Single Ended vs. Push Pull: The Fight of the Century

by Eddie Vaughn

If you're familiar with tube amplifiers, you know that the two major methods of power stage operation are single ended (SE) and push pull (PP). As with so many things in life, most people are highly opinionated when it comes to these two choices. While any confrontations between the two camps are less likely to end with blood spilled on the floor than say, vinyl versus digital or tube versus solid state (proponents of SE and PP are, after all, still "brothers in tubes"), their clashes can nevertheless become a bit heated at times. The real danger here exists to that poor, well-meaning soul who enters into the discussion and tries to play the peacemaker by extolling the virtues of each method and proclaiming them equals, which is the "raw nerve" for both camps. When this happens, you'll usually see Camp SE and Camp PP rise up in unison, beat him like a rented mule and toss his limp, lifeless body aside as fodder for the vultures, and then resume hostilities as usual betwixt themselves.

Yes, even among the brethren of The Brotherhood Of The Firebottle, there are many schisms. Paper in oil caps versus plastic film caps. Vintage carbon composition resistors versus fancy metal foils and metal films. Vintage amp designs versus modern amp designs. Grid bias versus cathode bias. And then there are various, very adamant "splinter groups." Here, you'll find the "Power Mongers" branch of the Brotherhood. The Power Mongers drive their gargantuan speakers with behemoth devices that give off an infrared heat signature detectable by satellites, and more closely resemble 800 megawatt nuclear reactors than tube amps. No "girlie man" amps for them! Sadly, many Power Mongers wind up in jail, after finally getting busted for the racketeering and smuggling operations they had operated in order to pay their electric bills.

Then, there are the Old Timers, those easy going, laid back fellows who love to enjoy their music with the "warm and fuzzy" sounds of the bygone Golden Era of Hi-Fi. Old Timer is not an age, it's a mindset. They are some pretty young Old Timers out there, who have usually been accepted into the Brotherhood under the wing of an older Old Timer. Every bit as dedicated as the other splinter groups, Old Timers would rather have their Eico HF-87 than a free lifetime membership in the Ferrari of the Month Club and a date with Norah Jones. Life is good.

Finally, there is the most hardcore sect of the entire Brotherhood, The Triode Junkies, which is further divided into the Transmitter Triode Vikings, the DHT Dandys, the Sweep Tube Ninjas, and the fractional-watt Spudmeisters.

Transmitter Triode Vikings are those aggressive individuals who want to crank it to SPLs that give you a nosebleed, but want to do so with class and finesse. Their way is an iron fist in a velvet glove. They differ from the Power Mongers in that they posess a seething dislike for push pull. As a matter of fact, some Transmitter Triode Vikings began life as Power Mongers, but were lured by the seductive sounds of SET, and became Triode Junkies. They do however still retain some of their old Power Monger ways, and 2 watts just won't cut it for them. Their quest for single ended finesse along with enough power to arc weld often drives them to the brink of madness. They do not at all mind the fact that there is nothing more than a thin ceramic cap standing between them and 1500 volts, neither do they mind a level of radiated heat that would leave even Power Mongers reaching for an asbestos suit.

The DHT Dandys are a crew of refined tastes, who enjoy the finer things in life. They mate their Avantgardes with their $35,000 monoblocks via speaker cables that cost more than the yearly Gross Domestic Product of most developing nations. If invited to their home for a listen to their system, you'll find most of them to be very hospitable and civil gentleman, with all the prim and proper ways of a British butler. Alas, there is the occasional DHT Dandy who can be a bit snobbish about the fact they spent more on their stereo than you did your house, but most are true gentlemen who simply love music. However, the situation can deteriorate very quickly if you mention that your favorite amplifier does not contain a DHT power tube, leading you to be escorted from the premises by some very scary looking bodyguards with large bulges under the torsos of their suits. So please, keep your opinions to yourself, be quick to compliment their system, and even if you do not like the caviar and truffles they offer you, pretend to like them anyway.

Sweep Tube Ninjas are "the thinking man's triode addicts". Cunning and crafty, with a million tricks up their sleeve, they are true to their namesake. Being the thinkers and doers thay they are, they can recite the pinout of every television vertical amplifier tube in existence from memory, and delight in building a chintzy looking amp for less than $250 (including $6 tubes) that will amaze and confound the DHT Dandys by it hanging with their $35,000 monoblocks.

The Spudmeisters are a relatively young group that is small and unique, just like their amplifiers. Their common bond is the spud (one tuber) amp. One resistor, one tube, and one output transformer in the signal path per channel, nothing else. Any more than that, and a Spudmeister will fall to the floor convulsing, ears covered, at the thought of putting anything extra in the signal path to contaminate the signal! Coupling capacitors are the ultimate statement of blasphemy to the Spudmeister. Though they only have one scant watt on tap, they don't seem to mind. It's all about quality, not quantity.

Various individuals within the rank and file of The Triode Junkies belong to a dark, fierce, and shadowy sub-sect, the Triode Zombies, who are the mindlessly dedicated Dark Overlords of the Three Element Tube. People who listen to pentode or ultralinear amps are no friends of theirs and are not welcome in their home. Enough said.

Getting serious for a moment, there are two broad categories that all these sects fall into, whether you're an Old Timer or a Spudmeister. Those two categories are Single Ended and Push Pull. Much has been written on how each works, but very little on their *direct comparison*. So, in this article I'll attempt to bring out the virtues and drawbacks of both single ended and push pull operation, lay them out side-by-side, and try to avoid raising the ire of both Camp Push Pull and Camp Single Ended along the way. Enjoy. :^)

How SE Works

Single ended means just that. In layman's terms, a single power tube does all the work. Actually, SE amplifiers may have more than one power tube per channel (parallel single ended, or "PSE"), but each of the parallel power tubes function exactly alike, and handle the entire AC music signal sinewave form together in unison. In other words, they together function as a single tube. Because it must handle the entire waveform, the SE output stage runs in Class A operating class. This means it is biased so that it conducts plate current throughout the entire 360° of the AC signal cycle, and maintains the tube in a highly linear region of it's operation at all times.

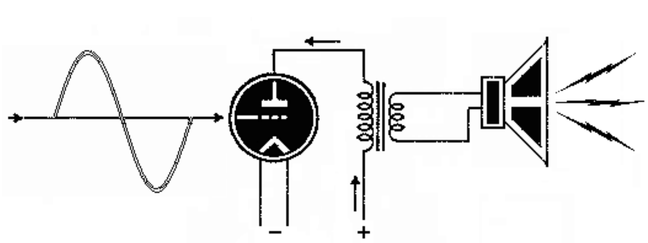


Fig.1 *Single Ended Output Stage*

Fig. 1 above is a typical single ended output stage. This 1940's schematic diagram depicts a directly heated triode, it's output transformer, and the loudspeaker. You'll notice that the single output device handles the entire waveform. As a side note here, some people are shocked (no pun intended) to find out that the music signal in their system is nothing more than an AC voltage, just like your home's AC wall voltage! Depending on where you live, your AC wall voltage has a frequency of either 50 or 60Hz, a single musical tone. The only difference in it and the music signal is that the music signal of course consists of many simultaneous frequencies.

Back to our SE output stage. The reason it must be biased so hot is because of the effect that the AC music signal's negative half-cycle has on the tube itself. Any tube stage must have it's operating point, or "bias" set by making it's control grid some negative voltage value versus it's cathode voltage. This sets the idle point of the tube, or else it would run wide open and quickly burn up. The AC signal's negative half-cycle has the effect of adding to the negative bias voltage present, which reduces current flow just as if you had increased the negative bias voltage. You can think of this as waves in water. Every wave has a crest and a trough. We can liken the crest to our positive AC half-cycle, and the trough as our negative AC half-cycle. The trough is a depression, and when it is added to another depression, the sum result is an even deeper depression! In Class A operation, the tube is biased hot enough that it does not stop conducting current at the peak of the negative half-cycle. If it did, it would not be able to amplify that portion of the signal. It must not only be biased hot enough to merely conduct at the peak of the negative half-cycle, it must still be conducting *enough* current at this point to keep the tube out of the highly non-linear (high distortion) region near it's conduction cutoff (zero plate current). To do so is very inefficient, and Class A's lower distortion and sweet sound come at a high cost of low power, high heat, and faster tube wear.

How PP Works

In push pull operation, two tubes work together as in PSE, but they do so on *alternate* cycles of the input waveform, instead of the *same* cycle. Each tube of the push pull pair bears half of the work, so to speak, instead of SE where any and all power tubes bear *all* the work. To accomplish push pull operation, the tubes of the push pull pair must be fed signals that are in opposite phase to each other, but are otherwise identical in all respects. The "phase inverter" does this. There are several different tubed phase inverter geometries, and each has advantages and disadvantages. Each of them works by splitting the signal into a copy of the original that is in-phase, and a copy of the original that is 180° out of phase. The two phase inverter outputs are simultaneously fed to their respective power tubes. This means the output signals of the power tubes are of course out of phase with one another. They are summed together in the OPT as the signal that drives the speakers. Fig. 2 below illustrates this.

