the distortion at the output is high, but high efficiencies (up to 90%) are possible. Some applications (for example, megaphones) can tolerate the distortion. The usual application for class-C amplifiers is in RF transmitters operating at a single fixed carrier frequency, where the distortion is controlled by a tuned load on the amplifier. The input signal is used to switch the active device causing pulses of current to flow through a tuned circuit forming part of the load.

The class-C amplifier has two modes of operation: tuned and untuned. The diagram shows a waveform from a simple class-C circuit without the tuned load. This is called untuned operation, and the analysis of the waveforms shows the massive distortion that appears in the signal. When the proper load (e.g., an inductive-capacitive filter plus a load resistor) is used, two things happen. The first is that the output's bias level is clamped with the average output voltage equal to the supply voltage. This is why tuned operation is sometimes called a *clamper*. This allows the waveform to be restored to its proper shape despite the amplifier having only a one-polarity supply. This is directly related to the second phenomenon: the waveform on the center frequency becomes less distorted. The residual distortion is dependent upon the bandwidth of the tuned load, with the center frequency seeing very little distortion, but greater attenuation the farther from the tuned frequency that the signal gets.

The tuned circuit resonates at one frequency, the fixed carrier frequency, and so the unwanted frequencies are suppressed, and the wanted full signal (sine wave) will be extracted by the tuned load. The signal bandwidth of the amplifier is limited by the Q-factor of the tuned circuit but this is not a serious limitation. Any residual harmonics can be removed using a further filter.

In practical Class-C amplifiers a tuned load is invariably used. In one common arrangement the resistor shown in the circuit above is replaced with a parallel-tuned circuit consisting of an inductor and capacitor in parallel, whose components are chosen to resonate the frequency of the input signal. Power can be coupled to a load by transformer action with a secondary coil wound on the inductor. The average voltage at the collector, drain or plate is then equal to the supply voltage, and the signal voltage appearing across the tuned circuit varies from near zero to near twice the supply voltage during the rf cycle. The input circuit is biassed so that the active element (e.g. transistor) conducts for only a fraction of the RF cycle, usually one third (120 degrees) or less.

The active element conducts only while the collector, drain or plate voltage is passing through its minimum. By this means, power dissipation in the active device is minimised, and efficiency increased. Ideally, the active element would pass only an instantaneous current pulse while the voltage across it is zero: it then disspates no power and 100% efficiency is achieved. However practical devices have a limit to the peak current they can pass, and the pulse must therefore be widened, to around 120 degrees, to obtain a reasonable amount of power, and the efficiency is then 60-70%.

Class D

These use switching to achieve a very high power efficiency (more than 90% in modern designs). By allowing each output device to be either fully on or off, losses are minimized. The analog output is created by Pulse-Width Modulation (PWM); i.e., the active element is switched on for shorter or longer intervals instead of modifying its resistance. There are more complicated switching schemes like sigma-delta modulation, to improve some performance aspects like lower distortions or better efficiency.

Block diagram of a basic switching or PWM (class-D) amplifier.

Boss Audio class-D mono car audio amplifier with a low pass filter for powering subwoofers

In the class-D amplifier the input signal is converted to a sequence of higher voltage output pulses. The averaged-over-time power values of these pulses are directly proportional to the instantaneous amplitude of the input signal. The frequency of the output pulses is typically ten or more times the highest frequency in the input signal to be amplified. The output pulses contain inaccurate spectral components (that is, the pulse frequency and its harmonics) which must be removed by a low-pass passive filter. The resulting filtered signal is then an amplified replica of the input.

These amplifiers use pulse width modulation, pulse density modulation (sometimes referred to as pulse frequency modulation) or a more advanced form of modulation such as Delta-sigma modulation (for example, in the Analog Devices AD1990 class-D audio power amplifier).

Output stages such as those used in pulse generators are examples of class-D amplifiers. The term *class-D* is usually applied to devices intended to reproduce signals with a bandwidth well below the switching frequency.

Class-D amplifiers can be controlled by either analog or digital circuits. The digital control introduces additional distortion called *quantization error* caused by its conversion of the input signal to a digital value.

