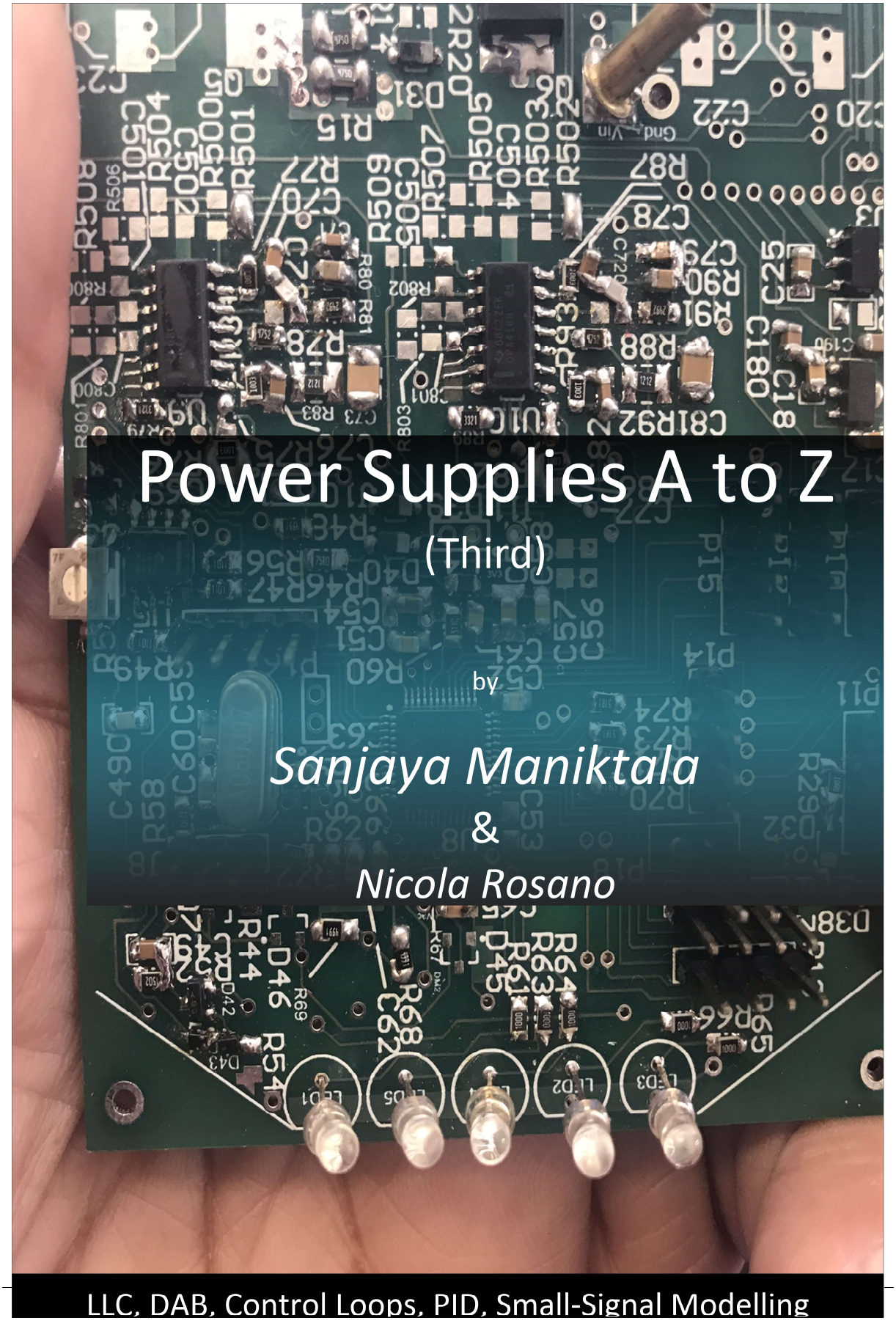


Power Supplies A to Z

(Third)

by

Sanjaya Maniktala

&

Nicola Rosano

LLC, DAB, Control Loops, PID, Small-Signal Modelling

This book invokes hidden-in-plain sight Scaling Laws, discovered by Sanjaya, further popularized by Nicola, to easily design wide-input, high-power, LLC converters (hinting also at the right way to design inductive wireless power transfer systems). , Also included here is the extremely tricky but powerful Dual Active Bridge (DAB), with several intimidating degrees of freedom, which engineers typically struggle with. Simple curves unveil a very simple way of designing them too. Also included is guest chapter by Nicola, on small-signal modelling, where he goes on to show how his low-level models can produce results matching Sanjaya's curves/calculations in A to Z Second Edition. In the last two chapters, Sanjaya focuses on truly understanding analog control loop theory first, then reveals an overlooked contributor to the actual step load response. He calls it the "Q-mismatch issue." He then creates a startling analogy between the impedance of a capacitor, and "PID" coefficients, to eliminate this mismatch and produce the best step load response possible, in a few easy steps.

Power Supplies A to Z (Third)

by

Sanjaya Maniktala

and

Nicola Rosano

ISBN: 9798588905015

Preface

These chapters were to have been the additional pages of a previous book, to make it the third edition. Unfortunately I felt the surprisingly short-sighted publisher handed me a young arrogant (IMHO) program manager from UK, who really spoiled my mood from Day 1! I felt I knew the production process well, having written three (very well-known) books through them… I was not in the mood of "proving" my writing capability to him all over again, by providing "sample chapters" etc, by designated "Date X", "Date Y"… etc…. In fact, I already had almost all the pages presented herein written out by the first date in my contract, but the guy was just waiting endlessly only for his "sample chapter", to first kindly "approve" my writing (hopefully not technical) skills, before giving me even basic dropbox privileges to upload the very same pages presented here… Yes, they never got to see most of the material as a result, but you can now. I walked out of the contract. Publishers seem to think they "make" authors.

In my personal life, things seem to have become very difficult/busy for several reasons, so I needed to get this out to the world quickly… in particular to acknowledge Nicola Rosano… the guy who caught a string of mistakes in A to Z, Second Edition (Note: I accurately call them mistakes, not "typos"). He also did a great job in popularizing my scaling laws, first publicly revealed by me in an IEEE PELS SFBAC seminar (on youtube). But that was in pre COVID-19 times… Nicola had to do a (very well-attended) webinar instead. But the type of guy he is, he always goes out of his way to acknowledge me. He is super at that! I do that too always (e.g. my constant acknowledgements of my mentor Dr GT Murthy in India). Yes, Nicola talked about my discovery of current ripple ratio, and my discovery of "hidden-in-plain-sight" scaling laws. He also kindly provided a superb guest chapter here.

In Chapter 5 you will find another hitherto unnoticed discovery of mine, dating back to 2015. Sept 17 to be precise. I call it the "Q-mismatch" issue. See how I solved it and overnight doubled the efficacy of a vendor's vaunted programmable "PID" controllers. It was an issue apparently overlooked by all control loop experts I still feel, and my solution was just a few lines of simple calculations. Spoiler alert: I compared PID coefficients to a simple capacitor! Resulting in spectacular (200%) improvement in loop response right off the bat, over just one afternoon in San Jose.

Many thanks too, to Gregg Clark, Clyde McKay, Christina Lu, Alex Soriano, Achim Doebler, and Sridhar Vallepalli in particular, for always standing behind me! Not to forget dear Ted Tewksbury for valuable, honest insights always.

Also, Eric Wen, a very capable English to Chinese translator of some of my books, is hereby free to translate this as he wishes, perhaps with China Machine Press… I welcome that too. He was a live translator accompanying me for the whirlwind 5-city Sanjaya2016 tour of China. Always gave great inputs too. As did Sesha Panguluri, a dear friend/colleague of mine over the years, who I really enjoy thinking along with, and even relearning power all the time…. He is a master at firmware, but very adept at understanding and discussing power too. Hence his contribution to my basic way of thinking too, that I must acknowledge here finally.

With that, all I will say is enjoy the book, hope you remember me for this too. And thanks for your overwhelming support of my previous books. Over the years. My reward was: you liked it! That is all I ever cared for. Besides my two beloved Maltipoos: Munchi and Cookie.

Sanjaya Maniktala
Dec 31, 2020

P.S. I suggest you print it out first since the figures may be all at the end of each chapter…. (sigh) if only I had the time to organize it better….! I just didn't. Also see the hints on how the entire world seems to have got it wrong in wireless power transfer (Inductive? Resonant? Ignorance!)

Intentionally Blank

Chapter 1
Resonant Power: LC to WPT

Introduction

We will start with a rather abstract discussion before we delve into finer details. Because resonant power is tricky—to state it mildly! The devil is *not* in the details! Not initially at least. One can easily mistake the forest for the trees, getting utterly lost and not even realize it. We recommend a *soft* transition from the relative comfort of our familiar world of classical ("PWM"-based) power, into the yet-nonintuitive world of resonant power, including wireless power transfer (WPT). Preferably with a stopover at the "LLC topology", for reasons which will become clear. The ultimate aim should be to cultivate and "fine-tune" a completely new type of reasoning—to deal with something, well…*completely new*. At some point, we will realize that our prized "prior "experience", is now merely excess baggage. We need to disband it quickly. To consciously "unlearn" much of what undoubtedly served us well over the years. Because first and foremost, we need to know, that we really *don't*!

For example: "*Lower the frequency to reduce switching losses and improve efficiency*". Or "*Minimize losses by reducing the leakage inductance*". Sounds convincing! Except that neither of these statements are valid when it comes to resonant topologies, implemented "correctly" of course. In the strange new world ahead, there are almost no switching losses. And the efficiency actually *worsens* if we *lower* the frequency! Why? Because, the conduction losses, *increase* when the frequency is *lowered*, not stay fixed, as we've always tended to assume—a sort of inverse switching loss. And energy residing in the "notorious" leakage inductance too, is almost fully recoverable. In fact, it facilitates *further* improvement in efficiency, by crowbarring out electrostatic stored energy buried in say, EMI-suppression capacitors placed on switching nodes, and cycles them back to the input, for convenient reuse! And *no* energy whatsoever is lost "mysteriously" *between* the coils either. It is all accounted for. Most surprisingly perhaps, it can also be shown that *low* coupling, i.e. high leakage inductance, actually enables *greater* power, thrown over *larger* distances in wireless power transfer ("WPT"), than tightly coupled coils. Because the rectified DC voltage on the Secondary (receiver) side, tends to reflect via basic "transformer action" back onto the Primary (transmitter) side, clamping the voltage across a certain *part* of its coil, thereby dictating the dI/dt in it, indirectly limiting the maximum power that can be drawn by the system. So, less coupling, less clamping, more dI/dt, more power. This is a world we were totally unprepared for perhaps…!

All the above advantages however accrue only *within certain definable limits*. But not necessarily, "hard" limits that we are familiar or comfortable with. Nature is bounded in its own responses—to keep *itself* in control too! Hot air rises, to allow cold air to rush in, to cool the hot spot. Summer gives way to winter, which eventually hands us over, back once again to summer, and so on. That is nature! Self-sustaining. We need to respect its processes, and boundaries. And if we want to go further and actually exploit the awesome power that resonance in particular, brings to our doorstep, quite literally, we must also realize it comes to us with a hefty price tag attached: it *demands total design expertise*. Certainly, not a convincing set of alibis…. Such as, when things didn't go as planned, instead of questioning our basic assumptions, we collectively shrugged it off as "inherently poor user experience of *inductive* (wireless) power systems". What if it was merely poor *designer* experience? Undeterred, we went on to declare: "*resonant* systems on the other hand, have inherently good user experience". Resonant being defined as "receiver and transmitter both tuned to *exactly* the same frequency". Did we forget that even in our "inductive systems" we were doing exactly the same thing already? Or at least intended to do so! Actually, in either case, that "double tuning" was always an erroneous concept to start with, as we will explain. Based on some radio-frequency (RF) gut-instinct of ours from a previous life. Completely inapplicable to coupled systems such as those we are dealing with here.

It got worse: because even accepting the "resonant-inductive" dichotomy for argument's sake, one might wonder what *else* was even left to "resonate" further with, other than the receiver (Rx) and transmitter (Tx) of course, for us to announce to the world an upcoming, "*resonant extension to the (inductive) Qi standard*"? Wasn't that to have been AirPower, by the way? What went so "unexpectedly" wrong?

Two wrongs never made a right, certainly not three. But what really seems to have happened historically, is that engineers working in classical power conversion, or RF electronics, maybe just Bluetooth, headed into a brand new, and very tricky, area called WPT, rather *unsuspectingly*. And much too *directly: t*hey did not have the luxury of that fortunate chance encounter with the *big brother of WPT: the LLC topology*, which this author had. Otherwise, they too might have realized somewhere along the way, that if the relatively better-understood LLC converter today, were to be designed the same way as WPT systems of today are, there would not be a single satisfactorily working LLC converter left on the planet either. Having set our minds to *really* understanding resonance, preferably via the LLC converter, we should be forewarned that it can still all seem very intimidating at first. There is hope though: things can get better quickly if we follow the process of correlating our developing "mental picture" to actual simulations, based on detailed math spreadsheets, then building something however small, and finally doing sanity checks, using lab data—a process of triangulation which this author recommends, as the one he has always followed: MATHCAD →SIMPLIS →LAB →SIMPLIS →MATHCAD→SIMPLIS →LAB, and so on. Not for the faint-at-heart. Certainly not for the gung-ho.

Maybe, we will finally see a pattern emerge too! The big picture! Because acquired mastery, if any, is ultimately best measured, by how effortlessly *we can communicate it back, even* to the relative novice on occasion. *Simplifying, what is perhaps benumbing complexity*!

That may become evident in the *powerful, hitherto "hidden-in-plain-sight" power and frequency scaling laws* that we will reveal here. Some of those were first discovered and published in Chapter 19 of the book, Switching Power Supply Design and Optimization, Second Edition edition, by McGraw-Hill, in 2014. In this chapter we will go further: we will add to those laws, and then *exploit them unabashedly*—to unveil an astonishingly simple, *barely half-page design process*. But one that is as thorough as can be. And it can be used right away for designing *any* LLC-type resonant converter or WPT system: any power, any frequency, any output voltage, any input voltage, in fact tolerant to *input variations* too. Oh, we forgot to mention: any coupling too!

A varying input voltage, incidentally, was supposed to be the bane, if not death-knell, of LLC/resonant converters. Not so anymore, we will be pleased to discover. No longer do LLC converters have to remain fearfully confined to that familiar, comfortable location just *after* a power factor correction (PFC) stage. From now on, LLC can be front-end, (and center-stage too!). And we will also be able to, design WPT systems—*correctly* for a change.

As a sneak peek, here are the scaling patterns that we will be using (bullet "c" below, being the brand-new member of our scaling family):

a) To double the power, halve the inductance and double the capacitance
b) To double the frequency, halve the inductance and halve the capacitance
c) To quadruple the power, double the input voltage

Instead of "double" or "halve", we can use other scaling factors to generalize quite obviously. The trend is obvious.

So, basically, in a few simple steps, none even involving the non-intuitive j=√(-1) for a change (!), we can scale an almost randomly chosen, but previously carefully-studied "kernel", to virtually any power level and frequency range we want! That is the true power of scaling, in movies, called "cutting to the chase"! And though it looks too good to be true at times, it has a happy ending: it is as accurate as can be. As can be confirmed very easily by an actual build.
It is important to keep in mind that if we change the power, we do not change the frequencies involved, since those depend on the product of L and C, and we divided L by the power scaling factor and also multiplied C by the same factor, so the LC product remains unchanged. Similarly when we scale frequency, we divide both L and C by the same frequency scaling factor, keeping the ratio C/L unchanged, and that is what power depends on. So the real beauty of the first two scaling laws is that they are "orthogonal". They don't interfere with each other. That is why this discovery is so fundamental and powerful.

Resonance Lessons back to Classical Power

And, though we usually recommend, not to believe anything that you may have read in Switching Power Supplies A to Z, Second Edition for example, resonance being a new frontier, the above scaling laws apply equally well to classical power.

Take a Buck, working well at 5V, 1A @ 100kHz. Keeping everything else the same, if your Boss asks you to double the power overnight, to 5V, 2A, you need to simply double the output capacitor (to keep the same output ripple, as that is proportional to load current, but inversely proportional to capacitance), and halve the inductance. Of course, though you halve the inductance, its size doubles, since size depends on ½ LI^2, and current has doubled too. Similarly, the size of a capacitor depends on ½ CV^2. Here, voltages were kept constant, so size of the capacitor will double. Size scales linearly with power, as we intuitively expected! The 5V, 2A converter will not be unstable either, because the crossover response depends on the LC break frequency (plant), which has not changed. Note also, that the current ripple ratio, *r*, remains the same, since ΔI has doubled in the inductor as we doubled I, since we halved L. It all fits well! We have the same current ripple ratio so the design is still optimum.
Similarly, if your Boss now demands you take the 5V, 2A @100kHz converter overnight to 200kHz, all you need to do is halve both the inductance (and size) of the L and C. In doing so, the LC break frequency which depends on the square root of LC, will also double, exactly how we want it to behave as a function of switching frequency. So, we will likely still be stable, but may need to scale the capacitances involved in the compensation circuitry. Indeed, we can just halve all of them, and things should be fine. Current ripple ratio has also remained the same here.

WPT is (or should have been) the LLC

Any perceived, or "physical", difference, between an LLC converter and a WPT system, is not what may seem "obvious" to our naked eye—such as the distance we are throwing, or transmitting, energy across, or in the shape of the windings/coils used. "Coil" or "winding", whatever we may call it, irrespective of how it looks, just becomes an "L" in our final electrical schematic. Similarly, distance or separation, whether mm or a few meters, just ends up as a different coupling coefficient ("K") in our math spreadsheet. LLC and WPT are really not much different, *electrically*.

The biggest difference between the two is actually the scariest one! The one you "see" but don't *observe*. Or "know" but don't comprehend. In an LLC transformer, the windings are fixed relative to each other (i.e. a constant coupling), but in a WPT system, the coupling coefficient ("K"), can change dramatically with varying placement, or different alignments, of the receiver (Rx) and transmitter (Tx) coils with respect to each other. The constantly varying K changes *everything.* The overall result is hard to predict! But it can be. In fact, it has not only been fully predicted, but solved and harnessed too. Those are however proprietary solutions which cannot be revealed. But a lot of hints will emerge as we slowly unravel the grand resonant puzzle, using the simple, graphical aids presented herein.

Design Goals

The graphical procedure we present, throws light on what was considered to be the most baffling piece of the resonant puzzle: *what is the correct, or most optimal L-C component selection—to handle a certain power, over a desired input range?* You would think that others in the field of resonant power would be very happy to know what those values are. Not really, because they likely don't even realize this is in fact the biggest challenge in resonant power. Because in classical power, "power capability" of the inductor was not the basic, overwhelming concern ever, for a designer. A simple buck regulator for example, has the simple "DC transfer function" $V_O = D \times V_{IN}$. It is implicitly *load independent*, at least in the first order. We can draw any load current *theoretically*! It is essentially "current unlimited"! And we could use almost any inductance too (on paper), and that simple equation was essentially still valid. No obvious "power capability" limits. If any, they were determined by parasitics. Indeed, in practice, there is a power limit, but mainly determined by *second-order* effects. Such as those arising from *parasitics*, like DC resistance of the inductor ("DCR"), current capability of the switch, its set "safe" current limit, the current at which the inductor may saturate, or get too hot, and so on.

But all that changes dramatically in resonant power, and you may never even realize it. The current Qi standard has approved a wide range of L's and C's, without a thought. All for the same target power level Yes, they "work", sort of, but were the component selections optimal? A hint here: every LC resonant curve has a certain inherent "bulge" or "peaking" in relationship to the R present. So, you might think you have a certain gain available, but the moment you change R, and try to draw more power, the shape suddenly shapes. Try walking across a carpet that sags as you walk over it, or even is moved sideways. We will see that the gain curve we thought we had, not only "flattens", essentially preventing us from drawing more and more power, but also shifts sideways, quite dramatically in some cases, not only making our targeted power possibly out of reach, but throwing any proposed control loop in doubt too….we may have thought that to increase power we need to lower the frequency, but now since the peak may have shifted to the "other side", now we may have to reverse our entire "direction of correction", or "DoC"….without realizing it, leave aside knowing how to do it. Basically, resonance and the awesome power it brings to us, is inherently self-limiting. That is what makes all the powerful forces of nature, stay bounded and self-sustaining.
A hint to our strategy to handle input droops too: to handle a definable droop in the input voltage, we need to *overdesign* the peak power capability at the maximum input voltage point! A bit more of a "bulge" But only a certain *definable* amount—to avoid straying into full-blown overdesign territory, which will impact cost and efficiency. *Careful* overdesign is our goal here.

The graphical aids not only point to the correct selection of the L-C components, but the best operating points too: *to do a meaningful "worst-case" system-level simulation*—for computing RMS currents, and so on. Because one of the biggest, unsaid problems facing all the simulation-fans out there, is that in resonant power, it is sadly no longer as simple as:—do the "min-load", "max-load", "min-Vin" and "max-Vin" corners! That was just child's play! In resonance, we are at the doorstep of an almost infinite array of possibilities to investigate. Can't be done without infinite simulation time! So, we desperately need to know precisely, *beforehand*, the best L-C values to use, and even the target operating frequency to set! At least somewhere very close to it—in the ball-park. By providing those in this chapter, we will finally be enabling meaningful, quick simulations too.

Simulations are in fact strongly recommended in resonant power for another reason. Because as we will see, all closed form equations in literature, base themselves on something called the "first harmonic approximation" ("FHA"). You may ask: *how approximate*? Hmmm, that depends! So, it is not an exact science, is it? Well, it is *tricky, as we had warned!* In this chapter therefore, we will focus not on creating accurate or impressive "equations", but more on unearthing the trends and pitfalls of resonance, thereby homing in on the best points for conducting actual simulations, so as to provide accurate reliable numerical estimates eventually, not approximations based on FHA. That should eventually lead to a more reliable product, designed in a predictable time-frame!

But a warning here: a lot of engineers likely did not understand WPT because they inadvertently used a certain *default transformer model* that is provided in most electrical simulator packages—which actually does *not* apply! It is based on certain assumptions which, in effect, do not correctly model either the LLC or WPT magnetic structures. So, we need to create our own simplified (transformer) model. The correct/best models can become really complex, and we'll reserve that discussion for a different day!

The Control Loop

After selection of the power components, we inevitably land up at another fundamental juncture: *the control loop*. Briefly: how do we change our operating point to control the output? Whether it is based on changing the switching frequency, or duty cycle, phase, or set input voltage (all being "traditionally" attempted nowadays). How do we get to the best, most optimal, *highest-efficiency* point, for any coupling, any load, and hopefully for a widely varying input too. Mission Impossible? Almost! Such questions typically lead to unpredictably long, trial-and-error design phases followed by rather clumsy solutions, even from major companies trying to "spearhead" resonant topologies. The truth is: *none* of the above control methods suffice *on their own*. You need *multiple* control loops working simultaneously, yet not "fighting" each other. However, now we are once again approaching the gates of proprietary solutions, which we need to steer clear of in this chapter.

But first things first: do we even understand "L" and "C", and their basic combinations? Probably not, or we would have known long ago why resonant circuits *inherently offer such high efficiencies*. Our relentless gut instinct warned against "inherent inefficiency arising from high leakage inductance (a.k.a. low coupling)". The reality being: we can easily get typical efficiency of over 90% from an LLC converter, and that can go up close to 98% if we use synchronous rectification. Oh really?! *Despite* the high leakage? No, probably *because* of it!

Resonant Circuits/High Efficiency

As a quick re-introduction, or refresher, to mundane "reactive components", inductance L and capacitance C, refer to **Figure 1.1**. Note the essentials here: the response to any (sine-wave) harmonic of any arbitrary applied disturbance means that, for example, in an inductor, the current lags the voltage by 90 degrees; and in a capacitor, it leads the voltage by the same amount. So, calculating the instantaneous power (i.e. V×I product) over a full cycle, for both L and C (considered *independently* connected to respective AC sources for now), we learn that both of them will return *all* their energy back to their source (twice every cycle too, see the shaded portions). We conclude that pure reactive components store, *but never dissipate* (i.e. lose) energy, except within their *resistive* parasitics, such as equivalent series resistance (ESR) in capacitors, or DC resistance (DCR) in inductors. Further, parasitics may not be explicitly shown in simple schematics, but are *essential* to natural processes. So, if we discover that something doesn't converge in simulations, or "explodes" in a math spreadsheet (divide by zero, for example), let us not ignore it. Most likely it really is incorrect/unnatural. And we need to fix it.

To create a standalone reservoir of energy, or a "tank" circuit quite literally, we exploit the fact that the two lossless candidates, L and C, are essentially of *opposite types*! Like holes and electrons. Male and female! They pair well! They are complementary. So, if they are combined in parallel for example, we are in effect forcing the voltage across them to be equal always. But their currents are "90 + 90 = 180 degrees apart", which simply means that when current is coming *out of*, say L, it is going *into* C at that very moment. And so on. Back and forth! In effect, the entire parallel LC network can present an *infinite* impedance to the applied AC voltage source, because the net voltage across the LC can equal the applied AC voltage instantaneously, and eventually, the LC network will draw zero current (there is no voltage differential left to drive any current!)—which implies infinite impedance of the LC network! When (and if) that happens, the LC will contain a fixed amount of stored energy always, constantly sloshing back and forth between the L and the C:

$$\frac{1}{2}LI^2 \Leftrightarrow \frac{1}{2}CV^2$$

*The sloshing occurs only at a certain natural frequency though—s*omething that is not so obvious from the figure perhaps! But look more closely: if we want the two opposite current components (the "180 degree apart" components) in the L and C *to also be equal in magnitude*, for *complete* matching, in effect we want their impedances ($2\pi f \times L$) and $1/(2\pi f \times C)$ to be equal. And that implies the applied frequency *be exactly* $1/(2\pi\sqrt{(LC)})$, called the "resonant frequency", to ensure a complete "match". A perfect couple! You can now leave them alone… A parallel LC tank it is!