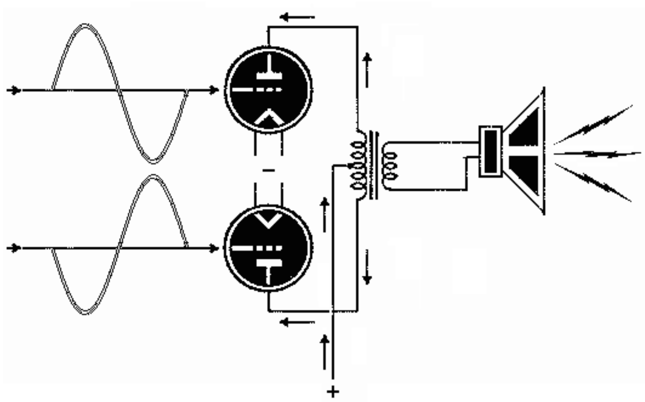


Fig.2 *Push Pull Output Stage*

Note that the input signal going to each tube of the push pull pair is a duplicate of the other, just 180° out of phase, or "inverted phase." You can see how that when one tube is at it's peak conduction when being fed with the peak of the positive half-cycle, his brother across the aisle is chilling out and resting as he is simultaneously fed with the peak of the negative half-cycle. Then, the cycle switches. Without going into a long technical explanation of precisely what happens in a PP output stage, suffice it to say that it does just that, one tube pushes while the other pulls, and back and forth. It can be compared to how SE operates with a very simple analogy.

Think of PP as two men in a small boat, with one rowing on each side. One rows while the other removes his paddle from the water and moves it forward to row again. This motion is repeated back and forth in alternating cycles. Though one is rowing on the starboard side and one on the port side, the sum of their rowing actions is a straight motion. Each expends only half the energy required to propel the boat at a given speed. You can liken SE to only one man in the same boat. He alone not only has to muster the same rowing energy that two men did together, but he must use a lot of his energy in turning his torso from side to side as he rows on each side of the boat in order to make it go straight, wasted energy that does not profit in making the boat move. It is wasteful and inefficient compared to two men rowing together, and the one man is simply not capable of expending the same level of energy as the two men combined are capable of.

Efficiency and Operating Class

Here we see the major reason for using PP, it's high efficiency versus SE. Actually, if the PP topology is ran at the same high current Class A operating point as a SE amp ("high" or "hard" Class A operation), it would produce the same power as a parallel single ended amplifier. No power is gained over PSE by using the PP topology if it is ran in high Class A, because high Class A is so inefficient (typically around 10% efficiency). To capitalize on the power potential of the PP effect, hi-fi gear is generally operated in Class A at a lower current, higher voltage operating point than SE, where the tubes *barely* conduct through the full 360° of the AC input cycle. The need to conduct substantial plate current for the entire 360° is negated by the PP effect, as each tube of the PP pair must handle only half the waveform (180°) anyway, not the entire 360° waveform as in SE operation. This "low Class A" operation is somewhat more efficient and wastes less power as heat, while still maintaining low distortion. Figure 3 demonstrates the opposing phase input signals, and that the tubes conduct throughout the entire 360° of the waveforms.

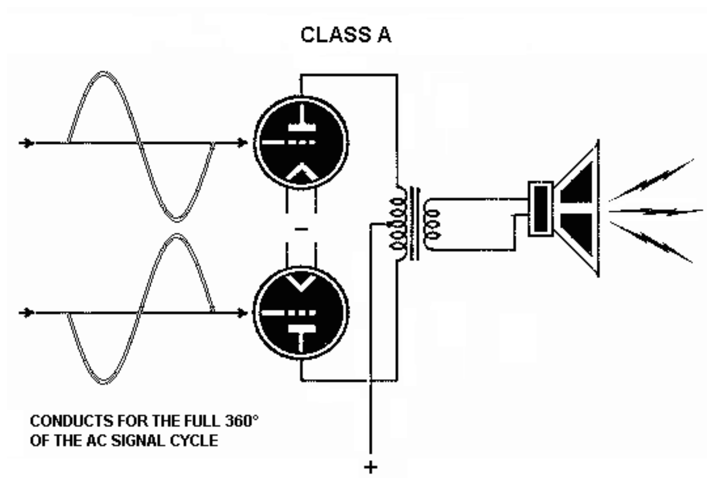


Fig 3 *Power tube conduction of a typical Class A PP output stage*

But if even more power is needed, PP is ran in Class AB operating class to further increase the efficiency. Class AB biases the tubes so that they conduct for less than 360° of the AC music signal cycle, but still more than the necessary 180°, in order to keep them out of the very high distortion region near their cutoff point. Because it idles at generally no more than 70% of maximum dissipation, and conducts full current only on loud peaks at high volume, the tube life is usually longer than with Class A operation. Figure 4 is a representation of the input signals to each power tube of a push pull amplifier in Class AB operation.

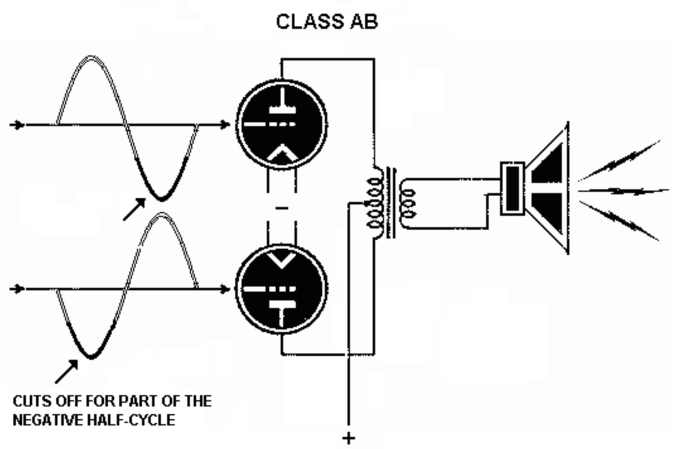


Fig 4 *Power tube conduction of a typical Class AB PP output stage*

Note that in the Class A push-pull amplifier in Figure 3, both tubes conducted throughout the entire 360° of the AC sine wave cycle. In Figure 4's Class AB operation amplifier, the tubes are biased "colder" and conduct for less than 360° but still substantially more than 180°. The black shaded area of the input signal sine wave represents where the tubes cut off on the peak of the negative half-cycle. Because the tubes are wasting less potential power as heat in Class AB than in Class A, the power output is increased substantially. However, due to the lower plate current, they are operating in a less linear region as the postive half-cycle nears the zero (center) line than with Class A. This increases distortion and gives a different sound. This is the conundrum of operating class. When you try for more power, you get increased distortion. When you try to lower the distortion, you lose power.

There are yet other operating classes, but they cannot be used for audio due to their high distortion. Class B operation conducts for exactly half the AC signal cycle (180°), and has extremely high distortion near the center zero line of the AC sine wave. Class C actually conducts for less than 180°. Class B and Class C are used only in RF applications, where the high distortion does not matter. Class AB is so named because it is intermediate between Class A and Class B in function, and behaves in a manner somewhat similar to (but not identical to) Class A at low output. A common urban myth is that "Class AB amps are Class A at low output". This is about as far from truth as it gets. A Class AB amplifier is just as Class AB at idle as it is at full power, and vice versa. As stated, they behave in a manner *similar* to Class A at low output, but they are still running at a Class AB operating point.

Biasing Methods Used in SE and PP

We mentioned bias in our basic primer on SE amplifier theory. Bias is simply making the tube's control grid (where the signal is applied) more negative than the cathode (where the electrons are emitted), in order to set the operating point of the tube. Without it, the plate current flow would be uninhibited and the the tube's internals would melt in mere moments if the circuit's supply voltage was very high at all. Not a happy thing, needless to say. Figure 5 illustrates the two major methods of biasing used in tube amplifiers, grid bias and cathode bias.

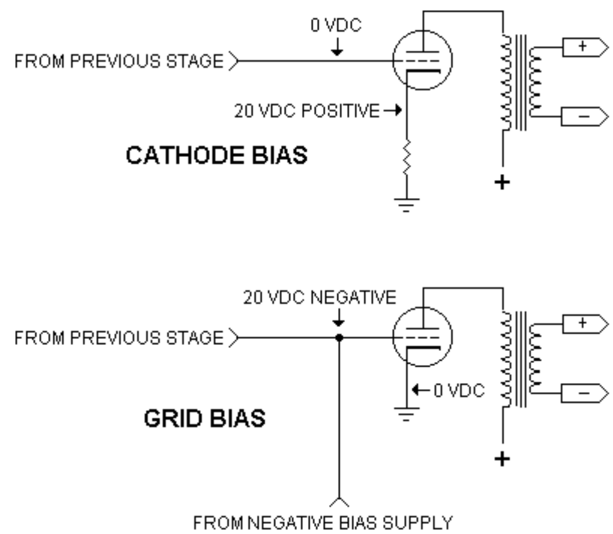


Fig 5 *Methods of biasing a tube amplifier stage*

GRID BIAS – Grid bias (also called fixed or adjustable bias) is achieved by applying the desired negative voltage to the control grid, from a special bias power supply. The cathode emits electrons, which have a negative charge. If the grid is "more negative" than the cathode, it will repel some of the electrons (like charges repel, remember high school science class?), keeping them from leaving the cathode. The higher the negative control grid voltage, the fewer electrons that will be allowed through to the positively charged plate, regardless of the plate's strong, high voltage electrostatic attraction. In Figure 4, we've applied 20 volts negative to the control grid, and the cathode is grounded. The cathode is at ground potential (zero volts), and the grid is 20 volts more negative than the cathode's zero volts, so we have 20 volts of bias.

Grid bias is easily adjusted by varying the applied negative voltage, hence it is often called "adjustable bias." So how can it be adjustable and fixed at the same time? This sounds like a contradiction, but it's really not. It's also called fixed bias because it's fixed to a certain voltage value and does not change regardless of the tube's operation or anything else in the amplifier. It does not mean that it can't be varied if one so desires.

CATHODE BIAS – Cathode bias (also called "self-bias") is a very different method from grid bias, but it accomplishes the same thing. Well, sort of... Grid bias makes the control grid more negative than the cathode. Cathode bias makes the cathode more positive than the control grid, which is exactly the same thing. The grid is still more negative than the cathode, it's just that we achieved it in another way. Remember, in a tube circuit, voltages are relative to *one another* rather than to earth ground. You can see in Figure 5 that the cathode is 20 volts positive with respect to the grid (which has zero volts), so we end up with 20 volts bias.