The main advantage of a class-D amplifier is power efficiency. Because the output pulses have a fixed amplitude, the switching elements (usually MOSFETs, but Electron Tubes and bipolar transistors were once used) are switched either completely on or completely off, rather than operated in linear mode. A MOSFET operates with the lowest resistance when fully on and thus has the lowest power dissipation when in that condition, except when fully off. When operated in a linear mode the MOSFET has variable amounts of resistance that vary linearly with the input voltage and the resistance is something other than the minimum possible, therefore more electrical energy is dissipated as heat. Compared to class-AB operation, class D's lower losses permit the use of a smaller heat sink for the MOSFETS while also reducing the amount of AC power supply power required. Thus, class-D amplifiers do not need as large or as heavy power supply transformers or heatsinks, so they are smaller and more compact in size than an equivalent class-AB amplifier.

Class-D amplifiers have been widely used to control motors, and almost exclusively for small DC motors, but they are now also used as audio amplifiers, with some extra circuitry to allow analogue to be converted to a much higher frequency pulse width modulated signal. The relative difficulty of achieving good audio quality means that nearly all are used in applications where quality is not a factor, such as modestly priced bookshelf audio systems and "DVD-receivers" in mid-price home theater systems.

High quality class-D audio amplifiers have now appeared in the market and these revised designs have been said to rival good traditional AB amplifiers in terms of quality. Before these higher quality designs existed an earlier use of class-D amplifiers and prolific area of application was high-powered, subwoofer amplifiers in cars. Because subwoofers are generally limited to a bandwidth of no higher than 150 Hz, the switching speed for the amplifier does not have to be as high as for a full range amplifier. Class-D amplifiers for driving subwoofers are relatively inexpensive, in comparison to class-AB amplifiers.

The letter *D* used to designate this amplifier class is simply the next letter after *C*, and does not stand for *digital*. Class-D and class-E amplifiers are sometimes mistakenly described as "digital" because the output waveform superficially resembles a pulse-train of digital symbols, but a class-D amplifier merely converts an input waveform into a continuously Pulse-Width Modulated (PWM) (square wave) analog signal. (A digital waveform would be Pulse-Code Modulated (PCM)).

Additional classes

There are several other amplifier classes, although they are mainly variations of the previous classes.

Class-E and class-F amplifiers are commonly described in literature for Radio Frequency (RF) applications where efficiency of the traditional classes is important, yet several aspects deviate substantially from their ideal values. *These classes use harmonic tuning of their output networks to achieve higher efficiency and can be considered a subset of class C due to their conduction-angle characteristics.*

Class E

The class-E and class-F amplifier is a highly efficient switching power amplifier, typically used at such high frequencies that the switching time becomes comparable to the duty time. As said in the class-D amplifier, the transistor is connected via a serial LC circuit to the load, and connected via a large L (inductor) to the supply voltage. The supply voltage is connected to ground via a large capacitor to prevent any RF signals leaking into the supply. The class-E amplifier adds a C (capacitor) between the transistor and ground and uses a defined L_1 to connect to the supply voltage.

Class-E amplifier

The following description ignores DC, which can be added easily afterwards. The above mentioned C and L are in effect a parallel LC circuit to ground. When the transistor is on, it pushes through the serial LC circuit into the load and some current begins to flow to the parallel LC circuit to ground. Then the serial LC circuit swings back and compensates the current into the parallel LC circuit. At this point the current through the transistor is zero and it is switched off. Both LC circuits are now filled with energy in C and L_0. The whole circuit performs a damped oscillation. The damping by the load has been adjusted so that some time later the energy from the Ls is gone into the load, but the energy in both C_0 peaks at the original value to in turn restore the original voltage so that the voltage across the transistor is zero again and it can be switched on.

With load, frequency, and duty cycle (0.5) as given parameters and the constraint that the voltage is not only restored, but peaks at the original voltage, the four parameters (L, L_0, C and C_0) are determined. The class-E amplifier takes the finite on resistance into account and tries to make the current touch the bottom at zero. This means that the voltage and the current at the transistor are symmetric with respect to time. The Fourier transform allows an elegant formulation to generate the complicated LC networks and says that the first harmonic is passed into the load, all even harmonics are shorted and all higher odd harmonics are open. Class E uses a significant amount of second-harmonic voltage. The second harmonic can be used to reduce the overlap with edges with finite sharpness. For this to work, energy on the second harmonic has to flow from the load into the transistor, and no source for this is visible in the circuit diagram. In reality, the impedance is mostly reactive and the only reason for it is that class E is a class F (see below) amplifier with a much simplified load network and thus has to deal with imperfections.