When that exact matching occurs, say by sweeping the applied frequency of the source, the currents in the L and C will be not only be opposite in direction, but numerically e*qual too*, and will thus be continuous (to each other), *self-contained* and self-sustaining. In this condition, no net current will theoretically *ever* need to be drawn from the input source, at least *eventually*. The voltage on the LC combination can, and will, rise to any voltage level equal to that of the source, including up to infinity if the source is up to it! But finally, the net current drawn from the source will always (eventually) be *zero*, *provided the source is exactly at the specific "resonant frequency"* above (otherwise not)! And finally, since no current is drawn from the source anymore, even if we remove it altogether, the resonant tank *can go on forever*, sloshing energy back and forth between the L and the C (at its natural resonant frequency). As a corollary: the only reason the voltage across the tank (and the current through it) decay, is due to *resistances*: ESR, DCR, and any "load resistor" that we may connect across it. Similarly, a *series* LC combination presents zero impedance, which is why, to avoid infinite currents, we prefer to explain it using a *current* source. Eventually, the net *voltage* across the combined series-LC will drop to zero (equal to the final voltage across the AC current source). The currents in the L and C have been forced to be the same as they are in series, but their voltage components are now opposite. In fact, those can be made exactly equal in magnitude too, but again, only at the specific applied frequency of $1/(2\pi\sqrt{(LC)})$. We thus have a series-LC tank!

Hypothetically, we can suddenly switch out (i.e. bypass), the current source, leaving the series LC components connected in series. Things will carry on as before! Thinking harder, we realize we have arrived at exactly the same configuration as the parallel LC tank if its voltage source were suddenly removed! Both are merely a simple loop consisting of L and C now, with circulating currents accompanied by pulsating/alternating voltages, with energy constantly sloshing between the L and the C, and with *no* dissipation ever. The only difference being that in one case we used a voltage source initially, to set the remnant voltage amplitude across each component of the now standalone LC tank, allowing the current to be what it is (based on impedances), and in the other case we used a current source to set the remnant current amplitude flowing in each of the L and C, but allowing the voltage across each to be what it is. In both cases however, we can think of it in the following way too: similar to the topmost Mathcad-based plots in **Figure 1.1**, the inductor "returns" energy to the capacitor (which is now behaving as the applied AC voltage source!). A bit later, the capacitor "returns" its energy back to the inductor (which is now behaving as the applied AC current source). Notice how it is all so complementary and elegant!

Note also that in the process, we can get huge "amplification" too: for example, if we had connected the series LC tank with a voltage source in series instead of a current source, we would get (close to) *infinite* currents. Similarly, a current source applied to the parallel LC would have forced it to infinite voltage.

Intuitively, in a series-LC circuit at resonance for example, the L and C impedances are equal and opposite, so it behaves quite like a "dead short" at that specific "resonant" frequency. This resonance action is basically very helpful to be able to create fields extending over greater distances, to instigate Faraday's Law, which is what we try to do all the time in WPT. That was exactly how Nikola Tesla was also trying to do it. He created very high voltages (millions of volts), instigated by a huge dV/dt spike, produced by interrupting very high current flowing in a huge inductor ($V = LdI/dt$ of course).

But note how nature tends to self-stabilize: the more energy you try to draw from the resonant tank, the less effective it becomes, because it "knows" that now you have introduced (load) resistance, and *the resonant voltage/current starts collapsing*. Like a carpet pulled from under your very feet as you try to walk across. But it is not that bad: without resonance, energy wouldn't have even made it over that distance to start with! That is why it is so important to characterize the *actual resonant response. How it changes!*
A big part of that natural response, *in effect the inherent "power capability" of an L-C combination*, is determined by the ratio of the resistance (i.e. the applied load mainly) and the ratio $\sqrt{(C/L)}$. This simple fact actually leads to one of the scaling laws which will be discussed in more detail later.

Resonant Circuits and Efficiency Nuances

It is clear: any inductance present in a circuit, discrete or "leakage", is nevertheless still just an "L" in the circuit, and it *will* attempt to return all its energy back to the source. It can't ever dissipate! Not within itself for sure. Maybe in external resistances! If present. And further, in combination with a nearby C, it can also form either a deliberate, or an inadvertent, tank circuit, constantly exchanging energy between the two. Because neither can dissipate! This is essentially why resonant circuits are inherently so efficient to start with.

Generally, if we focus on just reducing ESR, DCR, R_{DS} etc., we will get very high efficiencies. Quite automatically! Yes, instead of AC sources, when we introduce switches connected to DC voltage sources to push energy in at a certain rate, maybe with the intention of controlling the output voltage by choosing the switching frequency, either to be close to, or far away from the natural "resonant frequency", we do have to consider the possibility of *switching losses*. Because we will be turning our switches ON and OFF quite fast, simply to minimize crossover dissipation inside them! *But how "fast" is good enough*? Resonance is actually very forgiving in this regard too. Because it can be shown, as an example, that if we are using the popular half-bridge or full-bridge topologies, we can actually use the leakage inductance present, to coax "soft-switching", or "zero-voltage switching" (ZVS), thus reducing switching losses dramatically. Leaving us with only conduction losses, such as those from the DCR of the windings, or coils, to contend with, and minimize. No need typically, for "state of the art" wide band gap (WBG) switches such as Gallium Nitride (GaN) or Silicon Carbide (SiC). Cheap, low-voltage Mosfets, albeit with sloppy switching characteristics, *but with low R_{DS} to reduce conduction losses*, is the way to go in most resonant converter applications.

Keep in mind though, as another indication of the self-stabilizing, self-limiting advantages of natural resonance, it can be shown that as we lower the coupling, though that enables *greater* power delivery, it also comes at an increasing cost! The dominant circulating/excitation current component (the unused energy returning back to the source every cycle), can become rather high, leading to *higher conduction losses*, thus necessitating closer and closer attention to heat dissipation. And to thermal management. Eventually however, the dissipation term mentioned above is the specific one which ultimately impacts, and dictates, the efficiency of any resonant converter*, provided of course, ZVS is being ensured*.

The correct region to head to, for enabling ZVS, is indicated in **Figure 1.2** using a representative resonant curve. It can be shown that ZVS is possible only if you switch to the *right* of a (presumably *single*) resonant peak...on the left side of that peak, you get hard-switching, which is essentially lossy and gives out a lot of EMI too. Basically, to the right of any resonant peak, the LC network appears "inductive" (on the left it appears "capacitive"). We do remember that inductors "complain" whenever we try to suddenly interrupt current flowing in them—the very reason we need a catch diode in "square"/PWM topologies, as was explained in Chapter 1 of Switching Power Supplies, A to Z, Second Edition, published by Elsevier in 2012. As a result, inductances tend to maintain/force current continuity, searching for whatever path may be available to continue to push current through. So, in a typical half-bridge (totem pole), if we switch OFF the lower switch (Mosfet), the current in the "inductive" LC network being switched, will tend to freewheel—through the *body diode of the upper switch, the only available path!* We then just need to leave a deadtime of typically 100-200 ns before turning the top Mosfet ON. Because during that tiny deadtime duration, the voltage across the upper Mosfet will be forced to near-zero value, by the "inductive" network: its body diode will be forced into conduction to maintain current continuity. And so, when we finally turn the upper Mosfet ON (after the deadtime elapses), we get lossless, or zero-voltage switching (i.e. with no *overlapping* V-I crossover, since V was almost zero during the entire switch crossover).

Surprisingly, none of the WPT practitioners out there seem to have understood this aspect either. Especially those who try to control the output very simplistically: by trying to vary duty cycle, or phase or even the input DC voltage applied to the transmitter, keeping frequency fixed. All these were inspired by classical PWM power conversion obviously, where fixed-frequency techniques are preferred *for reducing EMI* in particular. They didn't seem to realize that with the typical natural *frequency shift* of the resonant peak to a higher level when you load it, as we will explain, they could easily land up, *without doing a thing*, on the wrong (left) side of the peak—where ZVS is no longer possible. Now we will get high losses and much higher EMI too! Just the opposite of what we were intending.

Energy is also not "automatically lost" *between* the receiver and transmitter coils, as some think. There is no "resistor" present in that space, for dissipation to occur! Keep in mind that no part of any energy balance-sheet is ever "un-accountable", especially not when we are in the *near-field* region, as explained further below. Energy is lost only in resistances (as heat), including in the DCR of the coils themselves, but not in the airspace *between!* Unless of course there is a *metal object* present therein (the "foreign object detection (FOD)" issue), which generates eddy currents and thus heat. But that too is still, in effect, just a resistance, an identifiable/quantifiable one, in the overall effective electrical schematic. It can all be accounted for. There is no black magic, nor black holes involved here!

It is however always helpful to mentally distinguish between *fields* and *waves*! Waves indeed, we may not be able to keep track of fully! They tend to disappear into outer space. Efficiency truly becomes questionable. But for fields, those restricted to the "near field" region, it is just all *stored* energy! *Fully deliverable, with the rest being fully recoverable!* None of it is ever inherently doomed or "missing". And luckily, this near-field region extends up to 480 meters for our typical 100kHz operation (it is precisely 0.16 times the wavelength, which is 3 kilometers in our case)! So, we can certainly disregard electromagnetic waves, which will appear only when, hypothetically, the 100kHz fields we generate are strong enough to reach into the far-field region, i.e. half a kilometer away! Which they clearly don't! Why? Because Nikola Tesla is not around anymore! Yes, he had tried all that! And nowadays, instead, we have UL (Underwriters Laboratories) and FCC (Federal Communications Commission) watching over us carefully! Quite a sobering thought!

Things will change dramatically only when we take our frequencies up to RF-levels. Indeed, RF-based WPT methods *do* have intrinsically low efficiencies, of around 5-10 %, because at their typical 2.4GHz, the near field region ends at 15mm! Less than a third of a finger away! As a result, a huge fraction of the energy those RF-technologies create, do leave forever as electromagnetic waves headed for distant planets. But not so in any of the near-field, low-frequency category of resonant topologies under discussion here. The underlying physics of resonance is shared by all the LLC/WPT methods that we are discussing here.
These clarifications, combined with the basic properties of inductors and capacitors, imply that on paper at least, any resonant system can approach the ideal of 100%. *Provided it is correctly designed,* as was hinted at in our ZVS discussion too. Usually, we just stop pushing efficiency at the point *where it stops making economic sense*. A product needs to be cost-effective after all! The ostensibly "fundamental" question of some: "What exactly is your efficiency", has a simple answer: "*How much do you want*?" Actual *implementations* of near-field WPT should certainly be questioned. To know for example, whether they are exploiting ZVS or not. That *will* affect efficiency. And distinguish technologies. But nothing arcane though.

Control Loop or not

As indicated, resonance is powerful, but inherently self-limiting too! Which is why it is very tricky to design in. The peak bulges, but then collapses as mentioned! Nothing obvious we can carry over from classical power for sure. That was all relatively "single-direction" and predictable.

We realize that in a simple buck regulator we didn't even need a control loop per se. For example, if we had 15V input, and we wanted 5V out, we just had to drive the buck converter at one-third duty cycle, *irrespective of load current, and we would always get 5V* (though indeed, the load must be reasonably high to enforce continuous conduction mode "CCM"). At a later design stage, when we finally introduced the (AC) control loop, its purpose was only to impart much higher *precision* (regulation) to the desired/set output voltage.

The "direction of correction" ("DoC") was unconditionally obvious too, as in all of classical power: for all three fundamental topologies, you just need to increase D, and in response, V_O (or the DC "gain" which is just V_O/V_{IN}), would definitely increase. No mystery there. See **Figure 1.2**. But coming to resonant circuits where traditionally, "frequency modulation" is used to control the gain and thus the output, we should actually *not* be making any of the three implicit assumptions made via the familiar "control algorithm" reproduced in **Figure 1.2**, from Fulton Innovation (an Amway subsidiary, the guys who actually created the first versions of Qi).

a) The height of the resonant curve is *not* fixed as it seems to be, from their diagram: it is in reality, a strong function of load. It will inevitably start *flattening out* if you try to increase the gain to deliver more power to the load (by lowering the frequency). That flattening, indicative of the self-stabilizing properties of resonance, was clearly not considered, or anticipated, in the Fulton/Qi curve. But the mistake is not fatal! It could still be considered implicit therein and in fact, even if that happened, it would likely be handled quite well by any simple control loop, with a fixed DoC, as indicated by the load transition cases shown in **Figure 1.3** (a to c)**.** But in fact, that was not even close to the actual shape of the resonant curve they had on hand, as we will reveal!

b) The resonant curve is also *not fixed* relative to frequency, as the curve from Fulton implies. It can move *sideways*. *To the right for higher loads*. And not just slightly, but dramatically, as we will learn. Precisely 41% for very high loading and a coupling of 0.5, as an example! It is a complicated function of both load and coupling in general.

   Note: This sideways shift of the peak is not even obvious in simulations if we use the built-in default "transformer" available in most commercial simulator packages.

   Now, with the lateral shift of the resonant peak, strange new things will happen! Qi for example assumes the peak is fixed and thus lays out a lower limit of 115kHz (to avoid rolling over it!). But what if the peak was actually at 130kHz? Now, if we headed towards 115kHz, we could also easily "roll over" the top of the (flattened), shifted peak, especially *if our C/L ratio was inadequate, in relationship to the R (load resistor)*, i.e. *insufficient power* available from the resonant circuit to meet our basic power requirement.

   There is in fact something almost intangible, called the "power capability of a resonant circuit", which depends on C/L as previously mentioned. Unknowingly, Qi "approved" many transmitter coils with widely-varying, almost random, C/L ratios, all supposedly for the same 5W/15W power level (and often for the same V_{IN} too). The overall power capability of the selected reactive components was all over the place! They were probably thinking "Buck" when they wrote their standard, which as we indicated has a transfer function independent of power (in the first order). But resonance is very different.

   Worse, by the flattening and roll-over, you would then land up in the region to the *left* of the peak, *where your entire DoC needs now to be reversed*: here you need to *increase* frequency to deliver more power, not decrease it! But if you didn't even know that you

had rolled over (and how could you?), the power would collapse increasingly, as you lowered the frequency further and further, eventually leading to a "mysterious" turn-off—due to power inadequacy, control loop confusion, or just too low a voltage on the receiver for it to even be able to communicate back its requirements. See **Figure 1.3d** for a similar example of what can happen!

c) Finally, why assume we have just one resonant peak? There were *two* capacitors remember? One from the Primary (transmitter), and another one *expectedly reflected from the Secondary* (receiver). Both would end up interacting with the existing L. So, *two* second order poles. *Two* resonant peaks—basic physics! But looking back, *all those engaged in WPT today*, seem to have based their efforts on a *fundamentally flawed "double resonator"* assumption, which unbeknownst to them, gave not one fixed peak, as they had assumed, not one *moving* peak (which they could have conceivably handled, say by exotic or proprietary control algorithms, had they known), but *two wildly moving* peaks, *not even amenable to a simple DoC, or even a very clever control strategy* .And that is precisely how the world possibly ended up with "inherently poor user experience of inductive systems". This is the critical flaw encircled in **Figure 1.4b** too. The entire WPT world subscribes to it! They just (still) don't know, that they don't!

As mentioned, in the early stages of development of WPT, engineers with RF backgrounds likely, instinctively felt they needed to *tune the receiver to exactly the same frequency as the transmitter*, each independently measured and set, for ensuring "maximum power" transfer. Or maybe, their theory was based on some transmission-line-inspired "impedance matching" concept. Also RF! Actually, it was all just prior mental conditioning, not applicable to relatively closely-coupled, near-field systems at all, i.e. those involving fields, not waves!
And by now, it may be too late to correct. Initially, they could have headed to WPT via the relatively well-understood LLC topology shown in **Figure 1.4a**. We now realize that *has only Primary-side resonance* (no "C_S") *for good reason!* We could then arrive at a better form of WPT shown in **Figure 1.4c**. After that, to improve the magnetics further, we could evolve towards **Figure 1.4d** or something similar. But unfortunately, by now, both the potential elimination of C_S as applied to the field of WPT in **Figure 1.4c**, and the magnetics of **Figure 1.4d**, are likely proprietary. A bit late in the day to play "catch-up".

The double-resonator flaw (i.e. the presence of C_S in effect), is shown encircled in the equivalent circuit diagrams presented in **Figure 1.5** too. But that figure has some new details, which will be helpful in our upcoming design calculations/procedure.

The First Harmonic Approximation

Figure 1.5 shows us how to go back and forth between equivalent AC and DC models, including reflecting impedances over to either side of the transformer, for enabling simple calculations. That is all based on keeping in mind that even if we apply a "square-ish" voltage shape to a general LC network, we need to mentally break it up into its Fourier (sine) components….and see the effect of each sine component on the network. Later we can reassemble the responses back. Turns out, it is usually sufficient to assume that of all the harmonic constituents of the "square-ish" waveform, only its fundamental (first) harmonic has the main, or dominating, effect. Therefore, resonant topologies are commonly studied using the First Harmonic Approximation" (FHA). We ignore other harmonics.

Using FHA, it can be shown that the peak-to-peak value of the first harmonic (sine wave component) of the input voltage, *exceeds* the peak-to-peak of the square wave responsible for it, by the simple factor of $4/\pi$=1.273. But in a half-bridge, as opposed to a full-bridge, we also have to understand that its resonant capacitor acquires an *average DC level of half the supply voltage*, subtracting from the effective input applied to the L-C network. So, in effect we need to halve the stated peak to peak value, down to $2/\pi$ times V_{IN} as indicated in the figure.
Using FHA, and further realizing also that if we "reflect" a given load resistor present on the output of a DC-DC converter, to an "equivalent" Primary side resistor, we must maintain the actual dissipation (energy) term even through the process of "reflection". The energy can't change by reflection! So combined with the FHA, it can be shown that a certain "R" present on the Secondary side (on its *rectified* DC output) can be reflected back to the Primary, to appear as an *effective (AC) resistor* of value $8N^2/\pi^2$, as shown in **Figure 1.5**.

Note: Keep in mind that the turns ratio used by us is "$N= N_P/N_S$" whereas most simulator packages prefer to call turns ratio as N_S/N_P. We may have to take the reciprocal!
So, the "equivalent AC load resistor" is actually of <u>*smaller*</u> value (in ohms) than the corresponding DC load resistor from where it came from, by the factor $8/\pi^2 = 0.8107$, all from FHA analysis. That equivalent AC load resistor, we can then reflect onto the Primary, *as a typically larger resistor* (since N >1 typically), by the factor N^2 (to keep dissipation the same too, despite reflection). Combined, we get the overall factor of $8N^2/\pi^2$, as indicated in **Figure 1.5**.

Similarly, maintaining the stored energy too, in the process of reflection, and realizing that voltage scales (typically increases), by a factor ×N in going from the Secondary to the Primary, and current scales (typically decreases) by the factor ×(1/N) in doing the same, a capacitor must reflect in going from the Secondary to the Primary, as C_S/N^2 (its value typically *decreases* since its voltage has increased in the reflection process, but energy is fixed). Similarly, an inductor reflects (typically *increases*, since the current through it has decreased) as $L_S \times N^2$. All this is based on maintaining the same "$1/2(LI^2)$" and "$(CV^2/2)$" i.e. energy terms, in going from the Secondary to the Primary!

Note: That is why in classical power conversion, we strongly recommend lowering the "trace inductance" on the Secondary-side of a Flyback converter. Even the typical 20nH/inch reflects into the Primary by turns-ratio squared, becoming a rather *huge* stray inductance whose energy has to be usually just burnt (dumped) in the Zener clamp at every turn-OFF transition. We will not derive any of these formulae here, as most references on the subject usually carry all that. We will focus on gathering and using all the tools necessary for a very thorough and complete design.

Note that in **Figure 1.5**, the "tuning capacitor C_S, which as mentioned, everyone seems to have rather incomprehensibly introduced in modern WPT, reflects to the Primary and creates another resonant peak! It is a fundamental flaw, as it leads to the *double-peak* gain profile possibility hinted at in **Figure 1.2**, which as mentioned, is not even amenable to any known DoC, or any apparent/exotic control loop strategy—to ensure ZVS for example, *or even a dependable output*. And worse, all the "fixed frequency" approaches, i.e. changing phase, or input voltage, or duty cycle, are even less promising than frequency modulation methods in this regard.

What is so special about the LLC?

As suggested, our first stop should have been the better-understood, more-familiar, LLC converter. From there, on to WPT. Because we realize that both LLC converters and WPT systems need to be designed in *much the same way*. They *are the same* topology! Or at least *should have been*, *had modern WPT implementations been done right*. LLC and WPT constitute two sides of the same resonant coin.

As indicated, what is so "wrong" in WPT today, and in fact totally unnecessary, and *not* there in LLC (for good reason), is the presence of the Secondary-side capacitor C_S shown in **Figure 1.4b**. This figure is exactly where *all* attempts in WPT are stuck today, without even knowing it.

As also indicated, it all came about perhaps based on a wishful attempt to tune C_P and L_P in the Tx, with L_S and C_S on the receiver, *radio-frequency style*, expecting maximum power. Both L_P and L_S were in effect measured independently, forgetting the effects of their mutual interaction when brought closer. But surprisingly, everybody is doing exactly that in WPT today, though the tuning frequencies may differ.

They *all* apparently assume what is now presented more clearly in **Figure 1.6a**: that if they tune the Tx to say 100kHz (as Qi does), and the Rx to the same 100kHz, *independently*, then the net resonant response when the coils are brought closer, *is still a single peak at 100kHz*. But what really happens, is also shown in the same figure, and it can be easily proven using the basic, but accurate-enough, "transformer model" shown in **Figure 1.5**. The gain profile not only shifts sideways, but splits into two virtually *unpredictable* peaks (**Figure 1.6b**) in their case.

Unpredictable in "height", just for starters (so how can you *ever predict* the "power capability" of any L-C network anymore?). Also, not only shifted sideways, but *two* shifting peaks (no bullet-proof control loop seems even possible anymore). Plus, the right of one peak is also the left of the other (so where exactly do we need to "aim" for, to get ZVS?). There are problems galore for **Figure 1.6b**.

In contrast, had we gone down the LLC path, we would arrive at **Figure 1.6c**. That recognizes and predicts both the extent of the flattening, and the lateral shifting. It has only one peak too! It is thus fully amenable to a properly-designed, smart control loop. Though of course, in WPT, to handle the variable K, we will need to progress on to proprietary control loops! That does get very, very complicated. Not to be discussed here either.

Keep it simple stupid (K.I.S.S.) as they say, but don't be stupid enough to oversimplify either! All those who did exactly that, and tried to keep a fixed switching frequency, resorting to duty-cycle, phase or V_{IN} modulation to control the output, didn't realize that they all *will* land up on the wrong side of the resonant peak (or peaks in their case) sooner or later, i.e. with or without any C_S present! So: poor "user experience" eventually? Probably worse than basic Qi too.

The Elusive Resonant Peak

Now let's explain the lateral shifting of the (single) resonant peak using the simple "transformer model" shown in **Figure 1.5**. The Primary inductance L_P splits into a Secondary-side coupled portion "$K \times L_P$" and a leakage or uncoupled portion $(1-K) \times L_P$. We are ignoring the encircled, flawed C_S now. See the equivalent circuit being carried over to **Figure 1.7**, to explain one of the key reasons why the LLC became so popular to start with, as compared to other "competing" resonant topologies such as the series-resonant converter ("SRC").

We now show in **Figure 1.7** how varying loads produce a predictable shifting of the resonant peak as indicated, *between two extremes*. One extreme occurs for no load, called "f_{LO}", the other for very high loading, called "f_{HI}". One peak location is based on the *entire* Primary inductance coming into play, the other on leakage inductance only. The gain curves for intermediate loads fall somewhere between the two operating frequency limits, f_{HI} and f_{LO}. Note that it is always a single peak in our case, but shifting between two extremes. *It is not a double peak*! No constituent curve in the set, has two "humps". Look closely.

This is what led to the big perceived advantage of the LLC. It simplified the control loop and the design of switches and all associated circuitry too (no need to go to very, very high frequencies just to reduce gain, say at very light loads, as was the situation with the "competing" SRC topology). This also leads to a rather narrow, manageable band of EMI. We could basically keep f_{HI} below 150kHz, just to pass EMI testing more easily, and restrict the operation between f_{HI} and f_{LO}, for *all* loads.

Notice how all the gain curves intersect *exactly* at the point described by: Gain=1 and f= f_{HI}. Let us call this the "magic operating point" ("MOP"). All the curves intersect here because at this frequency, the leakage inductance and the resonant capacitor are in complete resonance, and form a "dead short" in effect as explained previously! With that "dead short, the entire input voltage will appear unfettered, straight at the output, irrespective of load! So, $V_{IN} = V_O$. i.e. Gain =1, *for any load!*

Note the eerie similarities to the "load-independent DC transfer function" of a Buck as discussed earlier! So theoretically, now, if we operate *precisely* at f_{HI}, we would *not* need to do anything at all to adjust the output, for any load from zero to maximum! *Provided we had designed our circuit for a gain target of exactly 1*. What that means is the set reflected output voltage (often called V_{OR} in a Flyback for example), would need to be exactly equal to the applied DC input voltage. Then, clearly: Gain =1.

Theoretically, we could then even choose to operate precisely at f_{HI} *always*. No control loop in place. It does seem plausible and promising too: we would be operating to the right of *any* applicable resonant curve *always (for any load)*, and we would thus always get ZVS too, automatically. Since it is fixed frequency, it is good for EMI testing too! All we really have to do is to specify the *right* turns ratio so that $N \times V_O = V_{IN}$. As simple as that! It should work! A perfect topology it seems! What if anything, is wrong?

Like all things "power", nothing is ever straightforward! First, this series L-C "dead-short" visualization is true only in the first-order. When we introduce parasitics, the output starts to droop at high load currents due to simple resistive voltage-divider action between the DCR of the Tx coil (R_{DS} of FETs too), and the load resistor. So, we may opt to forcibly ensure that we have very, very low parasitics, just to make the output droop "acceptable" (whatever that means). But that will require very, very thick expensive copper! Instead, how about *using resonance* (more) effectively, to literally boost (amplify) the output voltage if it starts to droop too much? In fact, then maybe we can amplify it even further to compensate for a droop in the input too! That will be much smarter and more cost-effective, compared to using utterly thick copper windings! That is actually the right way to do it, the basic method underlying our graphical aids too. It does demand a lot of design expertise though, at least to generate those curves. But to use them is very easy, as we will learn.

In **Figure 1.7** we show the basic problem with the fixed-frequency (no control loop) method. If the set fixed frequency does not coincide *exactly* with f_{HI}, and/or the desired/set gain is not exactly unity, we will land up slightly to one side of the MoP, where we can see the gain curves can drop off dramatically (under our feet, quite literally) as we change load, especially if we are to the left of the MoP. If we land up there, we would likely experience a dramatic, droop in the output, for load variations. Typical tolerances in L and C, not to mention variations in coupling (which also affects the leakage inductance, and thereby f_{HI}), will play havoc with the desired output. The only way out is to manually tune every single transmitter in production, to the exact resonant frequency of the specific L and C we happen to have on a given board (yes coupling variations included).