Cathode bias also differs in that it doesn't use a separate power supply to apply a positive voltage to the cathode, the way that grid bias uses a power supply to give it a negative voltage. The cathode is connected to ground by a resistor of appropriate value, through which it draws it's current. According to Ohm's Law, a positive bias voltage will be created when the tube's current is drawn through the cathode resistor. There you have it, our positive bias voltage is created out of thin air, just like magic, eh? Even neater, the positive bias voltage will regulate itself, unlike grid bias which requires adjustment. If you put in a new tube that draws a little higher than normal current, the increased current draw through the resistor will create a higher bias voltage, shutting down the tube's current draw a bit and forcing it to "chill out." Likewise, a tube with lower current draw will create less bias voltage and allow the gate to open a bit more in compensation.

COMPARISON OF BIAS METHODS – We can see that cathode biasing's self-regulating property tends to prevent or at least alleviate "bias runaway," the condition in which the tube loses bias, conducts wide open, and melts. The more current the tube tries to draw, the higher the bias voltage created to stop it. On the other hand, grid bias can fail and cause the unthinkable to happen. Grid bias failure is caused by a malfunction of the negative bias power supply or the associated componentry that feeds the bias voltage to the tube, or loss of proper contact between the tube socket and the tube's control grid pin due to corrosion, loose contacts, etc.

You're probably asking yourself, "Why on Earth would anyone want to use grid bias, if it requires a separate power supply and can fail, causing a catastrophic, fiery thermonuclear holocaust?" The answer is fivefold.

(1) Grid bias produces more power. All your higher powered push pull amplifiers use grid bias (see below), although it can be used for Class A operation as well, and often is. Cathode bias is less efficient, and wastes some power as heat dissipated in the cathode resistor.

(2) With grid bias, the actual plate voltage of a tube is the measured voltage between the plate and ground. With cathode bias, it's the voltage difference between the plate and *cathode*. Since cathode bias applies a positive voltage to the cathode, the actual plate voltage is the measured positive plate voltage *minus* the positive cathode bias voltage. This means that in many cases, grid bias allows for a lower voltage main power supply than what is needed with cathode bias.

(3) Grid bias *must* be used for Class AB operation. Unlike Class A operation, which draws the same current regardless of signal level, Class AB operation proportionally draws much more current as the signal level increases. The self-regulating property of cathode bias fights against this increase in current. Grid bias works completely independent of everything else in the amp. It's called "fixed" voltage for this reason, remember? This turns out for the best anyway, as the whole reason for using Class AB is for higher power, and as we just read grid bias produces more power than cathode bias anyway.

(4) Grid bias produces lower distortion than cathode bias.

(5) Grid bias allows for ease of experimentation with different plate currents. You can turn the plate current up or down and find the sweet spot that sounds best to you, usually with the ease of turning a trim pot.

Well, now your question has probably been turned around in the opposite direction, and you're asking yourself, "Why on Earth would anyone want to use cathode bias if it wastes power, creates more heat, lacks adjustability, has higher distortion, and sometimes requires a higher voltage power supply? Who cares if it's self adjusting? That's too much compromise for gaining nothing but self-adjustment." The reason is simple. The sound. Cathode bias sounds *very* different from grid bias. Grid bias fans claim their method sounds clean, neutral, uncolored, and firm, while cathode bias sounds soft, mushy, and hopelessly colored. Fans of cathode bias claim their method sounds sweet, warm, musical, and natural, while grid bias sounds harsh, sterile, and artificial.

As we saw, grid bias is always used in Class AB PP amps (by necessity), and in a large percentage of Class A operation PP amps as well. It is even used in some SE amps, but only a few. Most SE amps use cathode bias, which is an essential ingredient of the SET sound. Most people who prefer the PP pentode sound also prefer grid bias, as it tightens and cleans the sound. For what it's worth, the author greatly prefers cathode bias in his hi-fi gear (both SE and PP) and grid bias in his geetar amps. :^)

PP OPT Phase Cancellation

Besides higher output power, PP exercises another huge advantage over SE, which is phase cancellation of certain artifacts. Here's an analogy: If you toss two rocks into a pond at exactly the right time and distance apart so that the crest of one wavefront exactly met the trough of the other, they would cancel each other out. The net result of a crest and a trough together is zero. In other words, the two waves met 180° out of phase with each other, and cancelled each other out. Because of the opposing directions of current flow in the PP OPT, phase cancellation of any "foreign bodies" such as power supply hum and noise occurs. The hum frequency from the power supply enters into the PP OPT's primary winding center taps, flows in opposing directions toward each tube of the push pull pair, and is cancelled by phase opposition. Figure 2 demonstrates this current flow from the center tap outwards. In a SE amplifier, it all goes straight through, and is transferred to the speaker. PP amplifiers therefore need much less power supply filtration to prevent hum that their SE counterparts do. Another benefit of PP is that second order distortion produced in the PP outputs stage is cancelled out in the OPT, which brings us to our next topic:

Harmonic Distortion Series of SE and PP

SE and PP amplifiers naturally produce different harmonic distortion series. Different distortion series sound quite different to the ear. In general, the higher the order, the more offensive to the ear, and in correspondingly smaller amounts with progressively higher orders. Also, odd order harmonics are more offensive up to a point in the series, where it all sounds ghastly horrible in very tiny amounts, whether odd or even order. For example, as much as 3% second harmonic is unnoticed by most people, while .3% (ten times less) fifth order is very noticeable to most people. Second order is manifest as being exactly one octave above the fundamental tone, and therefore it is not dissonant. In small amounts, it is perceived by most ears as adding "sweetess" or "liquidity" to the sound. Third order is a musical fourth above the fundamental tone, which is not terribly harmonious also but not terribly dissonant if held to low enough levels. Fourth order is two octaves above the fundamental, and is not extremely dissonant but blurs focus. Fifth order and up just sounds very dissonant and nasty, period.

A SET (single ended triode) amplifier's main distortion product is second harmonic, which is relatively benign to the ear. The single ended driver stages of a PP amp produce predominantly second order distortion, but the PP output stage itself produces mostly high/odd order distortion. Remember, the second harmonic created in the output stage is cancelled, but any distortion (whether second harmonic or not) produced in earlier stages is passed on through.

The reason why PP and SE output stages produce different distortion series is because they clip differently. This "clipping" occurs when the tube reaches the maximum output attainable at the operating point chosen by the circuit designer. The tops of the sinewave form cannot go any higher (higher amplitude), because the tube is at it's maximum, it cannot amplify any further. Still, on the oscilloscope screen the waveform amplitude tries to grow taller in response the increased input signal. Since it cannot, a flat line appears across the tops of the sine waveform at the "maxed out" point, giving them the appearance of having been clipped off with scissors, hence the name clipping. A PP amplifier clips symmetrically, which means the opposing sides of the sinewave clip alike and at the same time. This produces high/odd order distortion products, which are more offensive to the ear than second order. A SE amplifier typically clips asymmetrically (more on one side of the sinewave), which produces predominantly second order as it's chief harmonic distortion artifact. These different distortion series are one of the predominant reasons as to why PP and SE sound so very different.

Damping Factors of SE and PP

Damping factor is a function of an amplifier's output impedance and how it relates to controlling the loudspeaker, especially at low frequencies. It is expressed as a ratio, calculated by dividing the load impedance by the amplifier's output impedance (Z). For example, if an amplifier has an output Z of 2 ohms and the load is 8 ohms, that gives a DF of 4. This method can be used for estimating the DF, but is not highly accurate. In real life, a speaker's impedance curve is not flat, therefore the actual DF is more or less at certain frequencies. In other words, DF tracks the speaker's impedance curve. When calculating the DF, one must not only figure the output Z of the amplifier's output stage, but must also consider the resistances of the output transformer's secondary winding and the speaker cables as well.

So what does DF do and how does it apply in my system? In general, the higher the DF, the stronger the amplifier's "grip" on the woofer. A low DF makes for tubby, mushy bass with a slow, loose, boomy decay of notes. A high DF makes for better control over the woofer, and yields tighter, drier, faster bass without over-ring and sloppy note decay. A typical SE amplifier using directly heated triodes may have a DF of perhaps 4 or 5 at best. A SE amplifier using triode-strapped tetrodes or pentodes will have an even lower DF, depending on the tube type used. Some have a DF under 2. A vintage PP amplifier from the '60s may have a DF of perhaps 10 or 12. A typical solid state amplifier may have a DF of anywhere from around 100 to over 1000.

MU, TRANSCONDUCTANCE, AND PLATE RESISTANCE – So what does a tube's electrical characteristics have to do with DF? We'll get to that. We saw in the section "HOW SE WORKS" how the plate current and plate voltage of a tube swing up and down, back and forth in a reciprocating fasion in the output stage. The tubes themselves do not amplify, they manipulate the power supply by this voltage/current swing to produce a higher amplitude copy of their input signal across their load. A better term for amplifier would be "modulated power supply" and a better term for an amplifier stage would be "power supply modulator". If there is no load present across the plate of the tube, there will be no amplification. Again, it's *not* the tube itself that amplifies! Back to the voltage/current swing I just mentioned. The plate voltage/current swing occurs in an orderly manner in response to the AC input signal. As we had read in the "HOW SE WORKS" section, on the negative half-cycle of the AC input signal the plate voltage swings high and the plate current low. On the positive half-cycle, the plate voltage swings low and the plate current high. The width of this voltage swing determines the degree of amplification, and the symmetry of the swing determines the orders of harmonic distortion present. It also creates a difference in the DF of SE and PP amplifiers, which we'll see demonstrated shortly.