In many amateur simulations of class-E amplifiers, sharp current edges are assumed nullifying the very motivation for class E and measurements near the transit frequency of the transistors show very symmetric curves, which look much similar to class-F simulations.
The class-E amplifier was invented in 1972 by Nathan O. Sokal and Alan D. Sokal, and details were first published in 1975. Some earlier reports on this operating class have been published in Russian.

Class F

In push–pull amplifiers and in CMOS, the even harmonics of both transistors just cancel. Experiment shows that a square wave can be generated by those amplifiers. Theoretically square waves consist of odd harmonics only. In a class-D amplifier, the output filter blocks all harmonics; i.e., the harmonics see an open load. So even small currents in the harmonics suffice to generate a voltage square wave. The current is in phase with the voltage applied to the filter, but the voltage across the transistors is out of phase. Therefore, there is a minimal overlap between current through the transistors and voltage across the transistors. The sharper the edges, the lower the overlap.

While in class D, transistors and the load exist as two separate modules, class F admits imperfections like the parasitics of the transistor and tries to optimise the global system to have a high impedance at the harmonics. Of course there has to be a finite voltage across the transistor to push the current across the on-state resistance. Because the combined current through both transistors is mostly in the first harmonic, it looks like a sine. That means that in the middle of the square the maximum of current has to flow, so it may make sense to have a dip in the square or in other words to allow some overswing of the voltage square wave. A class-F load network by definition has to transmit below a cutoff frequency and reflect above. Any frequency lying below the cutoff and having its second harmonic above the cutoff can be amplified, that is an octave bandwidth. On the other hand, an inductive-capacitive series circuit with a large inductance and a tunable capacitance may be simpler to implement. By reducing the duty cycle below 0.5, the output amplitude can be modulated. The voltage square waveform will degrade, but any overheating is compensated by the lower overall power flowing. Any load mismatch behind the filter can only act on the first harmonic current waveform, clearly only a purely resistive load makes sense, then the lower the resistance, the higher the current.

Class F can be driven by sine or by a square wave, for a sine the input can be tuned by an inductor to increase gain. If class F is implemented with a single transistor, the filter is complicated to short the even harmonics. All previous designs use sharp edges to minimise the overlap.

Class G and H

Class-G and class-H amplifiers are marked by variation of the supply rails (in discrete steps or in a continuous fashion, respectively) following the input signal. Wasted heat on the output devices can be reduced as excess voltage is kept to a minimum. The amplifier that is fed with these rails itself can be of any class. These kinds of amplifiers are more complex, and are mainly used for specialized applications, such as very high-power units.

There are a variety of amplifier designs that enhance class-AB output stages with more efficient techniques to achieve greater efficiencies with low distortion. These designs are common in large audio amplifiers since the heatsinks and power transformers would be prohibitively large (and costly) without the efficiency increases. The terms "class G" and "class H" are used interchangeably to refer to different designs, varying in definition from one manufacturer or paper to another.

Class-G amplifiers (which use "rail switching" to decrease power consumption and increase efficiency) are more efficient than class-AB amplifiers. These amplifiers provide several power rails at different voltages and switch between them as the signal output approaches each level. Thus, the amplifier increases efficiency by reducing the wasted power at the output transistors. Class-G amplifiers are more efficient than class AB but less efficient when compared to class D, without the negative EMI effects of class D.

Class-H amplifiers take the idea of class G one step further creating an infinitely variable supply rail. This is done by modulating the supply rails so that the rails are only a few volts larger than the output signal at any given time. The output stage operates at its maximum efficiency all the time. Switched-mode power supplies can be used to create the tracking rails. Significant efficiency gains can be achieved but with the drawback of more complicated supply design and reduced THD performance.

The voltage signal shown is thus a larger version of the input, but has been changed in sign (inverted) by the amplification. Other arrangements of amplifying device are possible, but that given (that is, common cathode, common emitter, or common source) is the easiest to understand and employ in practice. If the amplifying element is linear, then the output will be faithful copy of the input, only larger and inverted. In practice, transistors are not linear, and the output will only approximate the input. Non-linearity from any of several sources is the origin of distortion within an amplifier. Which class of amplifier (A, B, AB or C) depends on how the amplifying device is biased — in the diagrams the bias circuits are omitted for clarity.

Any real amplifier is an imperfect realization of an ideal amplifier. One important limitation of a real amplifier is that the output it can generate is ultimately limited by the power available from the power supply. An amplifier will saturate and clip the output if the input signal becomes too large for the amplifier to reproduce or if operational limits for a device are exceeded.