We conclude that the constant, fixed-frequency approach, may make a "good (tuned) demo board" to impress the uninformed with, but it doesn't promise to ever qualify as a commercial product! Not without a control loop. That approach may have "sort-of" worked with a classical power topology, as we too indicated was possible for a Buck, on the basis of its simple DC transfer function, but resonance is not to be underestimated. We do need a control loop here, and for many *more* reasons than we ever needed it in a Buck!

If we have no tolerance even for L and C variations, how can we ever handle the second most vexing problem of all: input voltage variations. The most vexing of course being the variations in K, especially in WPT systems, as we had mentioned.

The graphical technique proposed here will handle input droops too. We realize that since Gain = V_O/V_{IN}, if V_{IN} falls, then for maintaining a fixed V_O, we just need to increase the Gain by the same factor! That should be easy to accomplish, almost automatically by a control loop, say one based on frequency modulation, as is commonly used in LLC converters. Provided we have the right "bulge" in the selected resonant curve, we can simply head towards f_{LO}, stopping wherever we reach the desired output. Indeed, handling K-variability is another huge challenge altogether.

Heading to a Practical Design: Setting Gain Target

In this chapter we are recommending that we must always, for the LLC and even for WPT, try to ensure that our system restricts its operating region to the region between f_{HI} and f_{LO}. This we can ensure by *designing it to have a "gain target" (i.e. VOR/V_{IN} where VOR = N× V_O.), set slightly greater than 1 (~1.05) at maximum load and maximum input voltage*. The rest of the procedure will become clear shortly. See **Figure 1.7**. We can see that the 1.05 gain target ensures we are slightly *to the left of the MoP!* Not to the right. *Ever*, as we can show! We will automatically move to the left of f_{HI} if the input voltage falls. But if the output is, say, suddenly unloaded, the output will jump momentarily higher (above the set 1.05). The control loop will then try to reduce it by increasing the frequency, but *won't need to go above f_{HI}*, since even the *no-load curve gain drops to below 1.05 near the MoP*, which we know occurs at f_{HI}. It will thus all work just fine between two frequency extremes. But *the key is to set the gain target (using the correct turns ratio in effect), to slightly above 1*. Most LLC designers do so, but rather inadvertently.

As a side note, all WPT systems around us today actually ignore even the above simple guidance concerning the set gain target/desired VOR. We already know that all the "approved" Qi transmitters *turned a blind eye to the critical C/L ratio*, so the "power capability of an LC network" was not understood. But now we must recognize that all their "approved" receiver coils, have *different turns ratios with respect to even a single transmitter*, and also, the receivers are virtually allowed to set output voltages of their choice (no real guidelines to set V_O, and to thereby establish the desirable VOR, which we just learned should be set slightly higher than V_{INMAX}).

In general, the gain target of any existing WPT technology out there could just end up being less than 1. Unknowingly! As a result, WPT systems doing "frequency modulation" will end up trying to control the output by operating to the *right* of "f_{HI}" as indicated in **Figure 1.7**. Not strictly to its left as strongly recommended. One obvious drawback of that, besides higher (broader) EMI, is that at very light loads, the transmitter can no longer reduce the gain any further by raising the frequency, so the receiver has to suddenly apply an internal overvoltage clamp (usually dissipative), just to protect itself. Keep in mind that Qi restricts its operation between 115- 205 kHz, to simplify the design of the power stage. But if the control loop tries to increase the frequency above the upper limit of 205kHz, internal clamps are activated within the receiver to try and protect itself. Not elegant! A heavy price to pay for ignoring both power capability, and VOR.

Besides, a key question for any clamp is: for how long is it safe? Or even practical? That in fact, is one of the reasons Qi can never do laptops! Imagine applying typically 50-100 Watt Zener clamps inside the laptop, for over a second (for that is how long the correction loop of Qi can be in practice, with a few missed packets, as has been commonly observed).
Now add to that mess the K-variability factor, *and the double peaking issue arising from C_S*, neither of which we have even considered while doing the gain plots in **Figure 1.7**, and all bets are off! Proprietary solutions are probably the only solution to these very complex issues and the unpredictable interplay of their influencers.

As we had declared initially: the awesome power and potential that resonance brings to our doorstep quite literally, comes with a hefty price tag: *it demands total design expertise*. It could have been corrected, but that window of opportunity has firmly closed. Only proprietary solutions exist to solve this issue (too).

LLC and the Power of Magnetics

Before we reveal our design procedure, we try to get a final "feel" for the magnetics of resonant topologies, the LLC in particular, and why that is another big attraction to use the topology, compared to the magnetics in classical power conversion. A lot of what we will learn applies in principle to WPT systems too.

Figure 1.8 provides a good visual impression of the key advantage: *the lowered size (and cost) of LLC magnetics*—making it akin to a Forward converter transformer in size, plus *without the need for any output choke*. Focusing on the transformer only for now, it is small, because no *fraction* of the useful power delivered to the load in an LLC topology, or in a Forward converter transformer, needs to be ever stored, even temporarily, within its magnetic core during the power transfer/conversion process. The reason is that the Secondary windings conduct *at the very same time* as the Primary windings, so the flux associated with the load current flowing in the Secondary, is fully canceled by the *additional* flow of current flowing into the Primary winding to support it. The input current component associated with the actual delivery of useful power to the Secondary, is obviously load-dependent. The magnetization, excitation component is not. The latter leads to the steady conduction loss term we discussed earlier, and the residual flux at all times within the transformer core. Since any flux attributable to the load-dependent currents flowing in the Primary and Secondary windings, cancel each other out, the net flux we are left with in the core is still, just the one related to the initial magnetization current component, which was and remains load-independent. Think of the flux as the one present at zero load current. It doesn't budge with load! Nature opposes any change. And it has done so very successfully here! Fully!
If a Forward converter transformer, or an LLC/WPT magnetic core, is made visibly larger for higher powers, it is really not to support higher load currents per se. It is made "bigger" only to accommodate the thicker copper windings constituting the Primary and Secondary (more window area, or larger surface area for better cooling). And for the same reason, the "core loss" term in a Forward converter's transformer or in an LLC or WPT system, is also just based on the zero-load-current case! Not for maximum load current. How convenient! If it gets hot, it is not core losses from within the core, just conduction losses from the adjacent copper, to consider and reduce.

To re-iterate: the transformer needs to store only the energy related to the *load-independent*, "magnetization current" component, which can be intuitively thought of as an excitation component necessary to initiate basic "transformer action", thanks to Faraday's law of magnetic induction. Nothing related to the actual power transfer process is stored in the core! So typically, in an LLC or Forward converter, the transformer volume requirement is roughly *half to one-third the size of a Flyback transformer of the same power capability*.
In a Flyback transformer, the Secondary windings do *not* conduct at the same time as the Primary windings. So, we basically need to store, *all* the energy meant for delivery to the output, inside the transformer core, during the ON-time, then pull it *all* out during the OFF-time. The Flyback transformer core is thus a bit like an Amazon warehouse, for energy! It needs to be *big*. For the same reason too: everything that goes out, must be stored there at some point of its journey.

But the Forward converter, unlike the LLC, needs an output choke, to finally store a certain fraction of the load-dependent energy too. That is because energy storage in a magnetic core is fundamental to the power transfer process of all PWM-based classical power topologies. Though by different amounts. That aspect was discussed in detail in Chapter 5 of Switching Power Supplies A to Z, Second Edition, in particular on Page 208. Reproduced here partially:

$$\Delta\varepsilon_{BUCK} = \frac{P_{IN}}{f}\times(1-D);\ \Delta\varepsilon_{BOOST} = \frac{P_{IN}}{f}\times(D);\ \Delta\varepsilon_{BUCK-BOOST} = \frac{P_{IN}}{f}$$

This tells us how much energy per cycle needs to be "cycled/stored" for the three fundamental (classical) topologies. Since P_{IN}/f, is the energy per cycle, we learn that the Buck-Boost (or the Flyback) need to temporarily store (and then release) 100% of the energy making its way to the output. In comparison, Buck inductors tend to be small, especially when the amount of "bucking (stepping-down)" asked of them, is less. And why Flyback (Buck-Boost) magnetics are so much bigger.
Basically, in the case of the Flyback topology, its transformer is its energy-storage element too, besides providing "transformer action" (voltage scaling and isolation in particular). It needs to temporarily store 100% of the energy transferred (not any fraction of the duty cycle, D), since a Flyback is just a Buck-Boost derivative. Similarly, a Forward converter, despite a simple transformer, *does* need an output choke, since it is a Buck-derived topology, and therefore needs to store a certain fraction (proportional to 1-D) of the power transferred.
All initial points are in perspective by now, hopefully providing us a much better "feel" for resonance. So, we will now get deep into the simplified but accurate design procedure suitable for any LLC and (correctly designed, i.e. *no C_S*) WPT system.

Defining the Kernel

As indicated, we just need to take a kernel, study it very well, and then use scaling laws to take it to any power, frequency and input voltage level. But we also want to introduce a factor for establishing a certain gain margin to serve us well when the input voltage starts to droop, i.e. to handle input variations! A lot of that was already demonstrated in Switching Power Supply, Design and Optimization, Second Edition, but the following key items, were not considered, which we want to do here:

a) How to scale as per input voltages. In other words, if the kernel was studied for an input DC range of 32V to 52V input, and now we want our LLC converter to work from 200V to 400V DC. How does that affect our overall scaling of components?

b) How to handle different coupling coefficients? In the previous exercise, we had a fixed K of 0.9 and when we scaled, we kept the same K. In reality, we want to consider any K, even very low couplings, as in WPT.

In our previous exercise we had initially started off with a certain very low frequency kernel, and then scaled that to become a 25.5W LLC converter for a Power over Ethernet (PoE) application (32V to 52V DC input). So just for historical continuity, we have for the graphical aids presented herein, used the same values of L_P and C_P (57.2µH and 225.8nF) that we had arrived at, as a result of scaling our very first very low-frequency kernel. Now, we will just use these suggested PoE converter values as the kernel for the next scaling exercise here. Since we are scaling anyway, we could use any initial kernel. So, we then introduced these L_P and C_P values into a general Mathcad spreadsheet to generate the graphical aids presented. The spreadheet is in the Appendix to Chapter 1.

Just to throw more light on the gain curves we will get in an LLC topology, as an intermediate step, we used the above Mathcad spreadsheet to suggest the "recommended" load resistor of 20.6 ohms. And then we used multiples, or fractions, of that to generate the stacked gain plots in **Figure 1.9**. Note that 20.6 ohms is the equivalent AC load resistor to be applied to the equivalent AC-AC, non-isolated model shown in **Figure 1.5c**. It was recommended by the spreadsheet, to be able to deliver the desired average power of 25.5 Watts in the final converter, over the desired input range of 32 to 52 VDC.

We have also used Mathcad to predict exactly where the entire LC network appears "inductive" (solid lines) and where it appears "capacitive" (dashed/dotted lines). It tells us where ZVS will occur and where it won't! A bit to our initial surprise we do see that the changeover from solid to dotted does not occur precisely at the "peak" value, as always somehow assumed. The reason is that the peak is determined where the leakage $(1\text{-}K)\times L_P$ and resonant capacitor C_P, are in perfect resonance ("dead short")… but on closer examination, that does not necessarily coincide exactly where the entire network (including $L_P \times (K)$) "resonates". Basically there is a slight shift on account of the magnetizing inductance and the load entering into the picture. The practical implication of that is it may not be enough to be merely "slightly to the right" of the peak voltage for ensuring ZVS. Maybe a little more than "slightly"! But in any case, the transition from lossless to lossy switching is a bit gradual anyway, so though the losses may go up slightly as we leave the solid region of the curves, and enter the dotted portion (on our way to the peak), they might still be acceptable enough, to *not* warrant more effort on our part to tweak the control algorithm.

In the lower part of **Figure 1.9**, we also show the complete mess the mere introduction of a reflected capacitance of 100nF (arbitrarily chosen), from the Secondary to the Primary side, causes. *On display here is all what is wrong with modern WPT*: the "double resonator" flaw we had talked a lot about! Notice there are two peaks, behaving differently. The one which shifted to the right, has an even more dramatic shift than anticipated without C_S. So in the case of Qi, a good question to also ask is: what if the resonant peak has not just moved to "141kHz", as we often say as a ballpark estimate of the shifting, but actually to above 205kHz, (the upper operating limit of Qi)? How can we even ever hope to regulate? We are stuck on the wrong (left) side of the resonant curve. If we lower the frequency, the gain collapses further! And of course we are not in a region where ZVS will occur. Unless the "double resonator flaw" helps us inadvertently in this case. Because we have another peak with C_S present. This has actually shifted to the left of the no-load peak location f_{LO}. Very hard to ever predict how any control loop will react to all this mess. Most likely, a bunch of "mysterious turn-offs". And then: "inherently poor user experience of inductive systems".
Let us move on, without C_S here-ever-after!

Practical Design Example of a wide-input 900W LLC/WPTconverter

With the above kernel, we have generated two key design curves, shown in **Figure 1.10**, and **Figure 1.11**. Keep in mind that the values used were L_P = 57.2µH and C_P = 225.8nF. So, here is the challenge: we want to deliver 900W into a 48V output. Select the best L and C values for the converter for an estimated coupling of K = 0.5. For EMI compliance reasons, we want to stay below 150kHz guaranteed. What are the best operating frequencies to carry out meaningful simulations, to test its performance? We are assuming an input DC varying from 200 to 400V.

On account of the high-power the obvious choice is a full bridge, not a half bridge. At 400VDC therefore, the equivalent AC wave applied to the input of the resonant network is as per FHA:

$$V_{INMAX\_AC} = V_{INMAX\_DC} \times \frac{4}{\pi} \text{ Volts}$$

$$V_{INMAX\_AC} = 400 \times \frac{4}{\pi} = 509.3 \text{ Volts}$$

This is the equivalent amplitude of the sine-wave applied as per FHA. (For a half bridge we would have divided this by 2).

From **Figure 1.10**, For a coupling of 0.5 and a V_{INMAX}/V_{INMIN} ratio ("gain factor") of 2 (to allow us to reach 200VDC from the original 400VDC), we see that the recommended resistor to be placed on the output of our non-isolated AC-AC equivalent circuit of **Figure 1.5c**, is 14.7 ohms.

This load resistor is actually valid, *irrespective of the applied/assumed input actually*, because it produces just the right amount of "bulge" in the gain profile curve, in conjunction with the selected C and L (the C/L ratio in effect), to allow for a Gain of 2, occurring somewhere between f_{HI} and f_{LO}, before the gain curve rolls off! So, that there is no hint of overdesign either! Just the right "bulge" in the resonance curve.

Now, had we applied an input of 400VDC through a full bridge, or an equivalent AC voltage of 509.3 Volts, the power at the peak of the AC sine wave input would have been

$$\text{Peak\_AC\_Power} = \frac{V^2}{R} = \frac{509.3^2}{14.7} = 17650 \text{ Watts}$$

Note that the average of any Sine-squared function is known to be half. In other words, the average power we got from the kernel, but with the desired maximum input applied to it, is

$$\text{Average\_Kernel\_Power} = \frac{\text{Peak\_AC\_Power}}{2} = \frac{17650}{2} = 8825 \text{ Watts}$$

But we want only 900W from our proposed LLC converter, So the Power Scaling Factor we need to apply is less than 1 in our case:

$$\text{Power\_Scaling} = \frac{\text{Desired\_Power}}{\text{Average\_Kernel\_Power}} = \frac{900}{8825} = 0.102$$

Coming to the frequency scaling factor, the f_{HI} of our kernel (for the desired K=0.5) is

$$\text{fhi\_Kernel} = \frac{1}{2\pi\sqrt{(1-K)L_P C_P}} = \frac{1}{2\pi\sqrt{(1-0.5)\times 57.2\mu \times 225.8n}} = 62.63\text{kHz}$$

To stay below 150kHz, we would like to set f_{HI} of the new converter exactly at 145kHz. So, the desired frequency scaling factor is

$$\text{Freq\_Scaling} = \frac{\text{Desired\_fhi}}{\text{fhi\_Kernel}} = \frac{145k}{62.63k} = 2.315$$

Note that the input voltage scaling factor was automatically accounted for when we learned to calculate the power we were getting from the kernel, at the desired maximum input voltage applied to it. Key to that exercise was recognizing that the applied input voltage doesn't change the basic shape of the resonance curves. Gain is simply output divided by input anyway! The input voltage scaling factor disappears from the picture, with our little trick of just calculating R, not power, from our kernel. The R was picked to create a certain "shape" (or "Q-factor" as others prefer to talk in terms of, but we decided not to, for various other reasons) to deal with a certain input droop that we want to handle. That is why the associated Mathcad spreadsheet and resulting graphical aids were created and presented to provide R, not power.

Final step, the recommended L_P and C_P values for our 900W converter are thus

$$C_P = C_{P\_KERNEL} \times \frac{Power\_Scaling}{Freq\_Scaling} = 225.8nF \times \frac{0.102}{2.315} = 9.95nF$$

$$L_P = \frac{L_{P\_KERNEL}}{Freq\_Scaling \times Power\_Scaling} = \frac{57.2\mu H}{2.315 \times 0.102} = 242.24\mu H$$

These are the values to try out in a simulation. They establish the desired power capability of the resonant network in effect, at the desired input voltage, allowing for the desired input variation too, and based on out estimate of K. In the simulation, the calculated L_P will need to be split up as follows

$$L_{LKG} = (1-K) \times L_P = (1-0.5) \times 242.24\mu H = 121.12\mu H$$

$$L_{MAG} = K \times L_P = 0.5 \times 242.24\mu H = 121.12\mu H$$

Now, if we had applied 400V DC at the input, then that would have given us exactly 400VDC output too, for the gain target of 1, if we introduced a 1:1 transformer and a diode bridge and output smoothing capacitors. In other words, we calculated it all for a VOR = 400V (or slightly higher as recommended earlier). So, with a 48VDC output instead, to appear as 400VDC in the equivalent DC-DC transformer-less circuit, the desired turns ratio is

$$N = \frac{V_{INMAX}}{V_O} = \frac{400}{48} = 8.33$$

That is exactly how voltages scale through transformer, and we are just using the same principle to reflect 400V into the Primary, from 48V at the Secondary.
Note: when we actually wind the transformer, we will be using full (or maybe half) integral turns, so we will not get exactly 8.33 as recommended above. Further, as mentioned, we want to aim for a slightly higher VOR than V_{INMAX}, to get a gain target slightly more than unity. Because we want to be able to do full regulation over the entire load range restricting ourselves to the left of f_{HI}. So if we manage to set the actual turns ratio N_P/N_S slightly greater than the recommended value above, it would work better. We may have to wind a few iterations of the transformer, to ensure that we stay within f_{LO} and f_{HI} always, with no "mysterious" dropouts.

As mentioned, simulators usually use the inverse of this as their "turns ratio". So, for them we need to enter:

$$N_{SIM} = \frac{V_O}{V_{INMAX}} = \frac{48}{400} = 0.12$$

We also realize that in the simulator, to test it for 900W (at desired output voltage), we simply need to put a resistive load of value

$$\text{Rload} = \frac{V_O^2}{\text{Power}} = \frac{48^2}{900} = 2.56 \text{ Ohms}$$

Finally, looking at **Figure 1.11**, we see that the gain factor = 2 curve intersects with the K=0.5 to give us a frequency ratio of 0.75. That is the ratio of the location of the resonant peak of the applicable resonant curve with respect to f_{HI}. Now, this *ratio* will remain unchanged through the scaling exercise, though the f_{HI} of course has changed, and is now at 145kHz. So assuming the same ratio, the location of the resonant peak, at which we should be able to get 900W at 200VDC input is 0.75 × 145kHz = 108.75kHz. That is the frequency we need to use at the lowest input voltage, to confirm maximum RMS current too.

Note: There is no point doing a simulation at maximum input and f_{HI}, i.e. 145kHz to "validate" our design. That simulation point is meaningless. We will definitely get the desired output voltage because we are assuming that at least the correct turns ratio has been set in the transformer. But we also get almost *any power* we want really! We can keep reducing the load resistor, and we can get 1800W, or 3000W, you name it! Because as we explained, at exactly 145kHz ("f_{HI}"), we essentially have a converter that can deliver as much power as we want, assuming ideal conditions (parasitics not included, or too small to matter).

What really determines that we have arrived at the right "shape" (no overdesign in selecting L and C, and also no insufficient power capability) is the ability of the resulting resonant curve *to bulge just enough* to give us the required gain factor, to compensate exactly for the droop we expect at the input. Very simply, if the input can drop to half its maximum value, we need to arrive at a curve that *just about* reaches Gain = 2 at its peak. If the curve is a bit "too flat", say reaching only Gain = 1.8, we will be unable to get maximum rated power at the minimum input. If it is too peaky even at 900W, we have "overdesign". We will be pulling 900W from the converter at minimum input, but still be well below its peak, so we can lower the frequency quite a bit more, and keep getting more and more power before it rolls off! Not recommended. As we indicated, every LC network has an inherent power capability, which in practice affects the input range we want it to handle. Knowing that fact, we can arrive at an optimal design. If a converter is inadvertently designed for more power by an incorrect choice of L-C components, then of course, if we understand basic power scaling, that has come about because we *have too large a C value, and too small an L value*. Because to double power for example, we always halve the inductance and double the capacitance! Keep in mind that resonant capacitors are more expensive than coils or windings! Overdesign will cost us money. Unnecessary money.

Validating our design through simulations

Before we build what we just designed, we still want to know the RMS currents and so on, to choose our FETs correctly for example. We mentioned that closed form equations in literature, though impressive looking, are all approximate. Why? Because they are all based on the First Harmonic Approximation! So, our effort has been to point us to the right selection of the LC components, and suggest a close-enough operating point, to run simulations at. Thereby we will get far more accurate estimates of RMS currents, for example. So finally in **Figure 1.12**, we present the results of the first-shot simulation, based on the results of our graphical method, iself based on our stunningly simple scaling laws. In the first iteration we are setting the frequency to the predicted 108.75kHz. The input voltage of course is the minimum of 200VDC.

The key waveforms are presented. We see that though we got a noticeably higher output, which is OK, there is something troubling too. We do not have ZVS. The transistor turned ON, into a pedestal of positive current through it, so there will now be crossover losses. We don't need to "see" it here, but they are present. Why is that? There are actually two contributing reasons

a) As pointed out in **Figure 1.9**, there is a possibility that the network appears capacitive at the peak value. We had discussed that.

b) Note that the output, even with diodes present, is higher than the predicted 48V. So we have more gain than we expected. However, if we simulate the "equivalent" AC-AC circuit, we do get exactly the predicted output voltage and thus predicted power. So part of the error is from the FHA itself. What we are also thus learning is that going to a full DC-DC circuit actually helps give us *more* power than we expected on the basis of the "equivalent" AC-AC circuit. The ignored harmonics seem to be helping us! So in that sense, the FHA is a good approximation: it errs on the conservative side. We won't be caught in a situation of less power than FHA predicted. But it also affects the estimate of the best operating point. As a result, we may need to tweak it for a second simulation run.

Since the switch waveforms revealed we were on the capacitive side (left side) of the resonant peak, we should try increasing the frequency in small steps, till we see that little negative glitch (encircled) in the switches just before they turn ON. That means that just before the switch turned ON, current was going reverse: through its body-diode (or through the Schottky diode placed across it: recommended for higher reliability, to avoid shoot-through). So we have ZVS now, because that body-diode currrent flow has forced the voltage across the switch too, to be almost zero at the start of its transition from non-conducting to conducting states. No crossover loss term.

In our case, after tweaking the frequency a bit higher, we arrive in the ZVS region at 115kHz. Note that the voltage is still a bit too high. But the control loop will take care of that by settling at a slightly higher frequency than even 115kHz. We don't need to know what that point exactly is though. The last simulation should give realistic estimates of RMS current ratings than we would normally ever get using closed-form equations available in literature, all based on FHA. For example, this one below for gain of an LLC as a function of frequency, surprisingly often found in literature to be propagating almost endlessly with a mistake/typo, going undetected from one "respectable" reference App Note to another, without anyone actually inserting it into a Mathcad spreadsheet and plotting it to check it, as we did. Here it is, now corrected:

$$\mathrm{Gain}(f) = \frac{1}{\sqrt{\left[\frac{1}{K}\left(1-\frac{1-K}{x^2}\right)\right]^2 + \left[Q\left(x-\frac{1}{x}\right)\right]^2}}$$

Where $x = f/f_{HI}$ *and* $Q = 2\Pi \times (F_{HI} \times LLKG)/RS$
LLKG IS (1-K) L_P AND RS IS THE EFFECTIVE AC LOAD RESISTOR TO BE PLACED IN THE EQUIVALENT SCHEMATIC OF **FIGURE 1.5C** (14.7 OHMS IN OUR DESIGN EXAMPLE ABOVE).

The PoE design Revisited and Improved

How about the PoE design example, using a half bridge? Let us reconfirm the values for that too. We want to ensure we have 25.5 Watts over the range 32-52 VDC. We want an output voltage of 12V. So, the gain (bulge) we want in our resonant curve is 52/32 = 1.625. Suppose we believe that the transformer we build can/will have a coupling coefficient of K=0.9. We are going to see if we need to scale the values used in the kernel of ours: L_P = 57.2µH and C_P = 225.8nF. Hopefully we don't have to scale them too much! Or there is something really wrong with our kernel if it can't even do what we think it can! Of course the graphs were written for any K and any input variation in the Mathcad spreadsheet but here, for starters, we are just sticking to the PoE example we had in Switching Power Supply Design and Optimization, Second edition. In other words, so far we are looking only at K=0.9 and 32-52 VDC input range.

A slight tweak now: in the book, we had set a nominal gain target of 1.05. And recommended, as we do here too, to set the turns ratio accordingly. In the book, as we are going to do here now, we calculated the gain based on an *initial value of 1.05*. So though 52/32 = 1.625, we did not use 1.625, but instead the "additional gain" over 1.05. So, we took the gain factor as 1.625/1.05 = 1.55 instead of 1.625. Either way it is all in the ballpark and we should be very close.

Now, looking at our graphical aid, **Figure 1.10**, for K=0.9 and a gain factor of 1.55, we get a little over 20 ohms. The full Mathcad spreadsheet presented in the Appendix gives an exact numerical value of 20.6 ohms. Which is what we had also used in **Figure 1.9** as the recommended value "Rtrial1". That was the basic value, from which we had then generated multiples of load, just for display purposes.