Let's see how this plate voltage/current swing in response to the control grid signal voltage works to define damping factor. To do that, we first need to look at the three main electrical characteristics that make a given tube type what it is, which are *mu*, *transconductance, and plate resistance*. They are inter-related and inseparably linked. If two of these are known, the unknown third one can easily be calculated.

Mu is the ratio of the change in plate voltage to a change in control grid voltage in the opposite direction, and is usually referred to as the amplification factor of the tube. It's symbol is μ.

Transconductance (represented by it's symbol *Gm*) is defined as the ratio of the change in plate current to a change in control grid voltage. The term "transconductance" is derived from "transfer conductance" and is measured in siemens (S), where one siemens equals one amp of plate current change per one volt of control grid voltage change. It is also commonly expressed as mho, which is equivalent to a micro-siemens (1/1000 siemens). Mho is ohm spelled backwards, which is a very fitting term because conductance and resistance are opposites. A *transconductance amplifier* is an amplifier in which a change in input voltage causes a linear change in output current. Tube amplifiers fall into this category. The other main type of amplifier is known as a *transimpedance amplifier,* where a change in input current causes a linear change in output voltage, exactly the opposite of a transconductance amplifier.

Plate resistance is defined as the change in plate voltage divided by the change in plate current it produces as a result. Plate resistance is expressed in ohms, just like any other resistance. Plate resistance is symbolized by *Rp*.

PLATE CURRENT AND DAMPING FACTOR – We can see how these three parameters are related and interactive. The property of their interrelation we are most concerned with here is how plate current affects plate resistance. Plate resistance drops with an increase in plate current, provided the other variables remain constant. To see how this relates to damping factor, let's build a theoretical SE amplifier and a theoretical PP amplifier, for the purpose of observing how each topology's inherit way of operating affects it's output impedance. Let's use the same triode power tubes in both amps, and run both in hard Class A operation at the same voltage and idle current. Both amplifiers use grid bias, and both their OPTs have the same secondary winding resistance.

Let's power up our SE amplifier first, and give it a signal. Remember, we just read where plate resistance drops with an increase in plate current, provided the other variables remain constant. On the positive half-cycle of the AC input signal, the plate current swings high, and therefore the plate resistance is lower on the positive half-cycle than the negative half-cycle, where the plate current drops. By this, we can see that the plate resistance goes up and down in sync with the polarity swings of the AC input signal. The higher the plate resistance, the higher the output Z, and the lower the DF. Here's where we start getting to the core reasons of why most PP amplifiers have firmer bass than SE amps. The DF flip-flops up and down with each 360° cycle of the AC input signal. This tends to give SE amplifiers an odd bass performance characteristic. The DF is different on the "out" movement of the speaker cone than it is on the "in" movement!

Let's try this same experiment on our PP amplifier. Each tube of the PP pair sees a signal that is in opposite phase to the other. In other words, when one signal is at the peak of it's positive half cycle, the other is at the peak of it's negative half-cycle. As a result, one power tube's plate current is rising while the other's is dropping, so that the two always average out to the same median combined plate current. To better illustrate this, we'll use integers and say that while at idle, both tubes are at zero. At the peaks of the signal cycle, one tube will be at say, +10, while the other is at -10, which averages to zero. Halfway through the cycle, one tube will be at +5 and the other at -5, which still averages out to zero. As a result, while one tube's Rp is dropping as current increases, the other's Rp increases as it's current drops.

Their Rp averages out to a median figure, and therefore the output Z also averages out to a median figure. This means (you guessed it!) the DF is steady as well, very much unlike our SE amplifier where the DF changed with each half-cycle of the input signal. This is what gives PP an inherit advantage in the bass department over SE, all other things remaining the same.

However, we used triodes in our test PP amplifier, and ran them in hard Class A. Triodes (especially low mu triodes) have far lower Rp than tetrodes and pentodes, and therefore will naturally yield a higher DF. But most PP amps use tetrodes or pentodes instead of triodes, which means very high output Z, a very poor DF, and almost non-existent bass. Why then do most PP amps have at least twice the DF of SE amps? The next section explains it all.

Negative Feedback

To further widen the sonic gap between SE and PP, consider that most PP amps use negative feedback (NFB), while *very* few SE amps use NFB nowadays. "Okay, that's nice" you say, "now, what *is* NFB?"

Negative feedback is a sample of the signal taken from the output of an amplifier stage and re-injected into the front of that same stage (local NFB), or into an earlier stage (global NFB), where it is 180° out of phase with the signal present at the re-injection point. In other words, "feeding back." Pretty crafty, eh? The circuitry contained between the two points is called the "feedback loop." Usually, the global feedback loop in a PP hi-fi amplifier will run from the speaker terminal back to the amp's input stage, thereby encompassing the entire amplifier circuit. Figure 6 points out the global feedback loop of a typical push pull amplifier schematic in bright red.

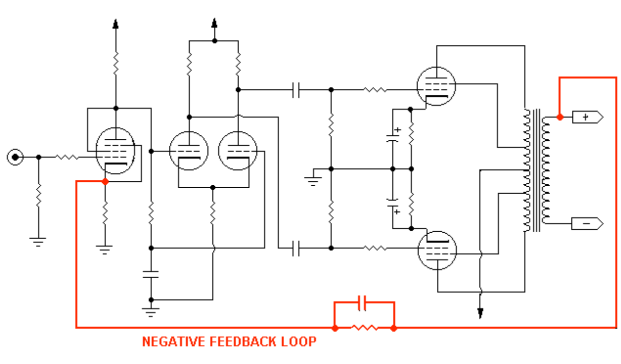


Fig 6 *Global negative feedback loop*

Your next question is probably, "Okay, now I know what it is. What does it do?" Good question. Well, it does a lot of good things.

(1) NFB reduces the distortion and noise in the stages contained inside the feedback loop. It cancels distortions in the same way our PP OPT cancelled power supply hum and second order harmonics earlier.

(2) NFB lowers the output impedance. This tightens and extends the bass.

(3) NFB flattens and extends the overall frequency response.

Most people at this point would be thinking, "Cool! Bring on the negative feedback!" Not so fast. NFB also does some bad things.

(1) NFB kills the space and air between the instruments, parts, and voices, rendering music dull and uninteresting.

(2) NFB reduces the stability of amplifier circuits

(3) NFB reduces gain.

Oh well...... Nothing's perfect, including NFB. *Especially* NFB. This first brings us to the early days of telephone service, where long distance telephone transmission lines were powered by tube amplifiers. Due to parasitic losses in the long, cross-country runs of wire, repeater amplifiers had to be used at intervals. The cumulative distortion of several of these repeater amplifiers was horrid, and telephone communications were difficult to understand as a result. NFB was invented at Bell Labs to lower this distortion, and it indeed worked very well. Without NFB, early long distance telephone communications across entire nations and continents would have been impossible.

Fast forward to the Golden Era of Hi-Fi, which was the original heyday of tube audio back in the 1950s and '60s, before mainstream solid state audio gear even existed. Just as it is in today's "big box store" world of mass consumer audio electronics, specs on paper meant everything back then. The lowest distortion at the highest power was what sold biggest. Lots of NFB was applied to everything, often even in multiple feedback loops. This was done because NFB lowers distortion, and to the uninformed masses the lowest distortion means the best sound. "Lower distortion *has* to sound better, doesn't it? I mean, distortion is a deviation from the original music signal, so lowest distortion means the most "perfect" sound, right?" No, it definitely *doesn't*. We saw this in the last section, "HARMONIC DISTORTION SERIES OF PP AND SE." The amp with the lowest distortion is not necessarily the best sounding. It more strongly depends on *which* orders of harmonic distortion are present, more so than *how much*. The human brain often disagrees with specs on paper as to what sounds best! Because NFB is taken from the very rear of the amp and injected back into the front, there is a frequency-dependent time smearing effect that occurs, which the ear is extremely sensitive to. It blurs the focus and homogenizes the sound. The excessive amounts of NFB that was used to get the best distortion specs on paper was what made many of these amplifiers of yore sound lifeless and uninteresting to listen to, and is a major reason why today's big box store electronics sound so bad as well. Both may have ridiculously low distortion, yet there are amplifiers with literally a hundred times more distortion that sound better. Human psychoacoustics do *not* obey the laws of science very often!

So why do most PP amplifiers use NFB and most SE amplifiers don't? The main reasons are distortion and output impedance. Power triodes and triode-wired power tetrodes/pentodes were not used very much during the Golden Era of Hi-Fi, as a matter of fact almost never. Triodes produce comparatively less power than tetrodes/pentodes, and remember, specs on paper were everything so the most power won the day (or at least the sale!) Practically 99.9% of amplifiers used beam power tetrodes or power pentodes. As we saw in the last section, both of these tube types have very high plate resistance and therefore high output impedance. Neither is as linear as triodes; both types have much higher distortion. A zero feedback amplifier built using these types has high distortion, with much of it being the high/odd order type, plus the very high output impedance means the bass will be weak and flabby. It sounds bright, harsh, and thin, and literally unlistenable. So, NFB was a necessary evil with amplifiers using these tubes, both to lower distortion and yield an acceptable damping factor.

To make matters worse, many of these PP amplifiers had very complex, multi-stage signal paths to improve the specs on paper. This created more distortion, which of course required more NFB to correct! NFB reduces gain, so higher gain driver stages were needed to recoup the lost gain. Higher gain also means more distortion is created than in a lower gain scenario, which means (you guessed it) more NFB is needed to lower the distortion! It's like a dog chasing his tail, a vicious cycle to which there is no end, and nothing good can ever come from it. Plus, the immense complexity of such amplifers also degraded the sound, due to the sheer volume of components, solder joints, and jumper wires the signal had to travel through. Your music got ran through a huge network of sonically degrading components, and it sounded like it.