Now, let us first confirm the power we can get out of this at the maximum input voltage. We are applying 52VDC through a half bridge. So, since this is a half bridge, not a full bridge, its equivalent AC voltage is

$$V_{INMAX\_AC} = V_{INMAX\_DC} \times \frac{2}{\pi} \text{ Volts}$$

$$V_{INMAX\_AC} = 52 \times \frac{2}{\pi} = 33.10 \text{ Volts}$$

$$\text{Peak\_AC\_Power} = \frac{33.1^2}{20.6} = 53.18 \text{ Watts}$$

$$\text{Average\_Kernel\_Power} = \frac{\text{Peak\_AC\_Power}}{2} = \frac{53.18}{2} = 26.6 \text{ Watts}$$

Which is pretty close to the 25.5W we want. We can think of this as a bit higher to account for estimated efficiency degradation from 100%! It is very much in the ballpark and we do not need any scaling for power, just as we had expected and hoped for! Because that is exactly what we had declared our kernel could do in the first place! So, treat this as a revalidation of the kernel.

As for any required frequency scaling, if we look at Switching Power Supply Design and Optimization, Second Edition, we will see that the kernel components were initially chosen such that with K=0.9, f_{HI} would be at 140kHz. Which is a good target for EMI reasons. Let us check

$$\text{fhi\_Kernel} = \frac{1}{2\pi\sqrt{(1-K)L_P C_P}} = \frac{1}{2\pi\sqrt{(1-0.9)\times 57.2\mu \times 225.8n}} = 140.0\text{kHz}$$

So, as we had expected, we don't need a scaling factor for frequency either. The kernel does exactly what the Mathcad spreadsheet said it would.

From **Figure 1.11**, we see that the frequency ratio predicted in this case is 0.38 × f_{HI}. That leads to 0.38 × 140 = 53.2 kHz. The Mathcad spreadsheet predicts 52.25kHz (and 115.1 kHz for the slightly boosted nominal gain value choice of 1.05, instead of 1). Close enough!
We want 12VDC output. But we do want a VOR of about 26VDC (since that is the effective input DC voltage in a half bridge). We recommend to set VOR slightly higher though, just to ensure a gain of 1.05, instead of 1). Anyway, the nominal turns ratio is around 26/12 = 2.167 (we may also want to leave margin for two diode drops if we are not using synchronous rectification and instead a standard four-diode bridge rectifier).
Note: Since we are applying 52VDC to a half bridge, and since the resonant capacitor will end up with a DC level of 26V, it is effectively only applying 26VDC to the bridge. Our VOR is based on that. That is why we got a turns ratio of 2.167 above, NOT 4.33! Beware!
If we want to guarantee 25.5 Watts from a 12V output, the desired resistor we want to put in our simulator is

$$\text{Rload} = \frac{V_O^2}{\text{Power}} = \frac{12^2}{25.5} = 5.65 \text{ ohms}$$

Note that **Figure 1.10** had recommended we use 20.6 ohms as the Primary-side AC load resistor. Had we reflected this to the Secondary side as an equivalent DC load resistor, in a DC-DC converter, we would get as per **Figure 1.5c**

$$\text{Rload\_reflected} = \frac{\pi^2}{8 \times N^2}\text{Rac} = \frac{\pi^2}{8 \times 2.167^2}20.6 = 5.4 \text{ ohms}$$

Now comes the big doubt: that is also very close to the 5.65 ohms calculated by simple V^2/R power cakculations. *So we have a basic question here once again, just when we had started thinking we had mastered resonance!* We ask: why did we even need **Figure 1.10**? We could have just reflected the desired DC load resistor, based on desired power, over to the Primary side as an equivalent AC load resistor (using FHA and turns ratio squared). That would give us almost the same answer as the 20.6 ohms suggestion we got from **Figure 1.10**! So did we even need **Figure 1.10**? Was this all in vain?

Wrong! Resonance is way trickier. Look closely at **Figure 1.10** again. We ask: what if we had wanted our PoE converter to work over a ***1:3.5 input range***? That would be 52VDC to 52/3=14.9VDC. Now, in this case, **Figure 1.10** says, you need to pick an AC load resistor of 49.5 ohms! Not 20.6 ohms or anything close. Intuitively, that is because we want the resonant peak to "bulge" a bit more at the middle, to compensate for the input droop! We see we now have to back off on the maximum allowed R, to create that "bulge". People call this "increasing the Q".

That is correct. Now, if we still want to guarantee 25.5W, we are in big trouble. This converter can do much less now. It only works with a much larger resistor.
At the peak AC voltage of $26 \times 4/\pi = 33.1$V, the peak power would now be much less, since we now have 49.5 ohms across it:

$$\text{Peak\_AC\_Power} = \frac{33.1^2}{49.5} = 22.134 \text{ Watts}$$

$$\text{Average\_Kernel\_Power} = \frac{\text{Peak\_AC\_Power}}{2} = \frac{22.134}{2} = 11.07 \text{ Watts}$$

This is much less that we need. So we need a power scaling factor of

$$\text{Power\_Scaling} = \frac{\text{Desired\_Power}}{\text{Average\_Kernel\_Power}} = \frac{25.5}{11.07} = 2.3$$

The Frequency scaling factor is still 1, since we are OK with f_{HI}=140kHz. That has not changed since we kept K the same, and since the power scaling factor does not affect frequency! So the new, optimum L and C to use are

$$C_P = C_{P\_KERNEL} \times \frac{\text{Power\_Scaling}}{\text{Freq\_Scaling}} = 225.8\text{nF} \times \frac{2.3}{1} = 519.3\text{nF}$$

$$L_P = \frac{L_{P\_KERNEL}}{\text{Freq\_Scaling} \times \text{Power\_Scaling}} = \frac{57.2\mu\text{H}}{1 \times 2.3} = 24.9\mu\text{H}$$

As per **Figure 1.11**, the frequency ratio for K=0.9 and gain factor 3.5 is about 0.34. So we need to smulate close to $0.34 \times f_{HI} = 0.34 \times 140 =$ 47.6 kHz. The inductance will need to be split for simulations as

$$L_{LKG} = (1-K) \times L_P = (1-0.9) \times 24.9\mu\text{H} = 2.5\mu\text{H}$$

$$L_{MAG} = K \times L_P = 0.9 \times 24.9\mu\text{H} = 22.4\mu\text{H}$$

The turn ratio should be 2.167 (or a bit more in the final version, as discussed). In most simulator packages we would need to input the reciprocal: 1/2.167 = 0.462. We are using a half-bridge and the only simulation point that makes sense to verify that we have indeed arrived at around the expected 12V output, with DC load resistor sized for 25.5W@12VDC, i.e. 5.65 ohms, is to run it at the lowest required: 14.9VDC at a projected frequency of around 47.6kHz. The C_P is 519.3nF. Let us do the simulation next.

In **Figure 1.13**, we see the schematic and the results of that in **Figure 1.14**. Note that to firm things up in our minds, we ran both the equivalent AC-AC circuit and the DC-DC converter. We have indicated how we need to change input voltage and output load resistors in going between the two. And since we ran both schematics in one simulation run, the results appear together, in **Figure 1.14**. Since the switch waveform has a very, very slight negative glitch at the end of its current waveform, it is running very slightly capactive. A slight increase in operating frequency will cause the little negative glitch to now appear at the beginning of its current waveform, implying ZVS.

And yes, despite the diode drops, we get almost perfectly 12V at the output of our DC-DC converter, even at 15V input, so our 1:3.5 input variation, hitherto unheard of for LLC converters (and now you know why!), was tackled completely and correctly, indicating that we are spot on in our understanding of resonance finally! Just graphical aids (two figures) and a couple of scaling laws!

The diode waveform always indicates zero-current switching. And it is usually almost half-sinusoidal. The vacant gap between the pulses occurs only when we are switching tr\o the left of f_{HI}. It disappears and starts looking more distorted too, not necessarily looking like ZCS, if we switch to the right of f_{HI} (as Qi, and most others too, typically end up doing, due to random gain targets and random C/L ratios too, not to mention the crippling double-resonator flaw).

Conclusion

With that we conclude this chapter. We have probably arrived! We have learned to scale a kernel to any power level, any frequency, and capable of being the “front-end” converter too, since it is tolerant to quantifiably wide input variations too. No overdesign either. The ability to change coupling too, to almost any desired value, means we can use it even for relatively loosely coupled wireless power transfer systems. All this power, quite literally, through two simple curves, and even simpler scaling laws, hitherto hidden in plain sight. And of course from resonance itself! We have learned how to tame it, and use it. We can start building proper converters now, fearlessly.

But the journey is still not fully over! Because we didn’t banish the devil, it just retreated deep into the details! Now we need a really carefully thought-of control algorithm too! To enable all this really well. And that is going to be proprietary, by all accounts.

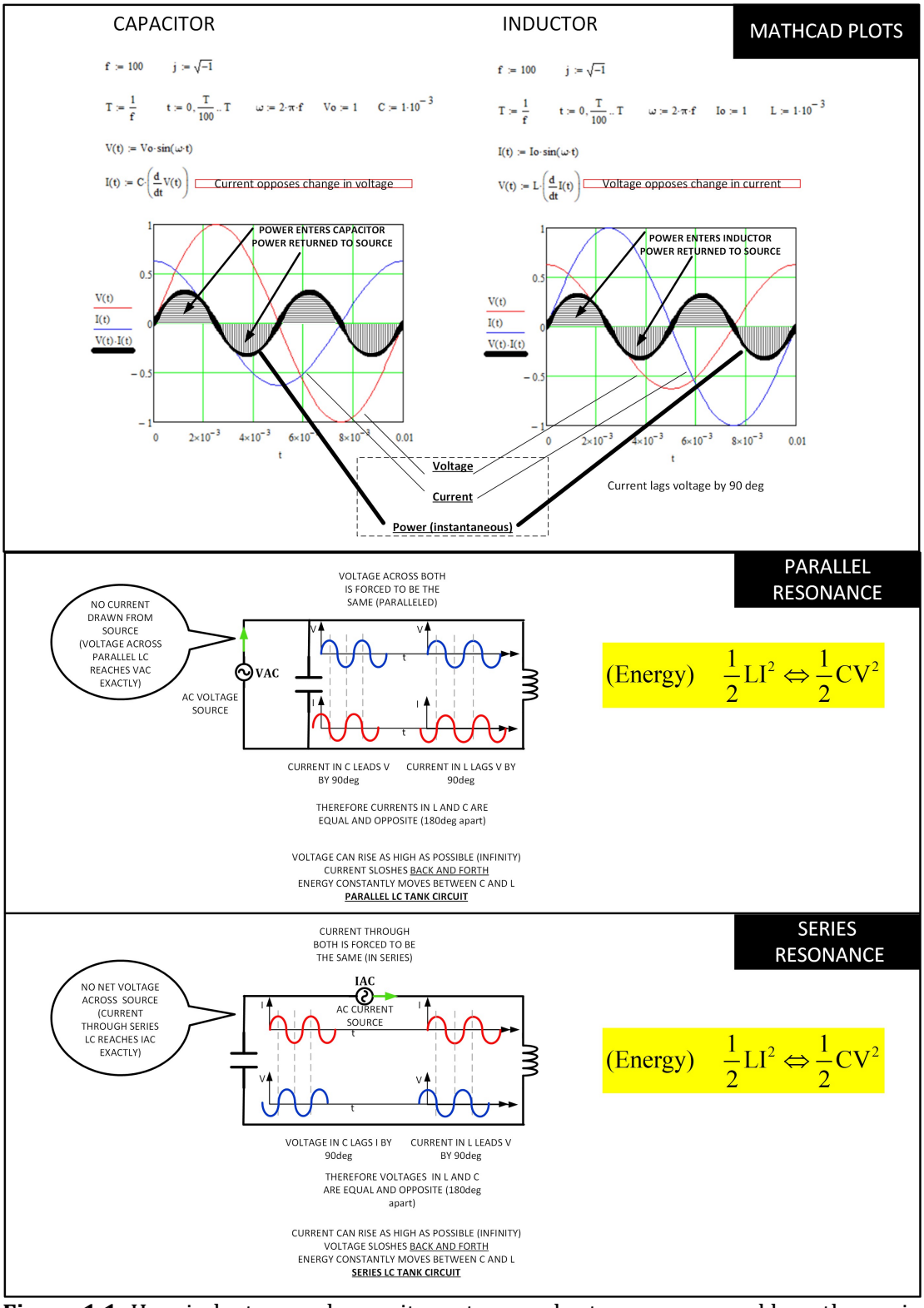


Figure 1.1: How inductors and capacitors store, and return energy, and how they pair up