What's even sadder is that most of today's commercially produced PP amplifiers are still nothing more than tweaked-over copies of these same vintage circuits! Fact is, many (most?) of today's companies producing these PP amplifiers are still caught in this same "specs on paper" rut and can't seem to get out of it. I have this fancy saying about this, "Test equipment ain't got ears". The oscilloscope is not gonna be listening to it. The harmonic distortion analyzer is not gonna be listening to it. So who is? Humans! *We have ears and brains, not screens and displays*. As I had said earlier, psychoacoustics usually doesn't obey numbers written on a piece of paper.

Now that we've beaten the high feedback PP horse to death, let's move back to triodes and single ended operation. Triodes sound less bright, thin, and harsh when not using NFB, and have lower output impedance and lower distortion. A SE amplifier with a single, highly linear driver stage and a low plate resistance triode power tube can usually get by without NFB, which means the average SET amp is usually more open, airy, and purer sounding than an otherwise identical SEP (single ended pentode) amplifier using NFB to correct it's problems of high distortion and output impedance.

In Summary

We've seen that single ended and push pull amplifiers both have advantages and disadvantages. It's difficult to make blanket statements about each, as there are many variations of circuitry and quality in each broad category, so keep in mind these statements are generalizations. There *are* exceptions to the rule

ADVANTAGES OF SE OPERATION

- [x] Less signal deterioration because of fewer signal path components
- [x] Better detail and coherency, especially at low volumes
- [x] More open and airy sound

DISADVANTAGES OF SE OPERATION

- [ ] Poor efficiency, with low power and high heat production
- [ ] Almost no rejection of power supply hum and noise, which necessitates more power supply filtering
- [ ] No cancellation of second order harmonic distortion (some do not consider this a disadvantage)
- [ ] (Usually) shorter power tube life
- [ ] Has speaker compatibility issues, is less tolerant of wide impedance curves and complex crossover networks
- [ ] Poor damping factor

ADVANTAGES OF PP OPERATION

- [x] More power
- [x] (Usually) more authoritative and powerful sounding
- [x] Better bass performance
- [x] Fewer speaker compatibility issues
- [x] Lower distortion

DISADVANTAGES OF PP OPERATION

- [ ] Homogenized, uninteresting sound
- [ ] More signal deterioration due to greater signal path complexity
- [ ] More power supply intermodulations between amplifier stages, and other power supply anomalies
- [ ] Poor low level detail and dynamics

Again, do note that these are generalizations. There are some horrible sounding SE amplifiers out there, and some very open and musical sounding PP amplifiers. It all depends on the design. If we apply SET design concepts to PP, using engineering practices aimed at achieving the best sound rather than the best specs on paper, the results can be astonishingly good. A PP amplifier using low Rp triode power tubes such as the 2A3 or 300B can get by with little or no NFB. For one, the triodes are an advantage in themselves, in both linearity and output impedance. Two, the front end voltage amplifer and phase inverter stages can be designed with much lower gain, if no NFB is present to reduce the available gain. Lower gain means lower distortion and lower output Z, so NFB is not necessary to attain these desirable properties. You end up with a much more linear amplifier, that has a simpler circuit with lower parts count and therefore less signal deterioration, and more air, breath, and vibrancy. See how much better everything gets when you set out to design it right from the beginning, instead of trying to fix it with NFB?

You're probably thinking, "There must be a catch." Well, there is. The power from such an amplifier is very low compared to high feedback PP amps using big pentodes. But, which would you rather have? 15 watts that sound scary good and leave you wanting more of the same, or 70 watts that bore you to death and leave you fatigued? Furthermore, in light of the huge number of highly efficient, great sounding speakers on the market today, amplifier wattage is a moot point. The pitfall to be avoided is not so much SE or PP as it is *high power*. By using very efficient speakers with minimalist or no crossovers, we can sidestep the whole power issue, and end up with much better sound, usually at lower cost as well.

High efficiency and low power is *the* way to go in this author's opinion. I grew up in a house full of vintage tube gear, and have owned both high and low powered tube and solid state gear. I've heard a lot of gear in my lifetime, and quite frankly, most of it isn't worth remembering. Most of what is put out well less than 10 watts. Nothing I've ever heard can compare to a good low powered tube amplifier with highly efficient speakers. SE or PP? Name your poison. Both have their place, and both can be fabulous *if* correctly designed and implemented. We can sum it all up in these words of wisdom, from one of the greatest tube amp transformer designers the world has ever known, Dutch electrical engineer Menno van der Veen:

"In reality, measurements are fine and can help you. But we have the finest measurement devices available on the two sides of our head: our EARS, and are not they fantastic? Well, use them and rely on them."

Eddie Vaughn

Single Ended vs. Push Pull (Continued): The Deep, Dark Secrets of Output Transformers

by Eddie Vaughn

A Burning Question

Here's a question I've been asked by several folks, including a gentleman who has a hearing condition that makes tizzy, grainy systems and loud listening levels painful. "Does the Carina SE EL84 amplifier have good detail and dynamics at low listening volumes, or must it be cranked up so that you can hear everything clearly?" I explained that it sounds very detailed and articulate at sub-watt output levels, and as a matter of fact it's inside the first watt where it truly excels. I immediately followed my statement with a question, "What amplifier are you using now?" It turns out that this particular gentleman has a push pull amp. Ah-ha, just as I'd suspected! While his PP amplifier is different (I'd say better) from most in that it operates it's four 6BM8s in strapped-triode and uses very minimal negative feedback, it still suffers from one of the drawbacks inherit to PP. He says that while it sounds positively wonderful at higher levels, he finds that it lacks detail and dynamics when played at very low volume. His main reservation about buying a Carina was the fear that it will exhibit this same trait.

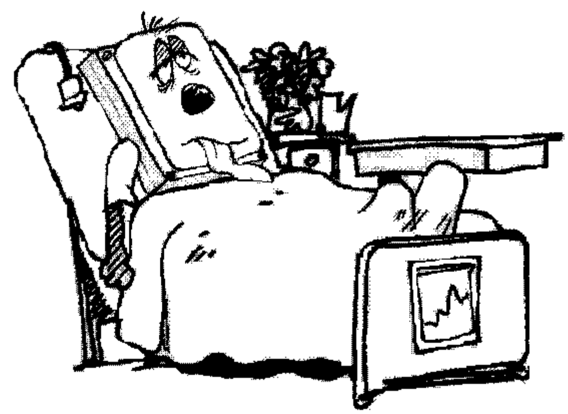
It won't. *Definitely* not.

This is a conundrum that confounds the "spec fanatics." Despite having much poorer distortion specs on paper compared to a push pull amplifier using feedback, a zero feedback SET usually produces better detail, *especially* at low volumes. Some of this lies within the amplifier itself, due to the fact that single ended amps usually have a lower signal path component count and hence less signal degradation, and less power supply intermodulations and such to due the fact there are usually less power supply stages. But to find the REAL reason why, you must go beyond the amplifier section and power supply, and look at.......the output transformer. The SET's output transformer (OPT) also confounds the spec fanatics, because at first glance to the untrained eye it shouldn't work well at all, because of it's air gap. The air gap is a "band-aid fix" that allows it to function with some degree of practicality. Because it reduces the inductive coupling efficiency, one would therefore think it would also only accurately convey the "macro" parts of the sound, and lose a lot of the fine details.

Again, it won't. *Definitely* not.

Neither the low volume fine-detail resolution of SET amplifiers, nor the seemingly inherit inability of PP amplifiers to properly render fine details and microdynamics at low volume are any kinds of secrets. Anyone who has owned both PP and SET amplifiers has observed it. Yet all too often, it gets blamed on the amp being PP. While this does have somewhat to do with it, it has a lot more to do with the *output transformer* than the amplifier itself, as this article on PP OPTs and transformer lamination materials will attempt to explain. We'll address the output transformer differences between PP and SE here, then we'll work towards the front end of the PP amp in consecutive articles. (WARNING: COPIOUS AMOUNTS OF WIGGY-HEADED, MIND NUMBING TECHNO-BABBLE AHEAD)

PP Core Saturation, or "Bass Iindigestion"



I knew I shouldn't have ate that 30 Hz bass note........

No transformer lamination material can achieve a theoretically infinite magnetic flux density. All "saturate" at a certain flux density. Flux density is achieved by applying a magnetomotive force, or *mmf*. The mmf in an OPT is *EMF* (electromotive force), more commonly known to us as *voltage*. Actually, mmf is the magnetic equivalent of electrical voltage! Here, it's the AC music signal from the power tubes that supplies the EMF, or voltage. The saturation point of a given OPT is mostly determined by the core dimensions and the lamination material used. The transformer's magnetic flux density increases proportionally with mmf, which increases proportionally with the applied EMF. Once the core reaches it's saturation flux density, any further increase in mmf does not (in fact, *cannot*) result in any further increase in magnetic flux density. An overloaded, saturated output transformer distorts the waveform in much the same way a saturated tube clips and distorts, creating harmonics in the secondary winding's output signal (NOT good!).

So as we can see, there is a limit to how much AC signal voltage a given OPT (whether PP or SE) can tolerate before core saturation rears it's head and everything falls off a cliff sonically. Actually, it's how much *total energy* is contained in each half-cycle, not just the voltage. A sinusoidal waveform of a given EMF (voltage/amplitude) increases in total energy as the frequency decreases. Why and how? The energy contained in a sinewave is proportional to the total area accumulated in the curve of that sinewave. Since each half-cycle of say, a 40 Hz waveform accumulates more area between it's ends (the zero line of the oscilloscope graph) than an 80 Hz waveform of equal amplitude (voltage) does, the flux density will attain higher values due to the higher mmf energy contained in the 40 Hz waveform. Therefore, the signal amplitude required to saturate a transformer decreases in a linear manner with a drop in frequency, and rises with an increase in frequency. In other words, if a transformer saturates at 20V RMS at 80 Hz, it'll saturate at around only 10V RMS at 40 Hz, and 40V RMS at 160 Hz.