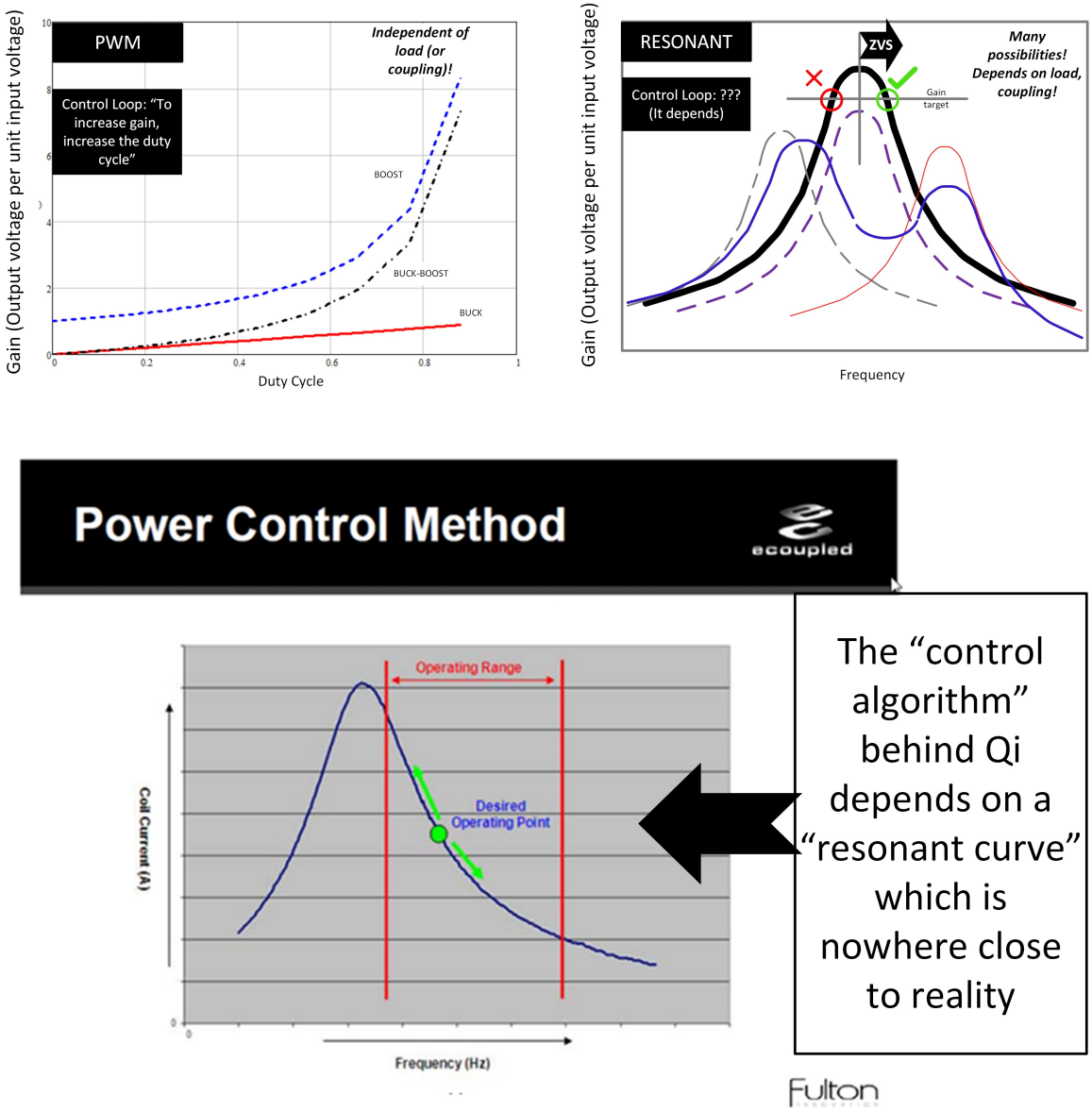


Figure 1.2: PWM versus Resonant Control Loops and Potential Issues

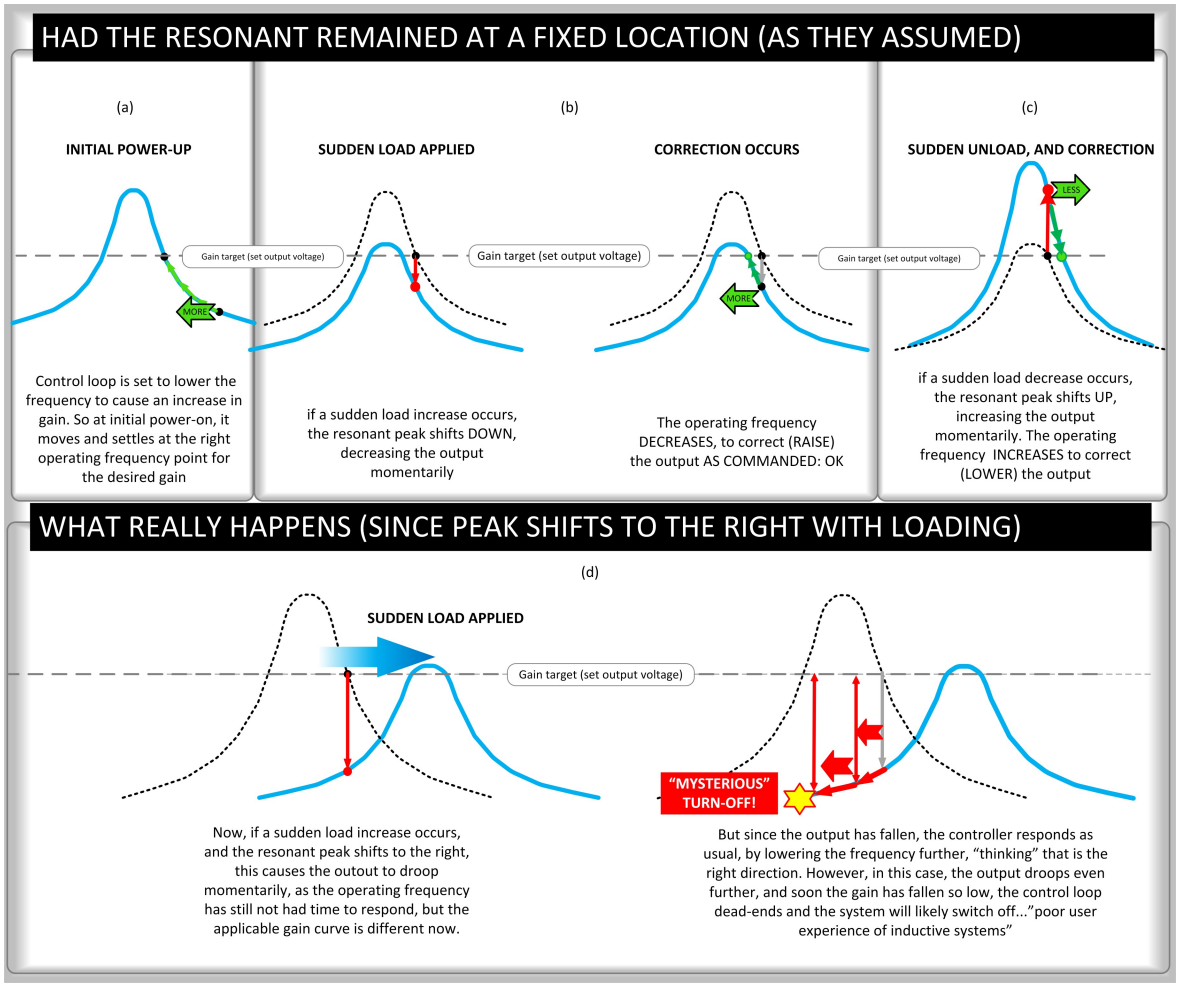


Figure 1.3: Frequency modulation will get stuck if the peak shifts

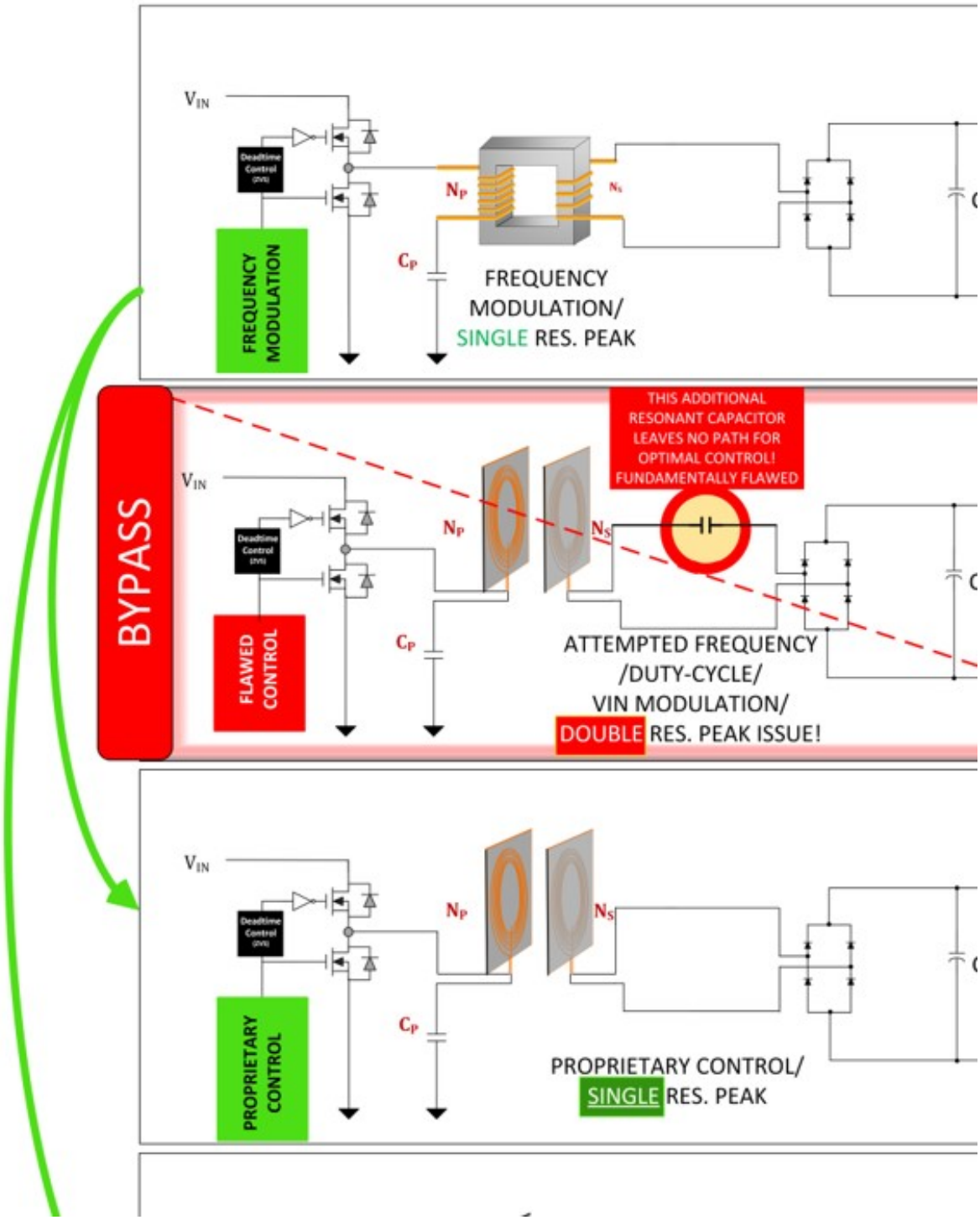


Figure 1.4: LLC to WPT, by avoiding a current prevalent mistake in the latter

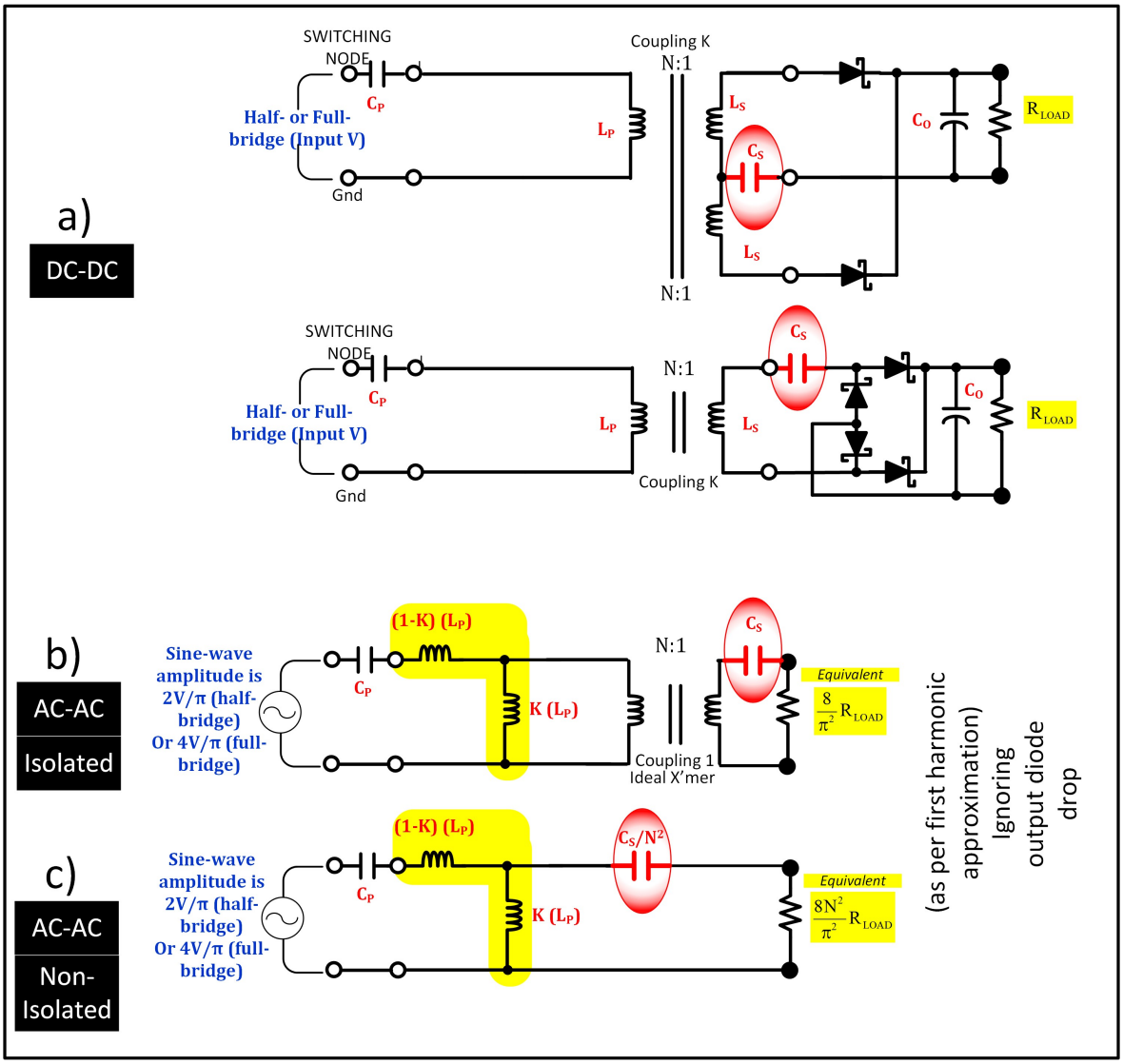


Figure 1.5: Equivalent diagram of LLC and (correct) WPT

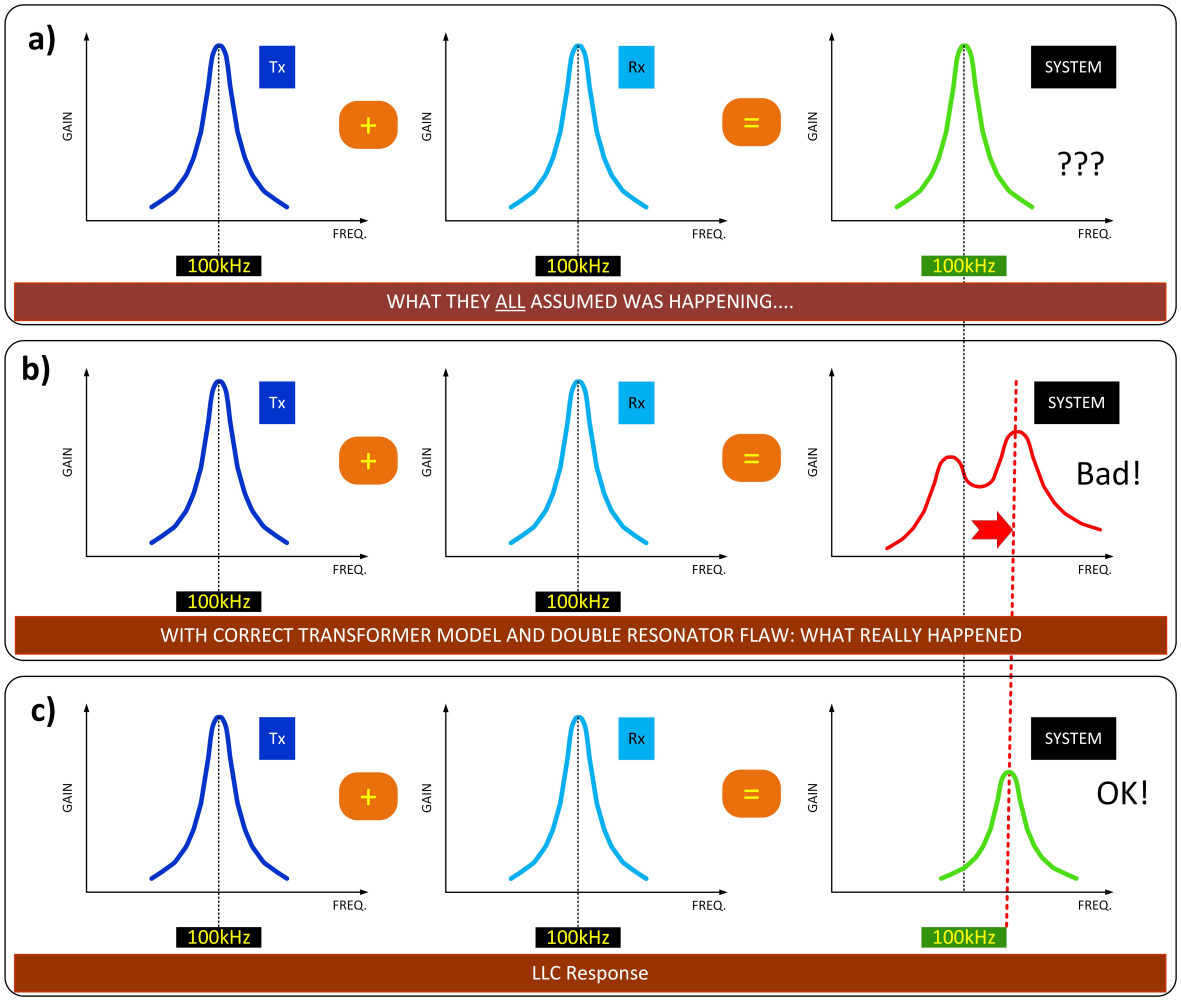


Figure 1.6: What the world assumed was happening in WPT, and what really did

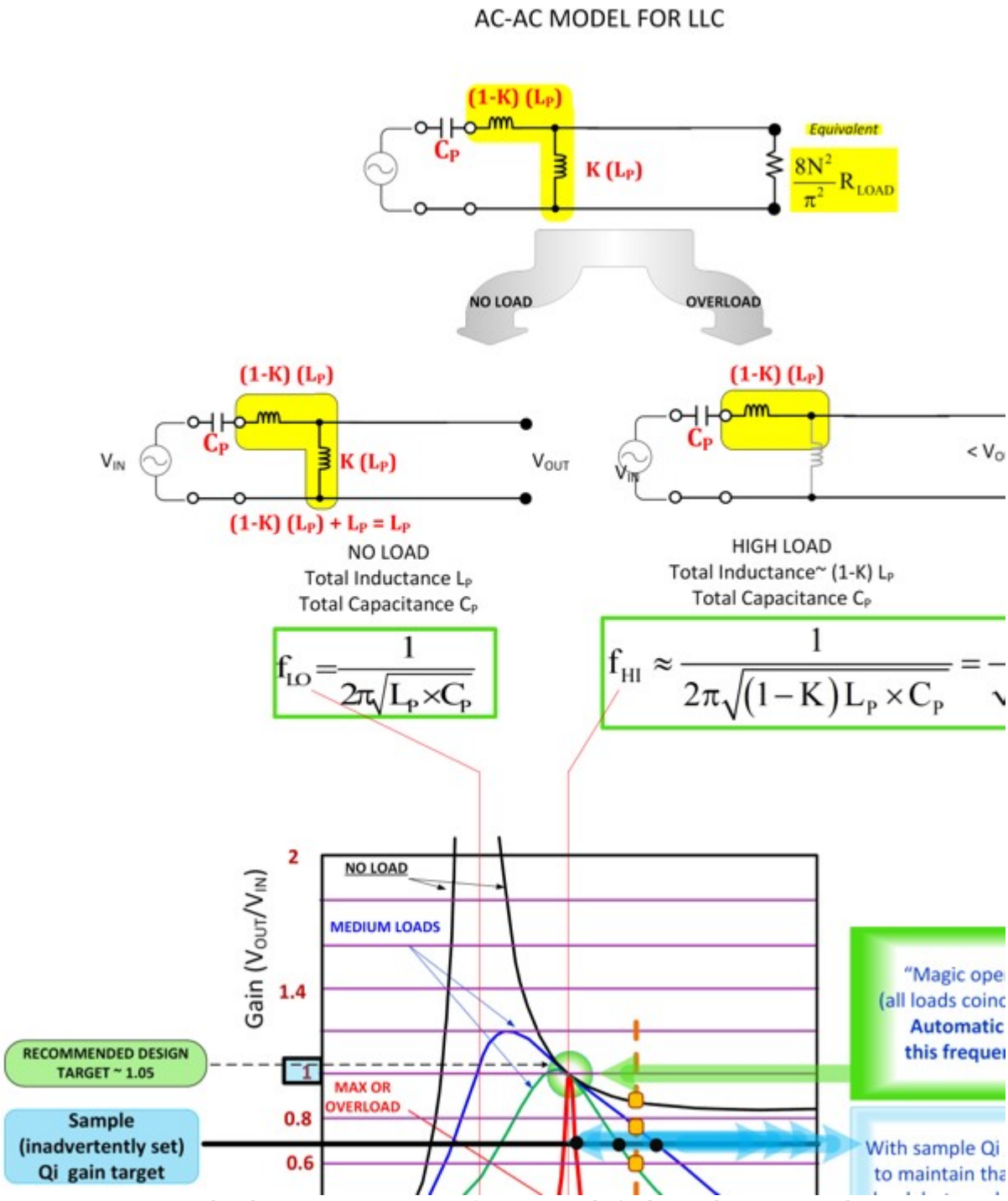


Figure 1.7: Varying load in an LLC, causes a frequency shift, depending on coupling

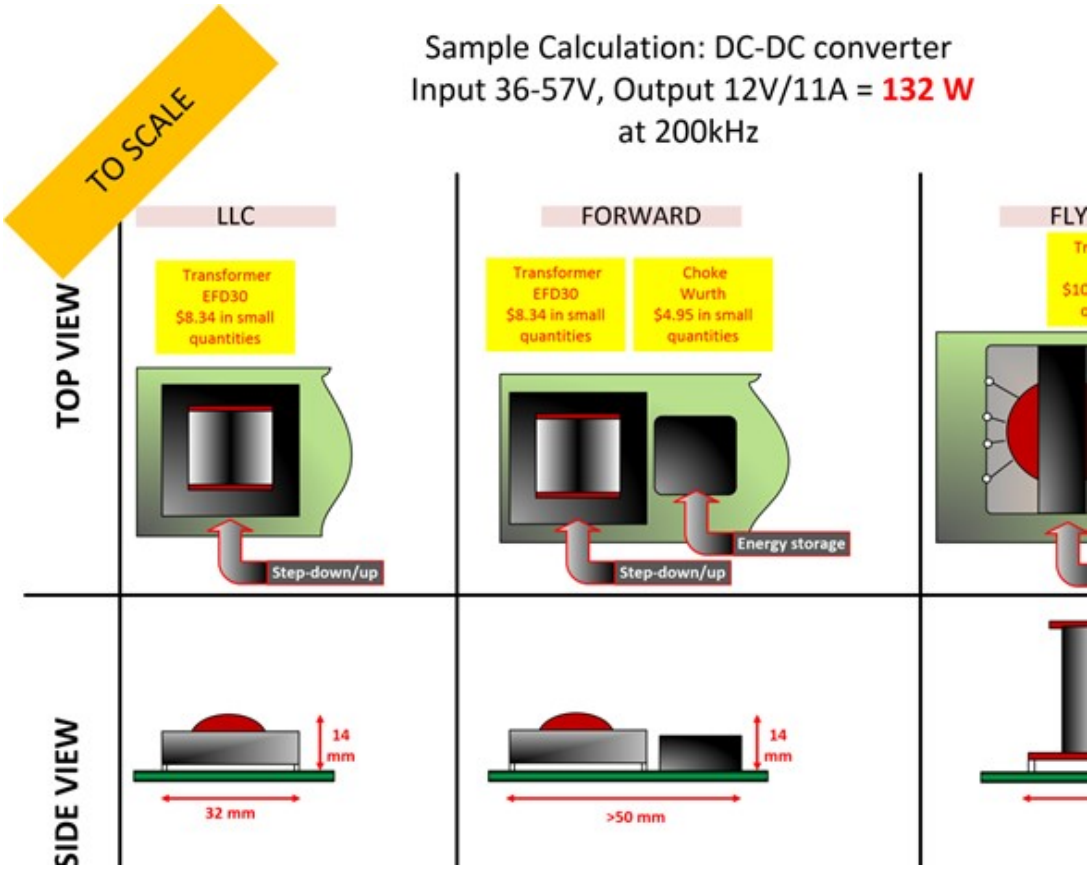


Figure 1.8: Getting a feel for LLC magnetics

PoE Example from Chap 19 of Switching Power Supplies Design and Optimization, 2nd edition.
Requirement 25.5Watts, 32V-52V input, 12V output, Half-bridge
Legacy values used: Lp = 57.2uH, Cp= 225.8nF, K=0.9

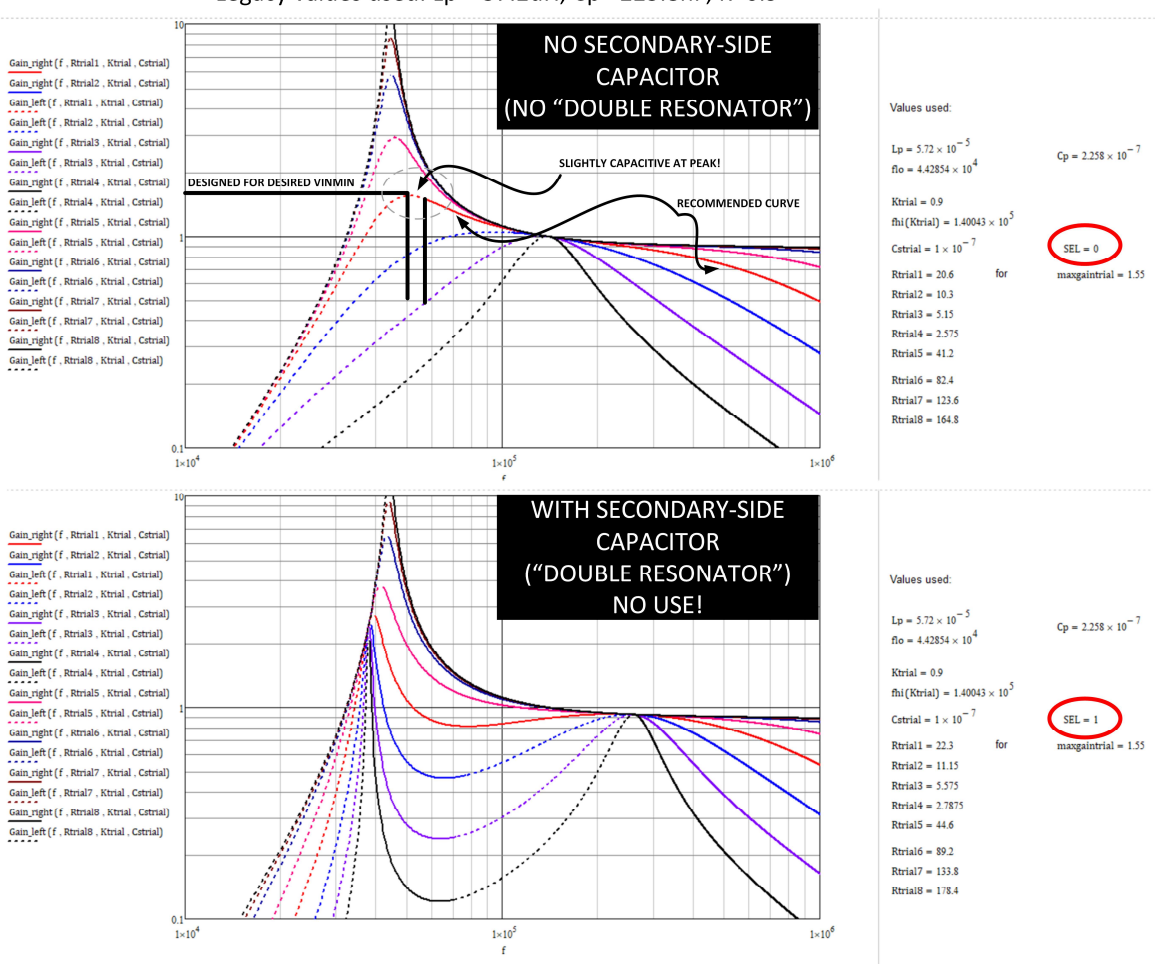


Figure 1.9: Mathcad plot of gain for an LLC with and without C_s

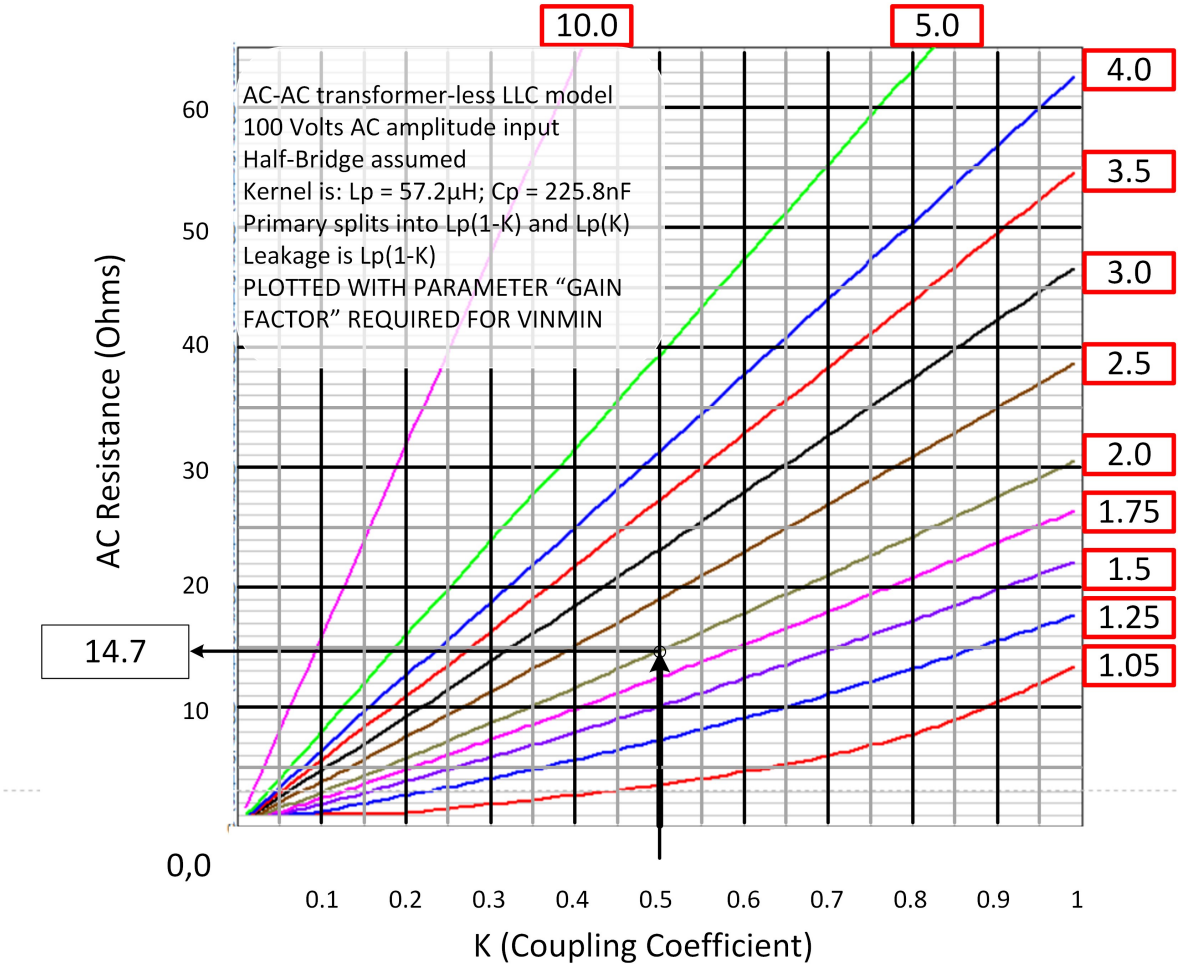


Figure 1.10: Resistance Lookup Aid based on Kernel

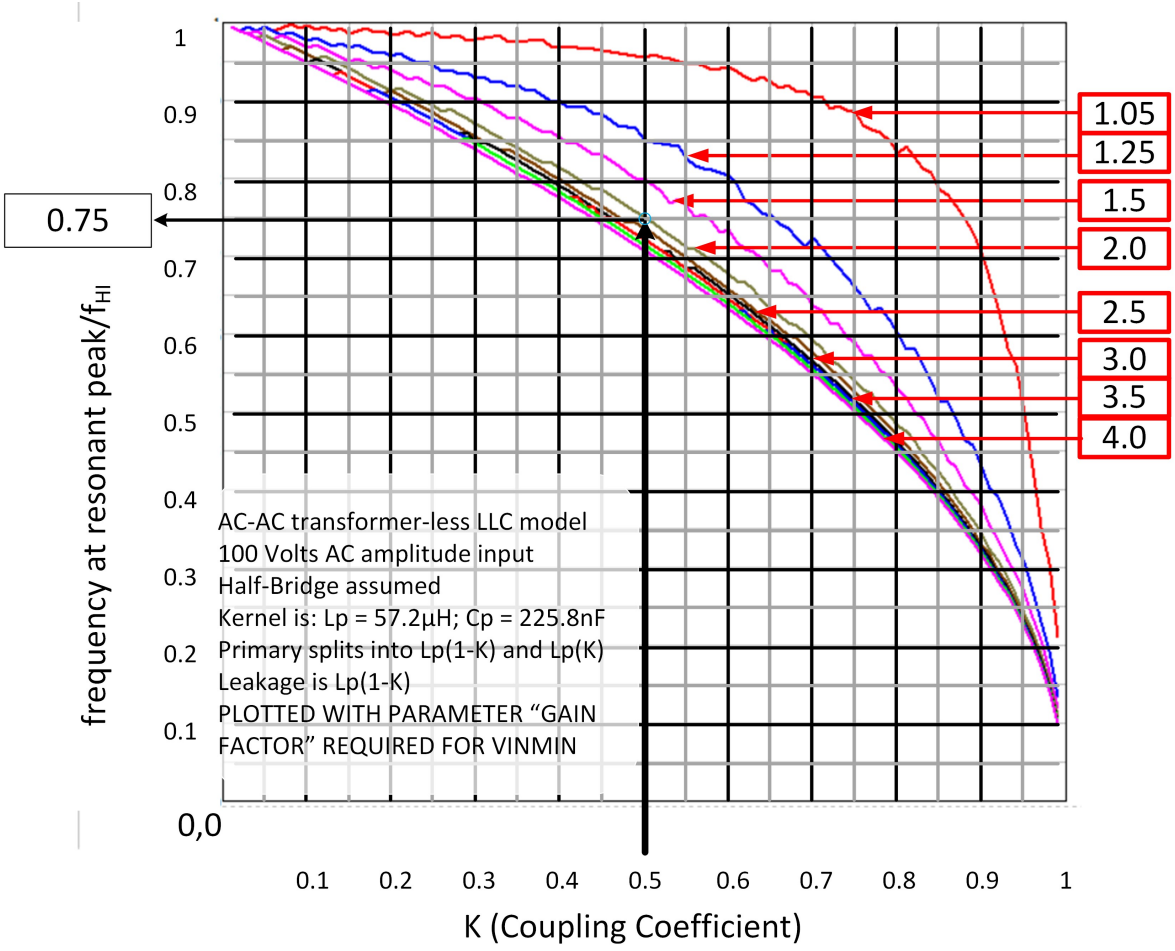


Figure 1.11: The location of the resonant peak

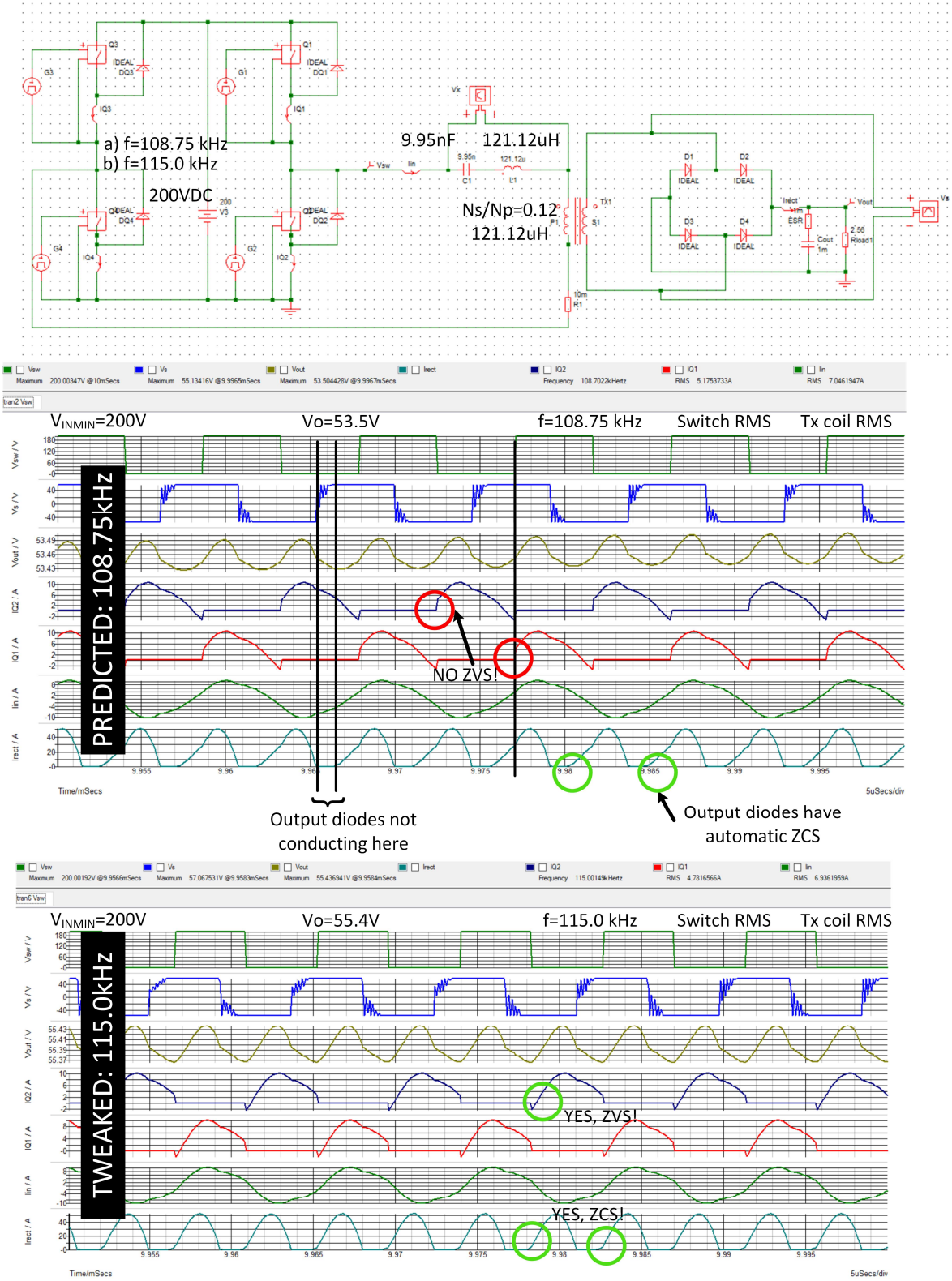


Figure 1.12: Design Validation via Simulations (900W)

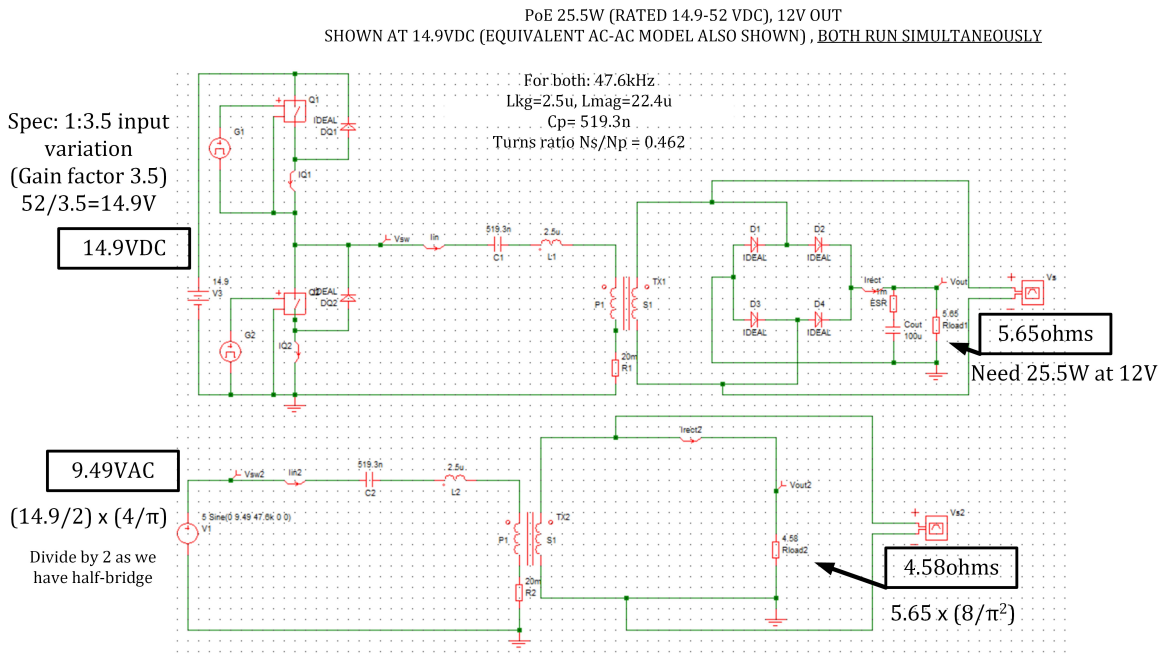


Figure 1.13: PoE converter with 1:3.5 Input Voltage Range (AC-AC equivalent too)

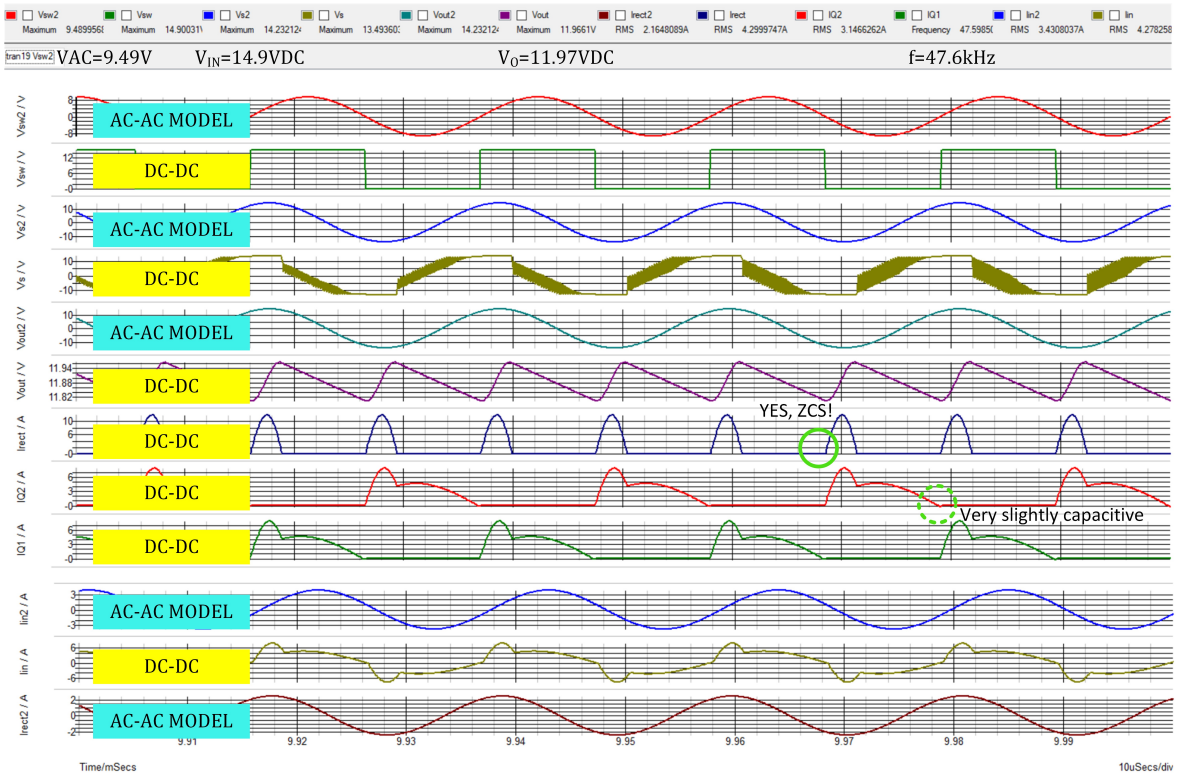


Figure 1.14: PoE converter with 1:3.5 Input Voltage Range (simulation results)

Appendix to Chapter 1

$m \equiv 10^{-3}$ $k \equiv 10^{3}$ $u \equiv 10^{-6}$ $j \equiv \sqrt{-1}$ $n \equiv 10^{-9}$

$Lp \equiv 57.2 \cdot u$ Set a primary inductance value Arbitrarily selected primary inductance (sum of magnetizing inductance and leakage, the split depending on "factor", in turn dependent on K

$K \equiv 0.01 , 0.02 .. 0.99$ Allow coupling K to vary (Typically 1, 1.1..1 to save compute time)

$Ktrial \equiv 0.9$ Set a trial numerical value of K for plotting

MATHCAD SPREADSHEET PAGE 1

$Lmag(K) \equiv Lp \cdot K$

$Llkg(K) \equiv Lp \cdot (1 - K)$

$factor(K) \equiv \frac{K}{1 - K}$ Define a "factor": ratio of magnetizing inductance to leakage

$Cp \equiv 225.8 \cdot n$ Set (primary-side) resonant capacitance Arbitrarily selected resonant cap (on primary side)

$gaintarget \equiv 1.05$ Set gain target for nominal operation: typically 105% gain (frequency to be calculated) on LEFT side of extreme resonant peak

$maxgaintrial := 1.55$ Set desired max gain target trial value (to ensure at resonant peak) (based on Vinmax divided by Vinmin)

$maxgain := 1 , 1.05 .. 10$ In general, set a desired input droop/gain requirement (range)

$Rs := 10 , 10.1 .. 1000$ Reflected (AC) load resistor varies

$factor(Ktrial) = 9$ Ratio of magnetizing inductance to Leakage Inductance, based on trial K selected $Ktrial = 0.9$

$Lmag(Ktrial) = 5.148 \times 10^{-5}$ $Llkg(Ktrial) = 5.72 \times 10^{-6}$ Calculated leakage and mag inductances, based on trial

$Lp = 5.72 \times 10^{-5}$ $Cp = 2.258 \times 10^{-7}$ Other values we have

$flo := \frac{1}{2 \cdot \pi \cdot \sqrt{Lp \cdot Cp}}$ No load resonant freq (entire Lp comes into play, with Cp) $flo = 4.42854 \times 10^{4}$

$fhi(K) := \frac{1}{2 \cdot \pi \cdot \sqrt{(Llkg(K)) \cdot Cp}}$ High frequency resonant freq, based on leakage only (magnetizing inductance literally shorted out by load resistor. Leakage depends on factor, in turn on selected K $fhi(Ktrial) = 1.40043 \times 10^{5}$

$flo = 4.42854 \times 10^{4}$ in Hz

$fhi(Ktrial) = 1.40043 \times 10^{5}$ in Hz

$div := \frac{flo}{100}$ Break up into arbitrary frequency into increments/divisions for plots/calculations

$fmin := \frac{flo}{10}$ Set minimum frequency for plots/calculations

$fmax := flo \cdot 100$ Set maximum frequency for calculations

$f := fmin , fmin + div .. fmax$ Allow frequency to vary, based on chosen range and increments

$\omega(f) := 2 \cdot \pi \cdot f$ Angular frequency

$SEL := 0$ Selector Switch for choosing secondary side (reflected) capacitor Cs (1 for include Cs, 0 otherwise)

$Cs := \frac{Cp}{100}, \frac{Cp}{99} .. Cp \cdot 100$ Range of Cs

$Cstrial := 100 \cdot n$ Trial value for Cs (if selected)

Impedances

$X1(f, K) := j \cdot Lmag(K) \cdot \omega(f)$ Impedance of magnetizing inductance X1

$Xs(f, Rs, Cs) := Rs - \frac{j}{Cs \cdot \omega(f)} \cdot SEL$ Sum of Rs and Cs reflected impedances

$X2(f, Rs, K, Cs) := \frac{X1(f, K) \cdot Xs(f, Rs, Cs)}{X1(f, K) + Xs(f, Rs, Cs)}$ Impedance of X1 in parallel with reflected load resistor Rs

$X3(f, K) := \frac{1}{j \cdot Cp \cdot \omega(f)} + j \cdot Llkg(K) \cdot \omega(f)$ Impedance on leakage and resonant cap in series

$Z(f, Rs, K, Cs) := X3(f, K) + X2(f, Rs, K, Cs)$ Impedance of entire network..

$Vin := 10, 11 .. 30$ In general, allow (amplitude of) input Vin to vary too

$I(f, Vin, Rs, K, Cs) := \frac{Vin}{Z(f, Rs, K, Cs)}$ Current entering entire network, based on impedance seen at entry point

$IL(f, Vin, Rs, K, Cs) := I(f, Vin, Rs, K, Cs) \cdot \frac{Xs(f, Rs, Cs)}{(Xs(f, Rs, Cs) + X1(f, K))}$ Current component that splits into magnetizing inductance

$IR(f, Vin, Rs, K, Cs) := I(f, Vin, Rs, K, Cs) \cdot \frac{X1(f, K)}{(Xs(f, Rs, Cs) + X1(f, K))}$ Current component that splits into reflected load resistor segment

$VR(f, Vin, Rs, K, Cs) := IR(f, Vin, Rs, K, Cs) \cdot Rs$ Voltage across reflected load resistor

$Gain\_left(f, Rs, K, Cs) :=$
- $v \leftarrow 1$
- $phase \leftarrow arg(Z(f, Rs, K, Cs))$
- $vr \leftarrow VR(f, 1, Rs, K, Cs)$
- $|vr|$ if $phase \le 0$
- j otherwise

Output voltage gain (input set to unity) where impedance of entire network as seen by input source is "capacitive", since argument of Z, i.e. its phase angle is negative, which is what we get when any impedance is predominantly capacitive

$Gain\_right(f, Rs, K, Cs) :=$
- $v \leftarrow 1$
- $phase \leftarrow arg(Z(f, Rs, K, Cs))$
- $vr \leftarrow VR(f, 1, Rs, K, Cs)$
- $|vr|$ if $phase > 0$
- j otherwise

Output voltage gain (input set to unity) where impedance of entire network as seen by input source is inductive

$Gain(f, Rs, K, Cs) :=$
- $v \leftarrow 1$
- $vr \leftarrow VR(f, 1, Rs, K, Cs)$
- $|vr|$

Overall gain curve

MATHCAD SPREADSHEET PAGE 2

MATHCAD SPREADSHEET PAGE 3

We had set gain targets

gaintarget = 1.05 The set gain target for nominal operation at calculated frequency below: we want to have a gain of of 1.05

maxgaintrial = 1.55 The gain target for worst case Vinmin operation

Calculate load resistor to yield desired gain target at resonance : calculate gain, starting from high frequencies, with ever-increasing loads, and break where gain (at resonance) just exceeds gaintarget

```
Rset(K , maxgain , Cs) :=  for   R ∈ 1 , 1.1 .. 4000
                             for   frev ∈ 0 , div .. 9·flo
                               freq ← 10·flo − frev
                               gain ← |Gain(freq , R , K , Cs)|
                               break  if  gain > maxgain
                             break  if  gain > maxgain
                           R
```

```
fres(K , maxgain , Cs) :=  for   R ∈ 1 , 1.1 .. 4000
                             for   frev ∈ 0 , div .. 9·flo
                               freq ← 10·flo − frev
                               gain ← |Gain(freq , R , K , Cs)|
                               break  if  gain > maxgain
                             break  if  gain > maxgain
                           freq
```

```
fnom(K , maxgain , Cs) :=  R ← Rset(K , maxgain , Cs)
                           for   frev ∈ 0 , div .. 9·flo
                             freq ← 10·flo − frev
                             gain ← |Gain(freq , R , K , Cs)|
                             break  if  gain > gaintarget
                           freq
```

Rset(Ktrial , maxgaintrial , Cstrial) = 20.6

fres(Ktrial , maxgaintrial , Cstrial) = 5.22567×10^{4} For trial values

fnom(Ktrial , maxgaintrial , Cstrial) = 1.15142×10^{5}

Ktrial = 0.9

maxgaintrial = 1.55

Cstrial = 1×10^{-7} but whether it is selected or not: SEL = 0

For plotting range of values

Ktrial1 := Ktrial

Rtrial1 := Rset(Ktrial , maxgaintrial , Cstrial)

Rtrial2 := 0.5·Rset(Ktrial , maxgaintrial , Cstrial)

Rtrial3 := 0.25·Rset(Ktrial , maxgaintrial , Cstrial)

Rtrial4 := 0.125·Rset(Ktrial , maxgaintrial , Cstrial)

Rtrial5 := 2·Rset(Ktrial , maxgaintrial , Cstrial)

Rtrial6 := 4·Rset(Ktrial , maxgaintrial , Cstrial)

Rtrial7 := 6·Rset(Ktrial , maxgaintrial , Cstrial)

Rtrial8 := 8·Rset(Ktrial , maxgaintrial , Cstrial)

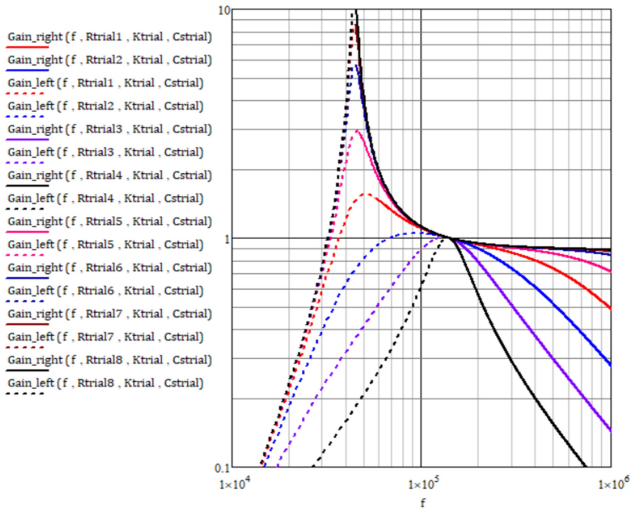
Rtrial9 := 10·Rset(Ktrial , maxgaintrial , Cstrial)

Can use closed for of LLC here just to check if desired

$$x(f , K) := \frac{f}{fhi(K)} \qquad Q(K , Rs) := 2 \cdot \pi \cdot \frac{fhi(K) \cdot Llkg(K)}{Rs}$$

$$GainCF(f , Rs , K) := \frac{1}{\sqrt{\left[\frac{1}{K} \cdot \left(1 - \frac{1 - K}{x(f , K)^2}\right)\right]^2 + \left[Q(K , Rs) \cdot \left(x(f , K) - \frac{1}{x(f , K)}\right)\right]^2}}$$

MATHCAD
SPREADSHEET
PAGE 4



Values used:

$Lp = 5.72 \times 10^{-5}$ $Cp = 2.258 \times 10^{-7}$

$flo = 4.42854 \times 10^{4}$

Ktrial = 0.9

$fhi(Ktrial) = 1.40043 \times 10^{5}$

$Cstrial = 1 \times 10^{-7}$ SEL = 0

Rtrial1 = 20.6 for maxgaintrial = 1.55

Rtrial2 = 10.3

Rtrial3 = 5.15

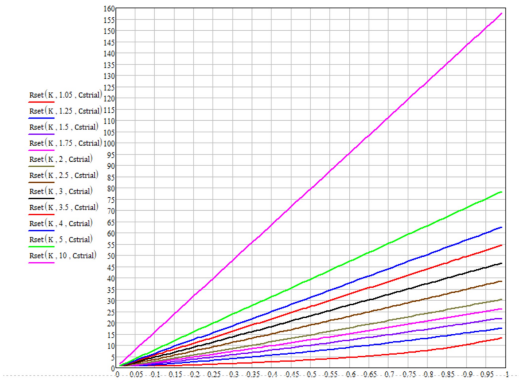
Rtrial4 = 2.575

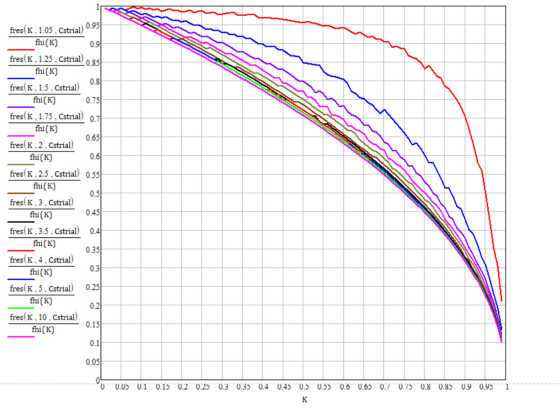
Rtrial5 = 41.2

Rtrial6 = 82.4

Rtrial7 = 123.6

Rtrial8 = 164.8





Chapter 2
Dual Active Bridge

Introduction

This chapter opens with a basic question. One that clarifies and re-iterates relevant aspects of Chapter 1 of A to Z, Second Edition, as a refresher, before we feel comfortable enough to head into more exotic and complex topologies like the phase-modulated full-bridge (PMFB) and the dual active bridge (DAB). The first question is: *why do we need different topologies*?
Various answers to that could be as follows, pointing to certain topologies as candidates; in the process indicating the huge promise that DAB, in particular, offers.

A) Setting an output voltage. For example, step down (Buck), step up (Boost), step-up/down (Buck-Boost). These are simple inductor-based topologies, with simple DC transfer functions, (independent of load current, to a first degree, in continuous conduction mode). We could also invoke the 4-switch Buck-Boost (basically a cascaded Buck and Boost), as discussed in Chapter 9, since that corrects the inherent polarity inversion that is unfortunately part of the basic Buck-Boost topology. Yes, for all three fundamental topologies, we have a simple direction of control or correction ("DoC") too: *to increase the output, increase the duty cycle*. Can't go wrong!

B) Isolation. Which brings in transformer-based topologies (Flyback, Forward, Half-bridge, Full-bridge). In these topologies, we can use the transformer turns-ratio to step up or down in addition to the underlying action accruing from the fundamental topology that it is based upon. For example, the Forward converter is based on the Buck, an inherently step-down topology, but we can use turns-ratio of the transformer to offset that voltage-lowering action, and instead even provide overall step-up action if desired. The Flyback is based on the Buck-Boost topology, but in a typical AC mains application (where the need for isolation is most prominent), the Flyback is typically operated in step-down mode. So not only do we have a turns-ratio to step-down (further), but the underlying Buck-Boost is also usually, but not necessarily, operated in step-down mode.

Note: Hence we get the concept of "V_{OR}", or reflected output voltage ($V_O \times n$) as explained in Figure 3.2 of A to Z, Second Edition, for a Flyback. We typically set it to about 130V, to stay within the typical 600V voltage rating of cheaper switches. In other words, as far as the Primary-side (underlying Buck-Boost stage) is concerned, it is taking in the rectified AC voltage, which may be say, 340 VDC (at 240 VAC), and is converting that to an "unseen" 130 VDC. Its duty cycle is based on the standard Buck-Boost DC transfer function, but using an input of 380V and a "virtual" DC output of 130V. That virtual DC output is further stepped-down by the turns-ratio to the desired (usable) DC output of say, 5V or 12V. For example, with a turn ratio of 10, we would get 13V from 130V. Of course, we then have to account for a typical ~1V of diode forward drop just after that, resulting in about 12 VDC. Or we often use half-integral turns too.

C) Higher efficiency. That quest leads us either to extremely fast switches, such as modern GaN (Gallium Nitride) or SiC (silicon carbide), or to inherently more efficient topologies such as resonant and quasi-resonant topologies. Fast switches help reduce crossover times (at the expense of higher EMI), and hence reduce switching ("V-I-t" crossover) losses, and are therefore helpful at very high frequencies of around 2MHz and above. But at the typical low frequencies (~70-150kHz) commonly associated with mains AC power converters, and those involved in WPT (wireless power transfer), it is preferred to use topologies such as the LLC topology and the phase-shifted full bridge (PSFB), because these enable (but not necessarily guarantee), zero-voltage switching (ZVS), and/or zero-current switching (ZCS). Hence the term "lossless switching" or "soft switching." Further: if we are doing soft switching anyway, why would we ever need ultra-fast switching devices?

Note: A lot goes into being able to extract or guarantee soft switching, but fundamentally, for ZVS for example, we make use of the fact that any actual inductance, or a switched network which appears "inductive" (i.e. we should be switching to the right side of its resonant peak), tends to force current to freewheel through the body diode of the switch (or through an external paralleled diode, preferably). So if we turn that switch ON after a short "deadtime" (typically 100-200ns), we can hopefully ensure that diode conduction has already occurred, in effect, bringing the voltage across the switch to almost zero when we actually turn it ON, so there is no V-I crossover term: *the "V" has already crossed over, before "I" even starts to move. No "V-I overlap" in effect. And that ensures ZVS* or lossless switching.

D) Bidirectionality. A modern emerging trend is to charge batteries from an available energy source, such as solar power panels ("PV", i.e. photo-voltaic cells), but under certain conditions, say during the night, to draw out some of the energy residing in the battery and deliver it back to the "grid" for example. Hence the search for bidirectional converters and the recent emergence of the DAB converter. Of course, just changing the direction of flow occurs in a synchronous Buck too, with a "pre-biased output", and that can transform the forward-Buck into a reverse-direction-Boost, as was discussed in Chapter 9 of A to Z, Second Edition. But to do something like that *with isolation included*, points to the DAB converter.

In fact, both the LLC and the DAB can combine all the above "requirements": Setting any output voltage (irrespective of turns ratio), isolation, higher efficiency and bidirectionality. A bidirectional LLC will pose some other "unanticipated" issues, especially based on the rather misguided "double tuning" approach used in some WPT implementations today, but it does hold future promise *if done right*. It seems a longer way off though. So, for now, only the DAB is considered within the scope of this chapter.

The DAB converter appears inherently symmetrical, and thus attractive as a likely/promising candidate for bidirectionality. And it is unique! Remarkably so, to be called a mere derivative, or combination/composite, of the three "fundamental" topologies, the Buck, the Boost and the Buck-Boost. It was probably inspired initially by the PSFB when used with a full-bridge synchronous rectification stage. Perhaps someone out there observed that there were now eight switches, in effect two full-bridges, positioned on either side of a transformer, so it must be possible, in principle, to exchange roles: the Secondary could become the Primary, and send current in a reverse direction. In fact, that is possible with the DAB, but is also just the easiest part of its design!

The most difficult part of actually designing a DAB, is not in terms of basic control in effecting a change in the direction of power flow, but even in a given direction, *how to design it optimally.* So, first question for the DAB is: do we even understand it *as a "topology"*?
But question at ground zero is: do we even understand why it is *stable*? To be ever considered a topology?

Stable Conditions to Topologies

As discussed in Chapter 1 of A to Z, Second Edition, the essence of any topology is that it is *stable*! And without a control loop too! So, under steady conditions, i.e. a steady input and a steady load, we must *automatically* arrive at a certain "stable condition": *one which ensures inductor "reset"*. As indicated, this must happen irrespective of any clever control-loop strategy, the intended purpose of which should only be to *add* precision to the desired output, not to make viable any inherently unviable topology!

Unfortunately, as regards to "stable condition", there is some confusion in related literature as to what even the related precondition, "inductor *reset*" means. Reset does *not* mean that the inductor current return to *zero* at the start/end of every switching cycle. That does happen in discontinuous conduction mode (DCM), quite automatically, in the three fundamental topologies. But in general, it only means that the start and end inductor current values *be precisely the same*: positive, negative, or zero, as the case may be. This is applicable to both CCM and DCM, or to synchronous converters (with negative currents possible). And thus, each cycle can become a simple repeat of the previous cycle. Otherwise there would be constant current/flux "staircasing", which is potentially destructive, eventually leading to inductor/transformer saturation and destruction of switches. To avoid that situation is a major objective, even the areas of PSFB and DAB, but surprisingly insufficiently addressed in literature.

In hindsight, the reason why the Buck, Boost and Buck-Boost emerged as fundamental topologies, is what happens at initial power-on, without any user intervention. See Figure 1.11 of A to Z, Second Edition for a Buck-Boost as an example. Its output voltage increases automatically, progressively increasing the magnitude of the down-slope of the inductor current, which is dependent on V_O, to force inductor reset ultimately. Very naturally! Basically, after a certain number of cycles, depending mainly on the value of the output capacitance and any load connected there, the down-slope increases to exactly the point where it matches the up-slope (in terms of magnitudes). The increment and subsequent decrement of current in each cycle are numerically identical, i.e. $|\Delta I_{ON}| = |\Delta I_{OFF}|$. And that results in a "stable condition" automatically: i.e. one corresponding to *inductor reset* (every cycle). It is a repetitive state. Hence it is a topology. Now, to precisely control the output, we can add active duty cycle control via a control loop.