Some transformer builders embellish their low frequency response specs by measuring them at very low signal levels, often only a few hundred milliwatts. In actual listening at substantial volumes, such an OPT will saturate rather easily at bass frequencies and distort. Quite often, an OPT nominally rated for 100 watts continuous RMS will distort at low bass frequencies at well under 25 watts. This problem also exists in power transformers. If a power transformer designed for 60 Hz line voltage is operated at 50 Hz instead, the flux density will reach a higher peak level on each AC half-cycle due to the greater energy contained in the 50 Hz sinewave, thus causing core saturation at a lower secondary current draw even if the voltage of each sinewave is the same.

On the flip side of the frequency band coin, it's nearly impossible to saturate any normal OPT with frequencies from the upper midrange audio band and up due to the low energy contained in the waveform. As a matter of fact, most OPTs have sufficient mutual inductive coupling to work *without a core* from about 2 kHz on up! The sound wouldn't be as coherent, but it *would* work without a core nevertheless.

SE Core Saturation and Air-Gapping

Now that we understand core saturation, let's hit on some basic PP OPT theory. In a PP amplifier's OPT, the power supply is connected to the center tap of the primary winding and a tube is connected to each end of the primary. This allows the tubes to conduct on alternate cycles of the AC input waveform, much like two persons in a boat with one rowing on each side. Since the DC bias currents to each tube flow in opposing directions in each half of the primary winding, they cancel each other out. Remember integers from high school math? Positive 10 plus negative 10 equals zero! However, in a SE amplifier's OPT, the primary has only two connections instead of the PP OPT's three. One goes to the power supply, the other to the plate of the power tube (or tubes if PSE). Therefore, the DC bias current flows *in one direction only*, and the end result is that there is none of the cancellation effect we saw with the PP OPT.

Those who know something of the fundamentals of transformer operation might be thinking, "But aren't we talking about AC voltage here, not DC current? Transformers turn AC voltage to magnetic energy and then back to AC, don't they? They don't pass DC from primary to secondary." This is true, but in the case of the SE OPT, a magnetizing voltage is created by the DC bias current flowing through the primary winding's resistance according to Ohm's Law. To your SE OPT, this DC voltage looks like one half-cycle of 0 Hz AC, if such a thing could actually exist. Low frequency AC was bad enough, but 0 Hz (infinitely low frequency) is *really* bad news for our OPT!

This voltage creates higher mmf, added on top of the mmf resulting from the AC signal.The resulting additional "bias" or "offset" will push the alternating flux waveform closer to saturation on one half-cycle than the other, thereby lowering the amount of AC signal a given OPT can handle before saturation. This offset DC bias effect is more easily understood by looking at in electrical, rather than magnetic terms. Figure 1 is an oscilloscope screen with a symmetrical AC sine wave, which for our demonstration purposes here depicts the magnetic flux created in a PP OPT.

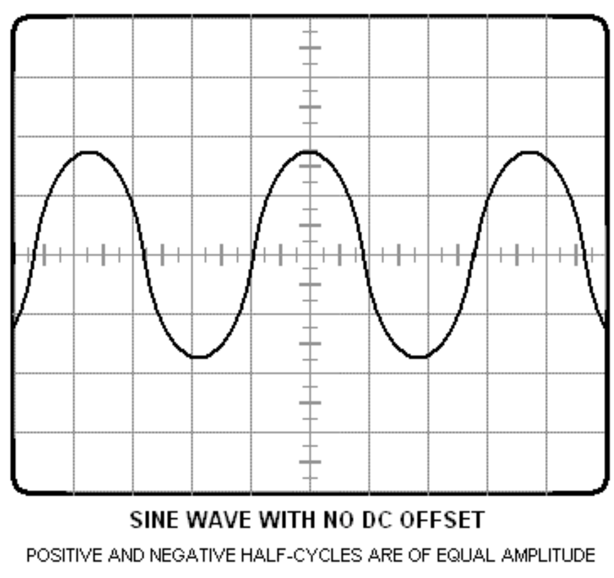


Fig 1 *Sine Wave with no DC offset*

We can see that both the positive and negative half-cycles are of equal amplitude, just as the magnetization in a PP OPT is equal in both directions. But what happens in a SE OPT, where offset DC is present? Figure 2 depicts our sine wave with an offset DC bias added.

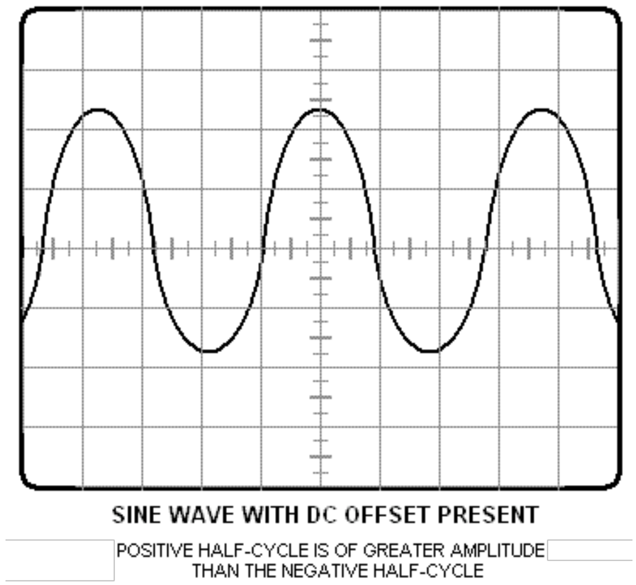


Fig 2 *Sine Wave with positive DC offset*

Notice that it is of greater amplitude on the positive half-cycle, which means that magnetic "headroom" before saturation is decreased in that direction in our SE OPT. Offset DC bias is bad news, to say the least. If you could operate your OPTs at extremely low cryogenic temperatures, the windings would become superconductive and you wouldn't have to worry as much about this condition. But, I for one prefer my listening room at around 72° F. :-) The inescapable reality is that the higher the DC current is in our SE OPT, the closer and closer the core is pushed to it's saturation point flux density.

Therefore you must do something, *anything* to raise the saturation point. Solution: A SE OPT has a shim in it's core between the stacks of E and I shaped laminations, to hold an air gap between them and reduce the inductance by breaking the magnetic "circuit". This helps to stave off early core saturation resulting from the offset DC. But, it also lowers the efficiency, limits bandwidth, and increases distortion and leakage inductance (more on that in a minute). When it comes to output transformers, you can't have your cake and eat it too!

We *could* just forego the air gap and make a bigger transformer, which spreads the mmf out over a larger core area, so that any given cross-sectional area of the core has a lower flux density than with a smaller core. But, this exacts a price in other areas as well. It increases stray magnetic fields in the core (fields not involved in the actual mutual inductive coupling), known as "leakage inductance." This leakage inductance "appears" to be in series with both the primary and secondary windings, so it manifests as a series impedance, reducing the signal level delivered to the speaker. Inductive reactance increases with frequency, so the effect is increased with rising frequency. It therefore rolls off the highs, just as in a speaker crossover.

Eddy currents are another problem. They are random, circular currents (just like eddies in a body of water, hence the name) induced in the lamination materials by the magnetic flux, that cause magnetic "ripples" in the flux field. Eddy currents are a major reason why transformers (and also electric motors) are built with stacked laminations and not solid blocks of metal for cores. The laminations are heat-treated to form a non-conductive oxide film on their faces, which electrically isolates each lamination from the next and prevents the conduction of the eddy currents throughout the core. This also helps to break up the eddy currents by limiting their physical size. The thinner each lamination is, the smaller the area that an eddy current is restricted to. The effect of lamination thickness on eddy currents is demonstrated in Figure 3.

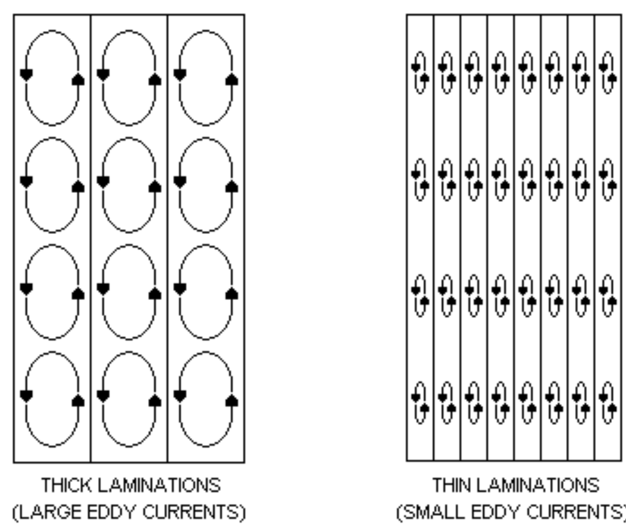


Fig 3 *Eddy currents and lamination thickness*

The eddy current effect increases with rising frequency, so the higher the desired maximum operating frequency, the thinner the laminations must be in order to minimize eddy currents and preserve high frequency coherency and bandwidth. A smaller core mass, thinner laminations, and high quality lamination materials all work to help reduce leakage inductance and eddy currents, and preserve high frequency coherency.