As shown in Figure 1.9 of A to Z, Second Edition, this also corresponds to the statement that the net *VoltSeconds in any cycle is zero*—a result of the simple inductor current law $V=L\,\Delta I/\Delta t$. Because we realize that to arrive at the same "|ΔI|" during ON and OFF durations, we demand that numerically, $V_{ON} \times t_{ON} = V_{OFF} \times t_{OFF}$. And the reason that this equality can occur in the first place, is that V_{OFF} increases steadily from its initial value of zero, on account of the increasing V_O, finally settling down to the steady, holding value of V_O corresponding to a "stable condition".

So that is the first thing we need to be assured of, in dealing with *any* topology, *especially the DAB*. Because things get very complicated here. Unlike simpler topologies, the DAB has not just two obvious segments of applied VoltSeconds (e.g. ON and OFF times), but even in the *simplest* case that we will discuss here, **eight** segments, none of which can definitely even be identified as distinct "power delivery" or "freewheeling/reset" durations. No simple ON and OFF durations! We can't even usually say which of these eight segments is definitely one with an up-slope (of inductor current) and which is a down-slope, leave aside assuring ourselves that the net increments and decrements of the eight current segments are always going to add up (signs included) to zero, as demanded in any "stable" condition. And if that is even assured, or guaranteed, can we at least answer: for what exact output voltage does that happen? In other words, what is the "simple" DC transfer function of the DAB? If any? Remember that for the three fundamental topologies, a *certain relationship between input and output voltages always needed to occur, for "reset" to result*. So, what if any, is that output voltage, for a DAB?

A "simple" DC transfer function expression is actually, for good reason, completely missing in related literature. There is none! Yet this surprising fact seems to have gone fairly unquestioned or unnoticed. Instead, the DAB seems to have spawned an overwhelming number of dissertations/theses, which on closer examination were "validated" purely by simulations and graphs under very specific conditions, rarely any built-to-design converter. All research papers seem to be focusing on how to control the output, which as we indicated, is helpful when we are already sure we have a "topology" on hand at least. But now we have exotic DAB control methods: SPS (single phase control), DPS (double phase control) and even TPS (triple phase control). All intended to be efforts to tame the powerful DAB, quite obviously. But the DAB, with so many independent variables, phase combinations, and load combinations to choose from, seems also to be *always* stable. Inherently so. For any output setting, any phase combination. Its inherent stability was however almost assumed as a fait accompli. But in reality, nature made that come true, through its inherent self-stabilizing tendency, as we had pointed out in Chapter 1 of A to Z, Second Edition too.

Fast forward: *we can in fact design the DAB for any output voltage, irrespective of the turns-ratio effect from the transformer. Fundamentally, it offers* Buck-Boost action, but one not so obvious, or predictable, based on any simple DC transfer function. Hence arises another question: *what really determines the output voltage*, if reset is assured irrespective of the output anyway?

We realize that this is not a resonant system where we could carefully tailor the resonant bump to provide us a step-up action, and use turns ratio to create a step-down action if desired. We are in effect saying that a DAB can, without taking recourse to resonance, or even turns ratio, produce both step-up and step-down action on demand, and still be assured that inductor reset will always occur! Unconditionally. A truly powerful topology if so, necessitating a complete relearn of classical power conversion on occasion. At least a serious re-clarification of basic concepts.

So back to some more basics: is there is at least a certain, obvious simple *direction of correction* ("DoC") to the DAB, such as "increase say, x or y or phase difference, to increase the output voltage"? In all three fundamental topologies we had a simple rule: increase the duty cycle and thereby increase the output. But in fact, no simple DoC exists for the DAB. The DAB is far more complicated. It "works", but even to get our heads around it, requires an astonishing amount of pre-study, which is what we are embarking on here.

The "underlying topology" of the DAB, is a buck or a boost action, but not the traditional Buck-Boost topology. It merely steps down or up, that is all we are saying here. As mentioned, we can use transformer turns-ratio on top of that underlying action. But unlike the Flyback, the DAB is also bidirectional, utilizing very small magnetics: more like an LLC, or a Forward converter with no output choke of the latter, but possibly with a much smaller, but additional discrete inductor like we prefer in the PSFB too. And it can all be very efficient too, like a PSFB or an LLC. But unlike an LLC, it can also be fixed-frequency, for obvious benefits in meeting regulatory limits for electromagnetic emissions.

Our ultimate mission: we will create a Mathcad spreadsheet to match simulations using SIMPLIS. We thus create a "kernel", and then, using our powerful scaling laws, we will scale that to any power and frequency level, not to mention any input voltage level too. In just a few lines.

The DAB Schematic

From **Figure 2.1**, we see that we have a Primary side, with four FETs, labelled Q1, Q2 (corresponding to switching node "a", or "Vsw_a"), and Q3, Q4 (switching node "b" or "Vsw_b"). Note that all odd-numbers, Q1 and Q3 here, are *high*-side FETs of their respective half-bridges. Similarly, on the Secondary side we have four FETs, labelled Q5, Q6 (switching node "c" or "Vsw_c") and Q7, Q8, (switching node "d" or "Vsw_d"), with Q5 and Q7 being the respective high-side FETs of the two Secondary-side half-bridges. Note the dots on the transformer windings, so in effect, the half-bridge consisting of Q5 and Q6 is the "cousin" or "next-door neighbor" of the half-bridge consisting of Q1 and Q2. The phase lag between the Primary and Secondary sides, is based on the lag between Q1 (master) and Q5 (its cousin). So, the high-side FET of the half-bridge consisting of Q1 and Q2, i.e. Q1, is the designated "master" always, and in fact all the other half-bridges are set up to lag this, as shown in **Figure 2.2**. The lag between the two Primary-side half-bridges (measured between the high-side FETs Q 1 and Q3) is D1 × π (in radians), and since D1 can be maximum 1, the maximum shift is π radians or 180°. D1 thus corresponds to what we may refer alternately to as "Angle 1", or "intra-bridge" phase difference, or "*Primary lag*".

D1 = 1 (i.e. 180°) corresponds to maximum power capability from the Primary Full-bridge, since whenever Q1 turns ON, Q3 is OFF, which means Q4 is ON whenever Q1 is ON, and full power is possible (not necessarily delivered). Similarly, when Q1 turns OFF, i.e. Q2 turns ON, it finds Q3, ON, for the entire duration too, and now current flows through the full-bridge network in a reverse direction, doing its bit to try and ensure inductor/transformer reset. Whether that happens or not is an entirely different matter, as we will explain.

Similarly, the lag between Q1 and Q5 is D2 × π, in radians. D2 corresponds to what we may refer alternately to as "Angle 2", or "intra-bridge" phase difference, or "Secondary lag". Like D1, D2 too can be a maximum of 1, so the Secondary essentially lags the Primary by a maximum of 180°. Indeed, it can be shown that under these conditions, power flows from Primary to Secondary. But *if D2 is "greater than 1"* (i.e. angle greater than 180°), or alternatively expressed, "is negative" (-180° to 0, which is actually the same thing as 180° to 360°), in effect we can think of Q3 as *leading* (not lagging) Q1. The power direction will then simply reverse. *But if D2 is between 0 to 1, the direction is always Primary to Secondary.*

To keep the number of variables limited, we are implicitly also assuming that the lag between Q1 and Q3, i.e. D1, *is the same as the lag between Q5 and Q7*. That means both full-bridges have the same inherent phase-shift. So, the "intra-bridge phases" are the same. But, measuring the lag of Q7 with respect to the master, Q1, we get (D1 + D2) × π.

We mentioned that when the Q1 and Q3 are offset by π (i.e. 180°), the Full-bridge is "capable" of delivering maximum power. But to actually get maximum power into the load, it can be shown that the Secondary has to have a precise "Secondary lag" of 90°. If the Secondary lag is less or greater than 90°, that will reduce the power further. See **Figure 2.3**. It can also be shown that if for example, D1 is such that the lag on the Primary side is less than 90° (D < 0.5), say 45°, the curves start flattening out. Now the Secondary side delivers maximum power when its lag equals the Primary lag (D1 = D2). So, if the Primary lag is 45° (D1 = 0.25), maximum power is obtained only when the Secondary lag is also 45° (D2 = 0.25). If the Secondary lag is made to vary from 45° to 135°, the power curve remains flat, *at the value it had at 90*°! In fact, if the Primary lag is reduced to say, 22.5° (D1 = 0.125), maximum power is obtained only when Secondary lag is also at 22.5 (i.e. D2 = 0.125). The power curve would now be flat over the entire range of Secondary lag from 22.5° to 157.5". The curves for 22.5° are not presented in **Figure 2.3,** since they are really of no use. But it is clear that all power curves are always symmetrical around Angle 2 = 90° (D2 – 0.5), and maximum power occurs exactly where Angle 1 = Angle 2 (i.e. D1 = D2).

What else does this imply? The answer is: there is *no point in trying to set D2 greater than D1 ever*. The power *capability* is specified by D1, but it can be controlled by D2. We get nothing by making D2 exceed D1. If the Primary is limited by a low value of D1, there is nothing we can do on the Secondary side to somehow "extract" more power. Keeping this in mind, our entire assumption from now on is: **D1 ≥ D2**.

In **Figure 2.2**, we also reveal that there are two sub-cases within the broader category of D1 ≥ D2. **Case 1: D1 + D2 > 1,** is accompanied by a "small edge" appearing in the Vsw_d waveform as shown. The other is **Case 2: D1 + D2 < 1**. The two cases produce different applied voltage waveforms and different time segments too, as we will see. The two cases need to be studied separately, though if the math is right, both cases must converge: give identical results for the case D1 + D2 = 1! That serves as a good cross-check point.

Understanding Transformer action

In the DAB schematic of **Figure 2.1,** we have shown an "ideal/DC-DC transformer", available in most simulator packages. It has a theoretical inductance of infinity, by definition. Which, indirectly, seems to imply that, using the inductor equation V=L ΔI/Δt, "no ΔI" can be injected into, say, the Primary winding, of this ideal transformer, and thus no current, "I", is possible either (because that would need a non-zero ΔI even to start things off!). Indeed, there is *no* current, at least not in the (infinite) "L" of the Primary winding of an *ideal* transformer. At least when there is no load present. But wait a minute: isn't there an actual measurable current flowing into the winding of any real transformer? With or without load present? Yes, there is an underlying "excitation/magnetization current" that every real transformer needs, just to start things off. More on this shortly, but to complete our transformer "model", we definitely need to connect a "magnetizing inductance" "Lmag" across it, as shown in **Figure 2.1**, representing the actual measured "Primary inductance" of the actual/real transformer. The reason for separating the constituents out is quite real: the magnetization current, though an enabler, does not itself participate in the "transformer action", which occurs (only) in the "ideal transformer" shown below Lmag.

That "no-load" current we measure, is the one that flows through Lmag: it is always flowing, irrespective of load. So, how do things change when there is a load present? Does that mean that we somehow manage to push current into the "infinite inductance" of the Primary winding of the ideal transformer? In a sense! But to get a good mental picture of how to understand all this, and thereby understand "transformer action" too, we look at **Figure 2.4** now.

We start with schematic "A" where we first have an arbitrary Full-bridge driving a *real transformer*. So, we are realizing that the current flowing into the Primary winding of this real transformer, actually has two components in general, not so obvious from this schematic. One is a load-independent current, corresponding to the current flowing in Lmag, shown in "B", as in **Figure 2.1**. This corresponds to the magnetization ("circulating current") component in the Primary-side, *not associated directly with power delivery*. The other "invisible" current component is the one that is actually a participant in the "transformer action". It is the load-dependent portion. In effect, it flows, not through the "infinite inductance" of the winding, but through the "load" which appears across the Secondary winding of our ideal transformer. If the load is not present (very high load resistor for example), the load-dependent current component falls to zero. At which point the residual current still present in the Primary winding, is the circulating "magnetizing" current component, and that continues.

So, what really happens is that the infinite inductance of the ideal transformer, goes out of the picture, *because the flux created by the Secondary-side (scaled) current, a result of Faraday's law of induction, exactly cancels all the flux created by the load-dependent Primary-side current component*. All this occurring within this ideal transformer model. Distinct to what happens in Lmag, even though physically, both are part of the same (actual) real transformer!
Note that no Secondary-side current can ever cancel the flux in the core arising due to Imag, though that is also flowing within the same real transformer. Imag is just an enabler, not a participant.

So, in effect the infinite inductance of the ideal transformer goes out of the picture, and this is best visualized in "C" within **Figure 2.4.** Here we have removed the ideal transformer entirely, and replaced it with the "reflected" load, or the "equivalent load" *as seen by the Primary side*. After this process of reflection, we can totally forget about the Primary-side (infinite) inductance of our ideal transformer. The entire transformer, with its inherent property of voltage and current scaling, is gone entirely now, being replaced by appropriately reflected impedances placed in parallel to the (still-remaining) *magnetizing inductance*.

All that takes us finally to "D" in **Figure 2.4**, where we have removed the infinite inductance, and so it is clear that all the current, except for Imag, heads straight to the load. And as indicated, we may *also* choose to (but perhaps *shouldn't always*) ignore the steady circulating current component associated with the actual, Primary-side inductance, Lmag, of the real transformer we started off with. That current component has ramifications as we will discuss, besides the obvious losses associated with this "circulating" current component.
Note: In reflecting impedances, from the Secondary side to the Primary side, we should remember always, that, based on energy invariance in the process of reflection, and since $Is \rightarrow Is/n$ and $Vs \rightarrow Vs \times n$, we get $Cs \rightarrow Cs/n^2$, $L \rightarrow Ls \times n^2$, and $Rs \rightarrow R \times n^2$, where $n = N_P/N_S$. Simply because $\frac{1}{2} \times (CV^2)$, $\frac{1}{2}$ (LI^2) and I^2R must all be preserved during reflection, despite the current and voltage scaling that occurs. And that is how we can scale anything from the Secondary side to its Primary side equivalent, with no intervening (ideal) transformer necessary anymore. The magnetizing inductance cannot be wished away.

So, we realize that from a high level, "transformer action" is a basic process via which: voltage scales as per turns ratio (Vp/Vs = n), and current, too, scales similarly, but in an opposite manner (Is/Ip = n). That "opposite" behavior incidentally, keeps the product "VI" unchanged on both sides of the transformer, as we should expect, since "VI" is energy/Watts, and that too cannot just appear or disappear, despite any tricks or reflection artifices of ours! Or our models are inaccurate and do not reflect nature.

The key difference between voltage and current scaling is clearly a bit more subtle. The voltage present across the ideal transformer Primary winding, and across Lmag is the same, so there was never any confusion as to what exact Primary-side voltage scales as per turns ratio. But most people get easily confused when it comes to understanding how *currents* scale. Basically, the *entire* current flowing into the Primary winding of a real transformer ("I_{TOTAL}") does NOT scale as per turns ratio! Only the load-dependent portion of that does. Basically, we have to remove Imag from I_{TOTAL}, and that is the current component which scales via transformer action. And finally, it is that scaled version which heads to the load and delivers output power. Whereas, the current associated with the magnetizing inductance, "Imag", simply goes back and forth every cycle and represents the circulating current component! It is not associated (directly) with any power delivered to the load, but is the enabler. Indeed, all the energy associated with Imag is theoretically fully recoverable. In practice, no recovery process is perfect, and there are associated losses.

Comparing to classical, transformer-based topologies, refer to Figure 3.6 of A to Z, Second Edition, showing a Forward converter. Note that the "magnetization current" component is just the shaded triangle riding on top of the reflected trapezoidal current waveform being reflected from the effective Buck stage consisting of the output choke, present on the Secondary-side. And that triangle is *not* related to the Secondary side current waveform by any scaling law. Its associated energy is recovered every cycle, by the energy-recovery ("tertiary") winding of the Forward converter. Since we usually set the tertiary winding to have the same number of turns as the Primary winding, but connected in such a way that it conducts only during the OFF-time, in effect we get the same voltage appearing across the tertiary winding during the OFF-time, as the Primary winding had across it during the ON-time. So from the basic inductor equation $V=L\ \Delta I/\Delta t$, we realize that the slope of the magnetization current during the ON-time is equal but opposite in sign to the slope of the recovery current during the OFF-time (albeit in different windings, but on the same core). So, to ensure "reset" we just have to ensure that we leave slightly *more time* for the magnetization current component to return to zero, than the time we gave it to ramp up in. Indeed, the ON and OFF time magnetization currents are in separate windings, but still on the same core, and thus can legitimately juggle the associated energy. And that is the reason why a standard Forward converter is always operated at a duty cycle slightly less than 50%, say 48%. That simple time-division approach ensures that the recovery current reaches zero, literally just in time, and stays there a wee bit longer, till the next cycle starts again. Inductor reset occurs every cycle. Note that the load-dependent Primary-side current component adds zero flux/energy to the core, because as we mentioned, the Secondary-side current produces flux which fully cancels the flux produced by that load-dependent Primary-side current component. That is why a Forward converter transformer core is so small, and usually left un-gapped too. The flux in it is always fixed. It does not depend on load current. So, if the size of the core is increased visibly for "higher power", that is only to create extra window area to accommodate the thicker copper which comes naturally with higher load currents.
In contrast, a Flyback "transformer" doesn't have "transformer action" per se. By energy conservation principles, we do get voltage and current scaling on either side too, but since the Secondary winding does not conduct when the Primary winding is conducting, there is no "flux cancellation" effect either, which is the very basis of "transformer action". We can't even talk in terms of a separate magnetization component and a separate load-dependent current component here. It is all just one current component. A Flyback transformer is best looked at as a multi-winding inductor, not as a transformer, despite exhibiting pseudo "transformer action." Another way to look at it is: it is *all* just magnetization current, and the Secondary winding is simply an "energy recovery" winding for that stored energy, except that unlike a Forward converter in which the "recovered" energy is delivered back to its input supply/capacitor, in a Flyback, the magnetization energy is dumped into an entirely different capacitor, the output capacitor in this case!

Why did we get into so much detail to recollect transformer action here? To point out that specifically, only *inductors* store energy, not "ideal transformers". Even the inductors which are inherently part of a real transformer. And for inductors, all inductors present, but *only* inductors, not for ideal transformers, we need to ensure that "reset" occurs every cycle. For it to be a viable topology.

In an ideal transformer, complete flux cancellation occurs at any load. So, there is no associated stored energy, ever. Nothing to "reset". It is pure transformer action. Any stored energy resides only in the associated magnetizing inductance. Though indeed, physically, that is part of the same physical (real) "transformer" that we are modeling. Which we now realize, is *part inductor, part transformer*! And we have to be very careful that we reset *the magnetizing inductance* part of it.

The same is typically true for the leakage inductance associated with a real transformer too. Though leakage is usually the result of flux straying through air, so it is in effect an air-gapped core structure, and we do not have to worry about saturating that leakage. On the other hand, in the PSFB and DAB, we usually prefer to use a separate, more predictable, "shim inductor", instead of the leakage from within the transformer, as shown in **Figure 2.1**. In that case, we have to be sure we reset the external/"leakage" inductor too.

In general, if we fail to ensure *all* involved inductors of a power converter reset every cycle, we will arrive at the phenomenon of "flux-staircasing", leading eventually to core saturation and destruction of switches.

In a DAB, the demand for inductor reset, both for Lmag and Llkg, seems difficult to ensure, or at least confirm to ourselves. First, because the way these topologies are constructed, with FETs in parallel to diodes, we really don't have the ability of using an "energy recovery" winding. But luckily, the two half-bridges of a Full-bridge alternate in action, and that can cause the magnetization current to ramp up, and then ramp down too, to achieve reset (though not in two distinct ON and OFF periods).

In any transformer-based converter, we always end up applying certain reflected voltages from the Secondary side to the Primary side. In fact, in a DAB, we will show that there are *eight* distinct segments per cycle, not just two (the ON and OFF times in a standard Forward converter). Those voltages are typically flat-topped, since the Secondary side gets clamped to V_O or V_{IN}, or some combination of the two. In turn these flat-topped voltage segments, "allow" certain rising or falling current slopes. These obey $\Delta I = (V \times \Delta t)/L$, but for different V's, as applicable. Yes, as the Full-bridge alternates between its two half-bridges, we can get various increments followed by decrements of current. To prevent "staircasing" over several switching cycles, we want to ensure that all these delta I's, with signs included, *add up to zero* every cycle. Just as automatically occurs in standard inductor-based topologies. As mentioned, the magnetization current does not need to "return to zero", as in a Forward converter. We can have positive or negative currents in this topology. The challenge therefore is: with all these variabilities, we still need to assure ourselves that the magnetization current component returns to *the exact value* it started the cycle off with (positive, negative or zero, as the case may be), every time, every cycle, for any phase combination or load combination. Even perhaps for any output voltage, as in a DAB. That would finally ensure "reset", and no "staircasing" and we would truly have a topology on hand.

There is another huge hint/lesson for us, from voltage-mode controlled full-bridge topologies, where if we were just depending on the associated time intervals of the rising and falling currents to be equal, in order to indirectly equalize say, $|\Delta Imag_{ON}|$ and $|\Delta Imag_{OFF}|$, and thus return the current at the end of every cycle to its starting value, we could be in big trouble. Slight inequalities, especially in timing, could cause flux-staircasing and eventual destruction of the switches. Same situation as a Forward converter inadvertently operated at, say, 51% duty cycle! You never heard about it, because it blew up right away.

For these reasons, it usually becomes imperative in Full-bridges, including the PSFB and DAB, *to incorporate a large-value DC blocking capacitor*, "C_DCBLOCK" as shown in **Figure 2.1**. Especially with voltage mode control. The purpose of such a large-capacitance value would be to do nothing to the high-frequency components of the switching current. Or even to the overall theoretical analysis. In practice, it would serve to merely absorb any slight inequalities, say those arising from slight differences in *timings* of the right and left side half-bridges. The DC blocking capacitor would take upon itself a slight DC voltage offset, thereby automatically re-adjusting the applied volts, to correct any inherent VoltSecond imbalance. Of course, we could similarly add a DC blocking capacitor to the Secondary side too, in series with the Secondary winding. It would maintain symmetry between the two Full-bridges, which would help in preventing staircasing when the power direction reverses, and the Secondary side becomes the new Primary side. And neither DC blocking capacitor would do anything at all to the overall expected performance, unless there was some inherent mismatch between any two half-bridges of any Full-bridge. In which case, it would delicately step in, by just the right amount, to ensure the desired/anticipated performance once again. So, it is more like a safety net. But a necessary one!

The underlying theory and explanation of the DAB and the PSFB however, does not involve, or need to involve, the DC blocking capacitor. *We too will therefore ignore it in our analysis too as in the simplified DAB schematic in **Figure 2.5***. But with a strong underlying recommendation to not forget the DC blocking capacitor in an actual build.

Now, we need to explain the reason for the presence of the critical component "Llkg", or the leakage inductance, in **Figure 2.1**. Yes, we will need to ensure that it resets firmly and unconditionally too, just as Lmag must. But with a DC blocking capacitor present, we would be receiving great assistance in the matter. Though still on our "to do" list: *we have to prove that the DAB is "inherently stable", and thus, the DC blocking capacitor actually stands a chance of correcting any unintended asymmetries between the two halves of the Full-bridges. Because even the DC blocking capacitor can't make any inherently unviable topology, viable. Just as no control loop can either.*

The Role of the Leakage Inductance "Llkg"

In **Figure 2.5**, by eliminating the transformer, we have the reflected output rail "V_{OR}" (= $V_O \times n$). The load across V_{OR} is actually the reflected load, equal to $R_{LOAD} \times n^2$, where $n = N_P/N_S$. The first thing we see from the simplified schematic in **Figure 2.5**, is that there would be literally nothing standing between the input source and the output capacitors, to regulate the voltage, or smooth it out, were it not for "Llkg". We would get a direct input-to-output "short", accompanied by spikes of very high current and eventual destruction. That too is not a viable "switching topology"! That is why in a Forward converter, we always need an output choke! Same as the PSFB in **Figure 2.6**.

Some early hobbyists in high-frequency switching power instinctively tried to use the simple "50-Hz" iron-transformer "principle": i.e. just apply turns ratio effect for attaining the desired voltage. Maybe add this new "duty cycle control" to tailor it precisely. What they didn't realize is it was now misapplied to an inadvertently well-coupled (low-leakage) high-frequency ferrite transformer, and with no input/output smoothening choke present to save the day! They all learned the hard way that what they had come up with was just a direct short from input to output, with no intervening parasitics like the leakage found naturally in the early iron-core 50-Hz transformers.

In the DAB too, we thus have good reason to introduce Llkg. For one, to prevent the "direct input-output short" effect. But to hopefully also produce smoothing and voltage regulation (full output control). In a PSFB, we use an output choke, as in a Forward converter, and also Lkg. The purpose of Llkg is, however, only to create soft switching, or ZVS by coaxing current to freewheel momentarily through the body diode of a FET during the deadtime, as discussed earlier. In a DAB, by using a relatively higher inductance for Llkg, we hope to acquire control of the output *and* also introduce soft-switching. So, the Llkg in a DAB, is basically one component for two functions. But the only way we can get it to work as we hope, is by no longer just switching the Full-bridge on the Secondary side synchronously (i.e. to aid diode conduction), as we do in the PSFB, but to actively switch the FETs, with phase control perhaps. And reverse the phases to reverse direction of power flow too. That is how the DAB really got off the ground.

Voltage and Current segments in the DAB

In **Figure 2.7**, we have for convenience, placed the entire Primary side to the left, and the entire Secondary side to the right. Note that the input DC voltage source feeds *both* the Primary-side half-bridges, though the actual wired connection may be implied, not drawn explicitly. Similarly, both the Secondary-side half-bridges are connected to the output capacitor and load (though the connections may be implied, not drawn).

After getting rid of the ideal transformer, by replacing the components with their appropriate reflected values (Rload × n^2 and $V_{OR} = V_O \times n^2$), and consciously deciding to ignore Lmag (with caveats), and any recommended DC blocking capacitors too, we see that the (only) intervening component left between the two sides is the leakage inductance Llkg. It will determine everything. Note that Llkg can be placed between the switching nodes Vsw_a and Vsw_c as shown, or between Vsw_b and Vsw_d. The results would be unchanged, since there are two separate grounds, the Primary and Secondary grounds, which are not connected to each other! The two sides are separate, except for Llkg. We can thus place Llkg anywhere in between the two full-bridges, as one component, or even distributed in any manner. All we need to know is the sum of all the intervening leakage inductances.