Another unwanted occurrence in inductive devices is "stray" or "parasitic" capacitance. This is capacitance from winding to winding or winding to core. The larger the core, the longer the distance around the circumference of the bobbin on which the wire is wound. The larger the bobbin circumference, the longer each turn of wire is, and the more winding area that exists to develop capacitance. This capacitance varies with signal voltage and frequency, and forms inductive-capacitive (L-C) resonances with the OPT's inductance. Needless to say, this wreaks havoc on our frequency response flatness. Not a happy thing. But, certain special winding topologies are used to reduce stray capacitance, and preserve the frequency response. Properly designed and wound, a good SE OPT can have a substantially large core mass before high frequency response suffers substantially.

The Hidden Glory of the SE OPT

Leakage inductance and stray capacitance are not the exclusive problems of SE OPTs however, they plague PP OPTs as well, and also inductors, power transformers, and autoformers. So if both PP and SE OPTs suffer from these same anomalies, then PP is superior because it doesn't need an air gapped core and is less susceptible to saturation, right?

Wrong. Although the SE OPT's offset DC is a drawback, it's also a blessing in that it keeps the core magnetized. The offset DC creates a static magnetic field that holds the transformer in a highly linear region of it's magnetization curve, even at very low signal levels. It avoids ever seeing a zero flux condition, whereas PP does not. In a PP OPT there is no offset DC, so with no signal present there is no mmf, and no magnetic flux. The PP OPT depends solely on the AC signal to provide the magnetizing or "excitation" current, while a SE OPT *stays magnetized 100% of the time*. It's very important to understand this, because it's a critical reason why SE amplifiers are more coherent and detailed at low volumes than PP! There is no hysteresis or core loss (more on that in the next section). Granted, this comes at a price, as the air gap and necessitated larger core mass do lower the efficiency and create more distortion and frequency response problems. But, it does knock hysteresis in the head.

I've read claims that the air gap is actually beneficial to the sound, and is what causes this phenomenon. This is 100% untrue! The air gap is detrimental in every way, and is a "necessary evil". it would be infinitely preferable to not have an air gap in a SE OPT if that were possible. In spite of escaping the need for an air gap, the PP OPT is faced with another, inescapable handicap in that the two halves of the AC cycle are of opposing polarity. This means the magnetic flux *is in a different direction on each half-cycle*. The PP OPT faces a grievous hurdle here, at the sinewave's zero signal crossover point where the effects of hysteresis take their toll on it.

Hysteresis, the Hysteria Inside Your PP OPT

Hysteresis is an unfortunate phenomena, and also an unavoidable one. Hysteresis is due to the existence of retentivity in the magnetic domains of the lamination material, the very thing that makes certain metals permanently magnetic in the first place. In simplest terms, retentivity can be thought of as residual magnetism, although it will be addressed more in depth later. Magnetic domains are regions of magnetic alignment in ferromagnetic materials, resulting from quantum mechanical interactions at the atomic level. A high degree of magnetization naturally occurs in ferromagnetic lamination materials within individual domains, but in the absence of an external mmf these domains are randomly oriented. An externally applied unidirectional mmf aligns the domains with the direction of the magnetization field it creates. The end result is a large net magnetic flux, a multiplication or amplification of the applied mmf.

Figure 4 depicts a hysteresis loop and it's parts. Examining it will help you to better understand the rest of this section, and how it applies to our OPT.

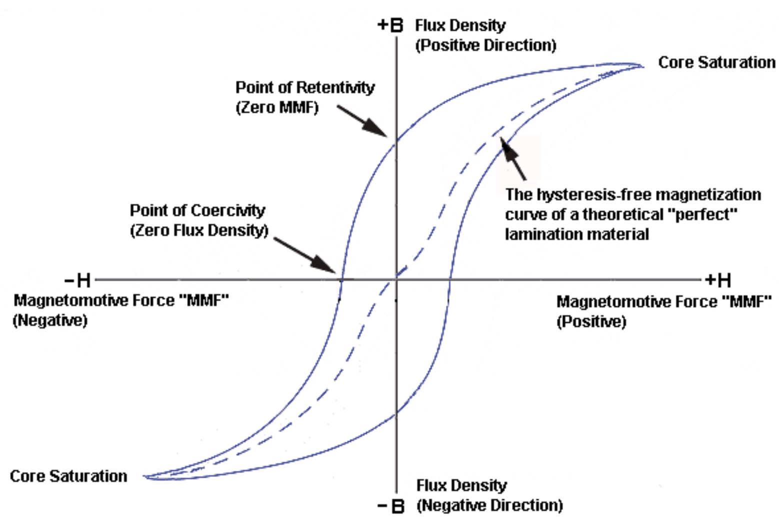


Fig 4 *A typical hysteresis loop, or "B-H loop"*

When the applied mmf in a transformer ceases (the zero voltage line of the AC sinewave), the magnetic flux should theoretically return to zero. However, because of retentivity (the tendency to retain residual magnetism) the flux does not fully relax back to zero when the mmf is removed, due to the fact that some of the domains will still retain their unidirectionally oriented alignment. Just as we read at the beginning that no transformer lamination material can achieve a theoretically infinite magnetic flux density, no ferromagnetic lamination material completely loses it's domain alignment when the mmf is removed, and returns to zero flux density. All exhibit a certain degree of retentivity. This retentivity makes possible the creation of permanent magnets, but in our PP OPT this is highly undesirable! The point at which the mmf ceases (the zero line on the oscilloscope graph) and flux due to only residual magnetism remains is called the *point of retentivity*. The level of residual flux at the point of retentivity is called the *remanence*. It requires considerable energy to overcome the remanence and drive the core to the point of true zero flux, called the *point of coercivity*. In our PP OPT the mmf expended to take the core from the point of retentivity to the point of coercivity is called the *coercive force* or *coercivity*. The coercive force can be considered wasted energy, and this inefficiency is called *core loss*. This is one reason why PP is less articulate, dynamic, and detailed at low levels than SE.

After full negative-direction magnetization has occurred, the cycle reverses again and begins it's positive magnetization swing. However, the point of coercivity is different on the return cycle, because it is skewed by what is now a different retentivity polarity. Because the magnetic polarity of the point of retentivity is opposite on each half-cycle, so is the point of coercivity. Because of this, the magnetization curve is not a uniform, linear line in both directions. This impossibility to precisely track a linear remagnetization curve when the alternating mmf is applied is the property called *hysteresis*. Hysteresis causes the magnetization curve to trace out an S shaped loop, which is where the *hysteresis loop* in Figure 4 gets it's name. The left-to-right width of the hysteresis loop at the center point or *waist* is the degree of hysteresis present. It is also sometimes called the *B-H loop*, where B is the flux density and H is the magnetomotive force, or mmf. All ferromagnetic materials exhibit a certain degree of hysteresis, as the relationship between B and H is never perfectly linear. The hysteresis loop can be simply thought of as a representation of the magnetic flux density (B) as a function of the mmf (H). The more magnetically linear the material is, the narrower the waist of the hysteresis loop.

Lamination Materials and Their Properties

OPTs demand lamination materials with high magnetic *permeability* (the ease with which it is magnetized) and low retentivity and coercivity. The higher the magnetic permeability, the lower the mmf that is required to attain a given flux density, and oftentimes a smaller core mass can be used as a result, along with it's benefits of lower leakage inductance and stray winding capacitance. If saturation becomes an issue, core size can be increased accordingly, and special winding techniques used to help overcome the increased stray capacitance that usually results from the use of larger core mass. The lower the retentivity of the lamination material, then usually less coercive force that is required to overcome it and remagnetize the steel in the opposite direction. So, the higher the magnetic permeability and the lower the retentivity, the better that details and microdynamics will be rendered at lower signal levels.

Such high quality lamination materials are very expensive. GOSS (grain oriented silicon steel) is one of the most commonly used lamination materials for transformers and electric motors, because of it's desirable magnetic properties. Grain orientation lowers core losses and extends the saturation point, and about 4 to 4½ percent silicon lowers the remanence (retentivity). GOSS laminations are annealed at about 1500° F (820° C) before use. Annealing relaxes the internal stresses caused by rolling out the sheet and then stamping the E and I shapes from it, and forms a non-conductive oxide film to reduce eddy currents. M6 is the grade of GOSS most often used in high quality power transformers and PP OPTs.

M19 is a non-oriented steel commonly used in filter inductors, low-end power transformers for tube amps, and in some SE OPTs. It has lower magnetic permeability than M6, and requires greater core mass per volt-amp of output in power transformers. It requires more energy for magnetization reversal, and has a wider, less linear hysteresis loop. Because of this, it is less efficient in power transformers than M6. Due to the fact that more of the AC sinewave's energy is wasted as heat in reversing the flux direction than with M6, power transformers built from M19 exhibit higher heat rise due to the higher core losses. Because of it's inferior properties, it should never be used in PP output transformers.

Some very high-end PP OPTs use exotic materials such as M4 GOSS, which has superior properties to M6 but is very expensive. Amorphous iron is another exotic material used in very high-end OPTs. It has an amorphous molecular structure similar to that found in glass. It has low retentivity/coercivity and high permeability, and is therefore extremely efficient, with a very narrow hysteresis loop. However, it saturates at a lower flux density than many materials, especially at low frequencies.

Nickel is added to iron to improve the permeability and lower the core losses. It's drawback is a far lower saturation point than GOSS. M6 typically saturates at around 20 kilogauss/2 Tesla, while iron with only 50% nickel content saturates at about 15 kilogauss/1.5 Tesla. Because of it's extremely low core loss, nickel iron is one of the finest materials available for transformers. 80% nickel iron was originally christened Permalloy back in 1914, when Gustav Elmen of Bell Laboratories discovered it's superb properties, but today these alloys may contain different nickel contents. The premium transformer grade Permalloy consists of 79% nickel, 17% iron, and 4% molybdenum. It has the highest permeability and lowest core loss of any common transformer core material, which makes it a very sensitive and low distortion lamination material. It's low saturation flux density of 8 kilogauss/.8 Tesla and extremely high cost limits it to certain specialty applications, such as recording studio microphone transformers, step-up transformers for low output, moving coil phono cartridges, and other critical small signal applications. It is also used as record head cores in the finest tape recorders.

79% - 80% nickel Permalloy is also known by the names Supermalloy, Mu-Metal, Ultraperm, and it's AISI/SAE classification, M-1040. 36% nickel Permalloy is commonly known as Radiometal, and 48% to 50% nickel is called SuperRadiometal. Permeability falls as the nickel content falls, and the saturation flux density point rises. 80% nickel Permalloy has a saturation flux density of only .75 Tesla, 79% nickel saturates at .8 Tesla, 50% nickel saturates at 1.5 Tesla, and 48% nickel SuperRadiometal has a saturation flux density of 1.6 Tesla. So, the transformer engineer must carefully consider his application and choose the appropriate nickel content to avoid core saturation. Nickel iron laminations must be correctly annealed to ensure the best magnetic properties, and care must be taken to not expose them to high impact after annealing. Hard, high-G impacts (such as dropping a completed transformer on a hard floor) can change the magnetic properties of nickel iron, especially high-nickel content alloys.

High-cobalt steels such are sometimes used in certain specialty applications. They can attain very high flux densities before saturating, higher than even M4 or M6, but have higher retentivity and therefore higher core losses. Were it not for their extremely high cost, cobalt alloy steels would be very useful for power supply filter inductors. Size for size, a cobalt steel inductor is capable of handling much higher DC currents before saturation than the commonly used M19 or M6 materials.

Because SE OPTs largely avoid the hysteresis effect from maintaining magnetization in a linear region, their construction permits lamination materials that could never be used in PP OPTs, such as M19. However, most high quality SE OPTs are made of M6 nonetheless. Because SE OPTs suffer less from core losses than PP OPTs, cobalt steels are sometimes used in certain SE OPT applications.

Torodial Output Transformers

Although toroidal power transformers are very common, toroidal audio transformers are not seen much, due to their inherit problem of high intolerance to offset DC. They will saturate very quickly because their flux density level typically is already higher than that of an E-I lamination transformer, and there is no air gap in their core. However, a few companies have "worked out the bugs" somewhat with winding topologies and core geometries that they are very secretive about. Despite their lower tolerance to offset DC bias, toroidal OPTs enjoy several specific advantages over E-I OPTs.

Toroidal transformer cores are made by winding a continuous strip of the core material into a donut shape, called a *toroid*. Because of the way the core is made, 100% of the grain is oriented in the same magnetic direction. This eliminates one of the major disadvantages of E-I lamination transformers, in which only about 60% of the grain is oriented in the same magnetic direction. Although the grain of a stamped-out lamination is oriented in one direction only, the magnetic flux in an E-I transformer runs in all directions, and is asymmetrical in different directions as well, as shown in Figure 5.

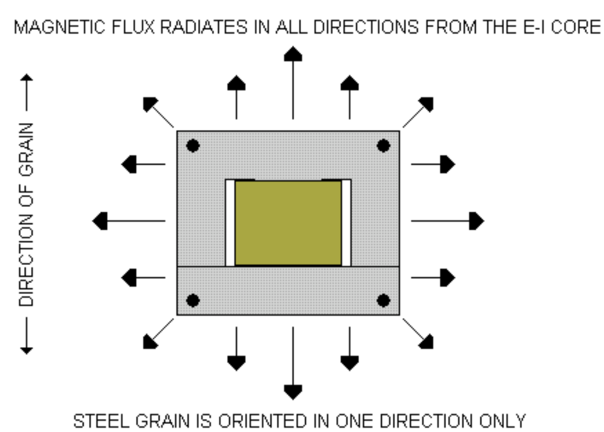


Fig 5 *Grain orientation direction versus magnetic flux direction in an E-I core*

The E-I core's magnetic flux is omnidirectional, but the lamination steel's grain is unidirectional, so much of the core is "out of sync" with the flux direction. In a toroidal transformer, the flux is symmetrical about the core, in one direction only. In other words, any given cross-section of an E-I transformer sees a different flux direction and density, where any given cross-section of a toroidal transformer sees the same flux direction and density. This gives toroidals very low leakage inductance.

The windings are then evenly wound directly on the symmetrical, gapless, round toroidal core, without a bobbin. Because the windings are symmetrical and even, the inter-winding capacitance can be very high unless special winding techniques are employed to reduce it. However, because the toroidal's windings and flux field are symmetrical, several advantages are realized. The flux density is much higher than in an E-I lamination transformer, and every turn of wire sees the same flux density. As a result, the level of mmf necessary to establish and maintain the minimum flux density required for coherent, full bandwidth mutual inductive coupling is very low. This minimal mmf, called the "quiescient" or "excitation" power, is typically over a dozen times lower for toroidals than E-I transformers. Not only are toroidal OPTs very efficient (typically over 95%) , they are ideal for push pull applications where little or no offset DC is present to cause core saturation. The toroidal OPT's low hysteresis and excitation mmf makes for much better detail and dynamic response at low volumes.

C-Core Output Transformers

C-cores are made in a manner similar to toroid cores, by winding a strip of material, usually M6, into a shape as shown below in Figure 5. It is bonded together and annealed just like a toroid core, but is then cut into two identical C shaped halves. The cut faces are precision ground exactly square and parallel, as it is of the utmost importance that they fit together perfectly, or a large loss of efficiency will result. The wound bobbins are placed over the legs of one core half, and the other core half inserted into the bobbin until their faces seat together. For SE operation, an air gap shim is added before assembly. Then, the core is banded tightly around it's circumference with a steel strap to ensure tight contact of the core halves.

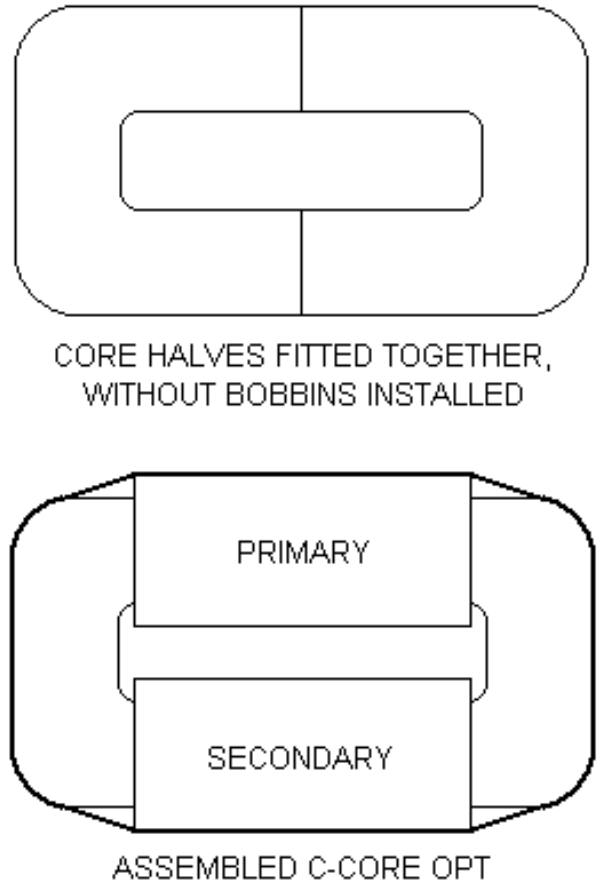


Fig 6 *C-core transformer construction*

Like toroidal OPTs, C-core OPTs are not highly common in today's tube audio world. However, like toroids, they do exhibit certain advantages over the much more common E-I lamination transformer. They are more efficient than E-I transformers, though not as efficient as toroidal designs. The bobbins can be wound on conventional equipment, unlike toroidal transformers. Because the primary and secondary are wound on separate bobbins and reside on different legs of the C-core, they are physically separated from one another and their inter-winding capacitance is practically nonexistant. The main disadvantage of the C-core topology is that it's leakage inductance is quite high.

To Parafeed or Not To Parafeed, *That* is the Question

Parafeed (parallel feed) is worth mentioning here in closing, while we're on the topic of OPTs. In a parafeed topology, the power tube is loaded by an inductor, resistor, or active constant current source, and a DC blocking capacitor couples it to the OPT. It has became very trendy due to several reasons. It has far better immunity to power supply ripple and noise (higher PSRR, or Power Supply Rejection Ratio), because the power supply is not routed through the OPT as in conventional SE output stages. The output stage's AC ground is the OPT primary directly to ground versus back through the power supply, which nets a lower impedance. There is no offset DC present in the OPT, because it is blocked by the "parafeed capacitor". In the absence of DC bias current, the OPT can be made much smaller, and without the air gap required by conventional SE OPTs. Due to the small core mass, leakage inductance and stray capacitance are very low. Manufacturers of parafeed amplifiers tout the parafeed topology as the be all, end all of output topologies, with the best specs and best sound. However, they do not tell you the whole story.....

It's worth noting here that although they are single ended, *parafeed OPTs suffer from hysteresis in the same manner as PP OPTs do*. This is of course due to the fact that no offset DC is present to hold a linear static magnetization. But, a high permeability, low core loss material such as Permalloy can be used to great effect in the parafeed OPT to minimize hysteresis effects. However, there is no way to avoid the other problems inherit to the parafeed topology cause by the DC blocking capacitor, such as frequency-dependent phase shifts, a charge/discharge curve that is non-linear with frequency, and losses caused by the capacitance being in series with the OPT's inductance. Although parafeed can sound incredible if done correctly, nothing is perfect. Especially transformers ...

Eddie Vaughn