We are ultimately interested in finding an easy way to predict/know the voltage *differential* that appears *across* "Llkg", because that is the actual voltage which determines the slope of the current during any particular time segment. Note that it also doesn't matter how that voltage differential is referenced, with respect to any of the fixed voltage rails (including grounds). Because only the *differential* voltage across Llkg counts. But it can certainly get hard to visualize. Because, back to physics, whenever we talk of a "voltage", or "potential", that itself is a *relative* concept. There is always an implied reference value against which we are stating any voltage. Usually it is the ground, but here we have two grounds, and they have no fixed, established relationship to each other either! So, what is that "reference value" here, for talking about the voltage at any given point?

The good news is that though the grounds are "separate", they do get referenced to each other in some fashion. *Always*. The way we are applying phase relationships, instead of, say, duty cycle, there is never any instant when all the FETs on the Secondary side are OFF, or all the FETs on Primary side are OFF. That can never happen in phase-controlled systems. So, in all cases, on both the Primary and Secondary sides, at least one FET is definitely ON at any given instant on any switching node (excluding deadtimes), and that serves to connect all four switching nodes (actively) to one or the other of their related DC rails. In effect, there is always a definable voltage relationship between the Primary ground and the Secondary ground too, though that relationship keeps changing as the switches transition between ON and OFF states. Yet, purely in terms of voltage *differentials*, we can definitely state that there is always a definable voltage across Llkg, and thus continuity of current is maintained at all times between Primary and Secondary full-bridges via Llkg.

In other words, the reference values of voltages do not really matter, only the differential across Llkg does. But to get to that, we need to first get a mental picture of how current flows through the system, and then come up with a very simple artifice to know exactly what Vlkg is. That is how we proceed here.

Back to **Figure 2.7**, in each half-bridge, to avoid some clutter, we have omitted the reference designators of the low-side FETs, but they are obvious and understood, based on **Figure 2.5**. So, we have Q2, Q4, Q6 and Q8 directly below Q1, Q3, Q5 and Q7 respectively. We start with schematic "A", at a certain switch configuration, marked "now" on the Primary side. It corresponds to Vsw_a pulled high to V_{IN}, (or simply: "Va is Hi"). Now, the input source, connected to the V_{IN} rail through Q1 which is ON, attempts to push current through Llkg. Assuming the current is flowing, after it goes through Llkg and arrives at Vsw_c, it finds that particular node is pulled low, because Q6 is ON. So, it heads towards Secondary Ground. Hypothetically, the current could have instead tried to push its way at Vsw_c through the diode placed across Q5 (even though it is OFF), and that would have taken it through the output capacitor and the load, over to the lower side of Q8, from where it could have returned to Vsw_d, either through Q8, because it is ON, or it could have even gone through the diode across Q8 were it OFF. But luckily, *current always chooses the easiest/shortest path*, and so the current prefers to go from Vsw_c, straight to the Secondary ground through Q6, and immediately return to Vsw_d through Q8 (which is ON) and from there on to Vsw_b. It is a circulating current, as it has bypassed the load entirely. From Vsw_b, since Q4 is ON, it returns to the Primary ground. And thereby completes the full current path through the input DC source V_{IN}.

So, it is fair to conclude that the input DC source, V_{IN}, is pushing current, but that current does NOT reach the load in this particular configuration, simply circulating back through the Secondary ground and returning. But even that aspect is of secondary interest to us here. We are primarily interested in working out the voltage *across Llkg*. For that we need to identify the voltages that appeared on the left and right sides of Llkg! But the question still is: *with respect to which rail? The Primary ground maybe*? OK, with that reference level assumed, we applied V_{IN} to the left of Llkg through Q1 being ON, and then 0V got simultaneously applied to the right side of Llkg, through Q8, Q6 and Q4 (all being ON).

A more general way of looking at this emerges as follows.

$$Vlkg = Vp - Vs$$

Where,

$$Vp = Va - Vb$$
and
$$Vs = Vc - Vd$$

In other words, stick to differences in voltages all through. First find the difference between Va and Vb. That difference is called "Vp" for Primary-side applied voltage. Then find the difference between Vc and Vd. Call it "Vs" for Secondary-side applied voltage. Then the difference of Vp and Vs gives us Vlkg.

Let us confirm this gave us the correct result in our above example. We had Va = V_{IN}, Vb = 0, so Vp was V_{IN} – 0 = V_{IN}. Similarly, Vc = 0, Vd = 0, so Vs = 0 – 0 = 0. So Vlkg = V_{IN} – 0 = V_{IN}. Which is exactly what we got by tracing the current loop all the way through.

We can do the same for schematic "B" in **Figure 2.7** and also for "C" and "D" in **Figure 2.8**. The math validating our simple difference method is provided within each schematic. It confirms that thinking in terms of Vp and Vs is a good way of knowing what the voltage across Llkg is. Then we can ignore the complexity of how the Primary and Secondary circuits get connected to each other at any given instant.

So finally, in **Figure 2.9**, we show that we can think in terms of a single pole, triple throw (SP3T) switch on either side of Llkg. On the left (Primary) side it chooses between V_{IN}, -V_{IN} and 0 V. As indicated, we get V_{IN} as Vp, if Q1 is ON, and Q3 is OFF (Q4 is thus ON). We get -VIN, if Q1 is OFF, but Q3 is ON. If the switching nodes "a" and "b" are both high, or both low, then Vp is zero (the "flat"regions between the mountains and valleys of Vp). Similarly, on the right side of Lkg, we can get V_{OR}, -V_{OR} and 0. And thereafter, Vp -Vs is the difference voltage across Lkg, from which we can calculate the current segments.

In **Figure 2.10** and **Figure 2.11**, on the basis of careful simulations matching our emerging Mathcad spreadsheet, we can quickly confirm our "SP3T"-based method for calculating Vlkg. We have further tabulated the durations of all the eight segments based on selected D1 and D2, and the corresponding Vlkg's in each of the eight segments. We have two cases as mentioned earlier: Case 1 corresponding to D1 + D2 > 1, and Case 2 for D1 + D2 < 1. Which case is on hand is quite obvious by looking at the waveform shape on "Vsw_d", and comparing it to the "master" node "Vsw_a". So if at the moment Vsw_a goes high, it finds that Vsw_d is also high, we have Case 1. Otherwise it is Case 2.

For convenience, we have compiled all the durations and applied voltages for both cases, in **Figure 2.12**. In that we also present a Mathcad "proof" that if we take the duration of each segment multiplied by the applied voltage during that segment (essentially the voltseconds for that segment), and sum over all the eight segments, for any D1 and D2, we do get the sum over any cycle as zero. This is very providential, as it implies that the fundamental voltseconds law is going to be upheld without user intervention, and so "inductor reset" is always guaranteed in the DAB. It is thus "inherently stable" as we had hoped, and qualifies to be taken seriously as a new, fundamental topology. Note that we made no assumptions about the relationship between V_{IN} and V_{OR} either, so the voltseconds law will be upheld for any output gain (Gain = V_{OR}/V_{IN}). Of course, to be assured that small timing asymmetries don't throw all this into doubt, we still strongly recommend the DC blocking capacitor discussed earlier.

Now we know that the inductor current always reaches exactly the same value that it started the cycle off with. Because we now know how to calculate all of the delta-I's (increments and decrements), based on applied voltages, and we thus determined that over a full cycle, the net increment in I, happens to exactly equal the net decrement. But we still don't know the absolute value of the current at the beginning and end of every cycle. *What exactly is that value?* How do we figure it out? For one, that critical number determines the output voltage we can get for a give load, as we will soon learn. That determines how much of the current circulates, and how much gets delivered to the load.

Discovering the Current Starting/Ending values

We will take a peek at the emerging Mathcad spreadsheet here, which matches simulations so well. One of the hints we had for discovering the exact shape of the leakage inductor current waveform was in the initial observation from simulations, that the presence of a large DC blocking capacitor in the Primary and/or Secondary paths, did nothing to the performance of the converter. Which meant that the average DC value of the inductor current *had to be zero*. Yes, that can always be confirmed. Think: the DC blocking capacitor will in effect subtract any DC value present. So clearly there was no DC value to start with (of course assuming no minor asymmetries in FET timings), or the actual insertion of the DC blocking capacitor would have changed everything. It changed nothing. In other words, the inductor current must have a zero DC value to start with: equal areas above and below the horizontal zero axis. That makes sense since an ideal transformer by definition, only responds to AC, by Faraday's law of Induction, not to DC. We can conclude that both the Primary-side and Secondary-side current waveforms in **Figure 2.1**, must always have an average value of zero, since they are pure AC waveforms going into and out of an ideal transformer. Yes, even if there is load present! No DC. Because in any typical Secondary side of a transformer-based converter, the very purpose of the rectification diodes is to take the AC current waveform (with no existing DC value at that point), then rectify it (and *that* gives it a DC value). Then we push that DC current through the load, and deliver power. There is no contradiction or confusion.

As an example, let us show how the Mathcad spreadsheet worked out the actual inductor current waveform. Let us use the phase angles we had for matching against **Figure 2.11** i.e. Angle 1 and Angle 2 being 70° and 50° respectively. We have also used a V_{IN} value of 400 V and a V_{OR} value of 275 V in the spreadsheet, matching **Figure 2.11** (the Mathcad suggested value of the load resistor for achieving that V_{OR} was 100Ω, what we had used in the simulation, so it all matches).

First step: We generate and accumulate all the current increments/decrements suggested by the applied voltages in the 8 segments, but we arbitrarily start from zero current. As expected, and hoped for, the last segment returns the current back to zero current, the starting value. So, inductor reset is assured. But that does not mean it is the actual inductor current shape! Because that current must have an average (DC) value of zero. And this one doesn't. So, the trick we use in our Mathcad spreadsheet is to calculate the value of the current at 0°, and at 180°. Call it the "offset". We now move the current waveform vertically by *half the offset* value, to reposition the current waveform, such that the currents at 0° and at 180°, are equal and opposite in sign. Since we expect the overall current waveform between 0 and 360° to be symmetrical around 180°, this little offset trick indirectly assures us that we will get equal areas of current above and below the zero-current horizontal axis, i.e. with net zero DC value. See **Figure 2.13**.

As we can see, **Figure 2.13** also indicates an almost perfect match between the Mathcad spreadsheet and the SIMPLIS simulation. But to get to know the overall power capability of this converter, we need to calculate the DC component of the input current next.

Calculating the Input DC current component and Power Capability (Scaling Laws too)

This is the critical part. It determines how much actual useful energy is being pulled into the converter every cycle, for delivery to the output. We need to separate the actual DC current component from the inductor current. So, the basic rules for that are actually very simple, and the same for Case 1 and Case 2. Here are the steps:

a) Keep in mind that DC current can only be drawn every half cycle, *in the first three* of the 4 segments. Not for the entire duration.

b) We may be starting the cycle (i.e. Q1 turning ON), with a negative current, quite like what we see in **Figure 2.13** after we apply the correct offset, as explained above. This represents a certain charge in the input capacitors, which is part of the *circulating component*. And because it is circulating current, we need to keep it "circulating." In other words, even though at the start of the cycle, the input voltage source is "available," it is not asked to draw any current till full charge equalization takes place, as indicated by the two small triangular shaded areas in **Figure 2.13**. Until charge balance (basically zeroing the current integrated over time), takes place, no DC current is drawn from the input source. But once charge balance is completed (for the circulating component), the input source then provides the current for "making up" the rest of the inductor current waveform, which in turn was created simply by the applied voltages during the segments.

Once we know the average of the DC input current, then input power is very simply just the input voltage, multiplied by the average DC input current, and assuming perfect efficiency, that is the power which will be delivered to the output. In other words, this simple equality will tell us the power capability of the said converter, for the assumed or targeted V_{OR} (and thus V_O):

$$V_{IN} \times I_{IN\_DC\_AVG} = P_{IN} \simeq P_O = \frac{V_O^{\,2}}{R_{LOAD}}$$

And that clears up the mystery of how much power we can get for a desired output. Clearly, to increase the power, say two times, if we just halve the leakage inductance, the slopes of all the current segments will double and in effect, $I_{IN\_DC\_AVG}$ will also double, and we will get twice the power! Another vindication of our scaling laws. Similarly, if instead of asking for twice the power, after halving the inductance, we double the frequency, once again the increments of current will be back to what they originally were, and so would $I_{IN\_DC\_AVG}$. So, we would just get the original power. Summing up:

- To double the power at a certain frequency (keeping all else unchanged such as the phase angles), we need to halve the inductance.
- To double the frequency, keeping to the same power, once again we just need to halve the inductance.
- As before, if we double the input voltage, allowing the output to rise proportionally too, we will quadruple the power.

Now, how did we actually "see" the DC input current component in simulations? All voltage sources in simulators and usually on the bench too, are fully capable of not just sourcing current, but sinking it too. But that doesn't easily tell us the net DC power coming from such a source. Therefore, we put in a series diode, and a relatively small bulk capacitor to only provide the AC component. A capacitor automatically charges up or discharges, if charge balance is not maintained, and that helps reverse-bias or forward-bias the series diode at the right moment, to provide DC input current component to the converter. See the SIMPLIS schematic we used, in **Figure 2.14**. Notice that we have a DC blocking capacitor, because even the simulator often misbehaves without it.

Sample DAB design: 7kW (for PV applications)

With our new-found confidence, let us use our scaling laws, combined with our basic power delivery curves in **Figure 2.3**, to do a quick design. We will weigh the merits of the two design options presented here, a bit later. But both cater amenably to the following requirement (quite arbitrary).

Design a DAB which to deliver 7 kW from a 48V output rail. The rectified input rail is 380V. We want it to operate at 100 kHz. This will become bidirectional automatically, if we simply change the sign of "Angle 1", the "inter phase angle" (i.e. between the Primary and Secondary full-bridges).

We can always manage the 48V step-down function, using basic turns-ratio. So, if we assume a desired gain of unity (Gain =1) on the Primary stage, then we can choose to basically design it for a V_{OR} of 380V (Gain =1). To get 7kW, we will need to apply a reflected load resistor of $(380)^2/7000 = 20.7\ \Omega$ across that output (or equivalently, 329 m Ω across the stepped-down 48V output). Let us, for now, simply assume we can get close to 100% efficiency.
The good news is that **Figure 2.3** was also generated using 100kHz, so no frequency scaling is required here. Just power scaling and input voltage scaling.

To set the (fixed) phase angle of the Primary stage ("Angle 1" or "intra phase angle"), which will also be the phase angle within the Secondary side as we always assume, we realize that if we set this phase angle to less than 90°, we are essentially starving the Secondary side unnecessarily, and if we set it higher than 90°, we may be overdriving the Primary. But in either case, we get maximum power when Angle 2 (the "inter phase angle") is exactly 90°. See **Figure 2.3** closely to confirm that. Let us break our solution into two basic design choices:
Option A: Looking at **Figure 2.3**, we see that for its "Case B" (i.e. Angle 1 equal to 90°), we get exactly 475W from the converter (Gain =1, and with Angle 2 set to its maximum power level of 90°). But the curves of **Figure 2.3** were all generated using 200V input. If we scale that to 380V, we will automatically get (by the V^2 scaling law): $(380/200)^2 \times 475 = 1715$ W from the same converter. But our requirement is 7kW. So, the desired power scaling factor is 7/1.715 = 4.082. In other words, all we have to do is scale the leakage inductance of **Figure 2.3**, i.e. 52 µH to 52/4.082 = <u>12.75µH</u>.

Option B: If we double Angle 1 to 180°, then as per **Figure 2.3** (using the alternate vertical scale shown), we will get exactly twice the power, i.e. 2 × 475 = 950 W. Scaling that from 200V input to 380V would give us $(380/200)^2 \times 950 = 3430$ W . In that case, the power scaling factor is only 7000/3430 = 2.041. To meet this, all we would need to do is reduce the inductance of **Figure 2.3**, by this scaling factor, to 52 µH/2.041 = <u>25.5 µH</u>.

Let us set up the simulator to test this out. We will use an input of 380V, and place a (reflected) load resistor of 20.7 Ω, and then drive it with the phase angles: Option A: Angle 1 = 90° Angle 2 = 90°, Llkg = 12.75 μH and Option B: Angle 1 = 180° Angle 2 = 90°, Llkg = 25.5 μH. In either case, if **Figure 2.3** is correct, we should expect to get a V_{OR} of about **380V** (because we picked the Gain =1 curve in **Figure 2.3**). And if so, it would guarantee 7kW. Otherwise, not. Naturally, we don't want over-design, but certainly not under-design!

Results of simulation:
Option A: We get 392V and 7.45 kW. Very good match. To recall, Option A is: Intra-angle 90°, Inter-angle 90°, Llkg = 12.75μH, fsw =100kHz, V_{IN} =380V, V_{OR} = 380V (Gain = 1.0). Use turns ratio of 7.9 for getting 48V.

Option B: We get 390V and 7.3 kW. Very good match. To recall, Option B is: Intra-angle 180°, Inter-angle 90°, Llkg = 25.5μH, fsw =100kHz, V_{IN} =380V, V_{OR} = 380V (Gain = 1.0). Use turns ratio of 7.9 for getting 48V.

So, within the slight fudge factor of our graphical interface, we are still getting an astonishingly accurate match to simulations. And as an added, inadvertent benefit, we actually get slightly higher power than the "ideal" calculation, and that helps us, since in practice we expect the efficiency to be around 95-98% , not 100% as we assumed initially. A bit later, we will analyze the differences in Option A and Option B, though both meet our basic requirement.

More Design Curves and Validation

In **Figure 2.15**, we have for convenience, design curves for Angle 1 = 180°, and various angles for Angle 2, not just 90° as in Option B above. It just confirms that we won't get more power by making Angle 2 greater than 90°. But this also documents the drop-off on both sides of the peak in **Figure 2.3**. We pick a trial point here and do the simulations to confirm the match. It is not a maximum power angle, but just for checking, we are picking Angle 2 as 45°. This corresponds to a Gain target of 1.25 and **Figure 2.15** says that we need a 70 Ω (reflected) load resistor to pull that off. Note that these curves do not depend on input voltage. Expressing the reflected output voltage as a ratio of input voltage makes the curve "universal" in a sense. ***Note that we have introduced a special case with Angle 2 = 26°, for reasons we will describe later.***

In **Figure 2.16**, we have compared the key waveforms generated by the detailed Mathcad spreadsheet (from which the easy graphical aid of **Figure 2.15** came about), against the SIMPLIS simulations. In the simulations we used an input of 400V, and used the suggested 70 Ω from **Figure 2.15**, though the exact value the Math spreadsheet gave us was 69.5 Ω . The simulations returned 505 V instead of 500 V, which is very accurate indeed.
In **Figure 2.17**, we look a little more in detail at the waveforms from the simulation and since the current goes negative in Q1, Q2, Q3 and Q4, just before turn-ON, with a little deadtime we are, in effect, allowing the body diode (or paralleled diode) across it to conduct, just before we turn it ON, thus getting ZVS. In Q6 to Q8, we see that the current is zero, just prior to turning ON, so we have in effect zero-current switching (ZCS).

Is the value of 70 Ω consistent with **Figure 2.3**? Let's check. For a gain of 1.25 at Angle 2 of 45°, the power is exactly between the grid lines for 425W and 450W. Let us say, 437W. Wait a minute, that vertical axis was for an Angle 1 of 90°. For 180° we will get exactly twice that power. So, we expect 2 × 437 = 874W. Since **Figure 2.3** was generated for an input of 200V, with a gain of 1.25, we expect V_{OR} = 1.25 × 200 = 250V. To get 874 W from this output would require a load resistor of $(250)^2$ /874 = 71.5 Ω. Very close to the 70 Ω predicted by **Figure 2.15**. So there is no contradiction between the two methods, as described by **Figure 2.3** and by **Figure 2.15**. We can use either method/graph, then apply scaling laws, to meet any requirement. Using resistor value graphs from **Figure 2.15**, we do not even need to scale the input voltage. Because the resistor value for a certain gain applies to any input voltage! It scales by itself, is one way of looking at this!

In **Figure 2.18**, we have some more design charts, mainly around the possibility of setting Angle 1 at 45° or 135°. We are focusing only Case 1 (i.e. D1 + D2 <1). Note that, by now, we know that we always get maximum power only for Angle 2 = 90°, irrespective of Angle 1. All power curves are symmetrical around Angle 2 = 90°, as per **Figure 2.3**. In **Figure 2.19**, and **Figure 2.20,** we simulate three test points, arbitrarily chosen from **Figure 2.18**. The resulting VOR is very close to what we expected. In **Figure 2.20**, though we see that in some cases, ZVS can break down, at some phase combinations. We will discuss this in more detail later.
What if: we had D2> D1?

Previously we had hinted that the "intra phase angle" of each full-bridge (Angle 1, i.e. D1 × π) should always be set greater than the "inter phase angle" (Angle 2, i.e. D2 × π). A key reason is that D1 also corresponds to the power capability of the Primary side, and reducing that, is counterproductive. From **Figure 2.21**, we can easily see that one impact of that is the angle available for drawing in DC input current is substantially reduced, because the AC (circulating current) is relatively higher and takes longer and longer to reset. This refers to the triangular areas mentioned when we discussed **Figure 2.13**.

The Two choices for Angle 1 for the same Power and Different Gain Targets

One of the most fundamental questions for a DAB designer who wants to keep it simple: should we set Angle 1 as 90° or 180°? We can see from **Figure 2.3** that both give an optimum shape for the power curve, incorporating no "flat" region around Angle 2 = 90°. We also realize that the Angle 1 =180° curve gives twice the power than the Angle 1 = 90° case, albeit with different leakage inductances. However we could consider running the Angle 1 = 180° converter with not Angle 2= 90°, but with a smaller phase angle, to deliver exactly the same power we would get for the case of Angle 1 = 90° and Angle 2 = 90°. So what is that phase angle?

Looking at **Figure 2.3** closely, we will see that all the curves for Angle 1 equal to 90° or 180°, irrespective of set gain, end up at half power when Angle 2 is reduced from 90° to exactly 26°! Which means that for a given desired power at a certain desired gain, we have two distinct possibilities. Let us call these "Converter A" and "Converter B" for convenience.

Converter A: Angle 1 = 180° and Angle 2 = 26°
Converter B: Angle 1 = 90° and Angle 2 = 90°

That is why in **Figure 2.15**, we specifically provided the curve for Converter A case above.

So for the base kernel on which **Figure 2.3** was based, we have calculated the required load resistor, and thus expected VOR and power output. The comparative results are plotted in **Figure 2.22**. In fact since another key question in mind is, what should be the target gain? We can of course get our desired output rail using turns ratio of the transformer, but the underlying DAB can be set for any desired gain, step-up or step-down. So what is a good choice? Also, to guarantee high efficiency.

We realize that we would like to introduce a higher circulating current component, just to instigate ZVS, but that also leads to higher conduction losses in the bargain. So, it is a tradeoff of course.

However, by staring at **Figure 2.22** a bit we can make the following conclusions:

a) In general, Converter A case is preferable
b) In general, it seems a good idea to target a gain of unity

With that choice we get not only soft-switching at maximum load, but also much lower transistor RMS currents, so lower conduction losses too.

It is also very interesting that the RMS current of all transistors Q1 to Q8 is identical in all cases! Despite the very different current shapes! A proof that nature attempts to distribute stresses evenly, if we let it! And the DAB does exactly that by providing eight transistors, distributed evenly on both sides.

What if: We run the DAB in synchronous mode?

In **Figure 2.23**, this last question is answered too. Basically, the power drops significantly, but we do get soft switching again, thanks to the presence of the inductor. Had we added an output choke, we could have added control of output voltage too, to this "synchronous DAB" of ours. That is how we go back in time and reinvent the PSFB!

Final Solved Example of Chapter 2

Here we will use **Figure 2.3** for a quick example to reinforce concepts.
Design a DAB converter for 200kW, from 300V input to 800V output. It is preferred to operate this at 85kHz.

Looking at **Figure 2.3**, the recommended reflected load resistor is 84Ω with 52μH for a Gain of 1, as recommended, starting with Case B. Note that this resistor value does not depend on input voltage! Or on frequency!

If we place this resistor across the reflected output voltage of 300V (since Gain =1), the power we will get (from the kernel on which **Figure 2.3** was based) is:

$$P_{INIT} = \frac{V_{OR}{}^2}{R} = \frac{300^2}{84} = 1.071 \text{ kW}$$

But we want 200kW. The desired power scaling factor thus is:

$$P_{SC} = \frac{P_O}{P_{INIT}} = \frac{200k}{1.071k} = 186.74$$

The frequency scaling factor is

$$f_{SC} = \frac{f_O}{f_{INIT}} = \frac{85k}{100k} = 0.85$$

Therefore suggested (scaled) leakage is

$$Llkg_{FINAL} = \frac{Llkg_{INIT}}{f_{SC} \times P_{SC}} = \frac{52uH}{0.85 \times 186.74} = 0.33\ \mu H$$

The reflected load resistor we need to test this out is

$$R_{FINAL} = \frac{R_{INIT}}{P_{SC}} = \frac{84}{186.74} = 0.45\ \Omega$$

Finally, we present the results of the simulations (with a 1:1 transformer). So, we expect to arrive at a V_{OR} of ~300V, since we set Gain =1 (preferred). We have used 0.33μH leakage inductance, and a reflected load resistor of 0.45 Ω, and we can confirm we achieved a bit over 200kW. **Figure 2.24** is for the baseline case of Angle 1= Angle 2 = 90°. We get exactly a bit over 200kW with 300V reflected output. But not necessarily soft-switching! In **Figure 2.25** we show exactly the same circuit, but driven at the preferred setting of Angle 1= 180° and Angle 2=26°. As before, we see we reached a reflected output voltage of around 300V, commensurate with a target gain of unity. And thus, an output power of a bit over 200kW as we expected. With this latter configuration, we can see we have full soft-switching assured! So, this is our final design target: 0.33μH, 85kHz! That's all!

Indeed, we want to deliver this power to 800V, so we need a turns ratio of $n=N_P/N_S = V_O/V_{OR}$ = 800/300 = 2.667.

This will deliver 200kW from an output of 800V with full soft-switching if we prefer to set the intra-phase angle of both full-bridges as Angle 1= 180°, and the phase lag between Primary and Secondary full-bridges as Angle 2=26°. That completes the 200kW design example! Based on just one graphical aid, namely **Figure 2.3**. Such is the power of scaling as applied to DAB too.

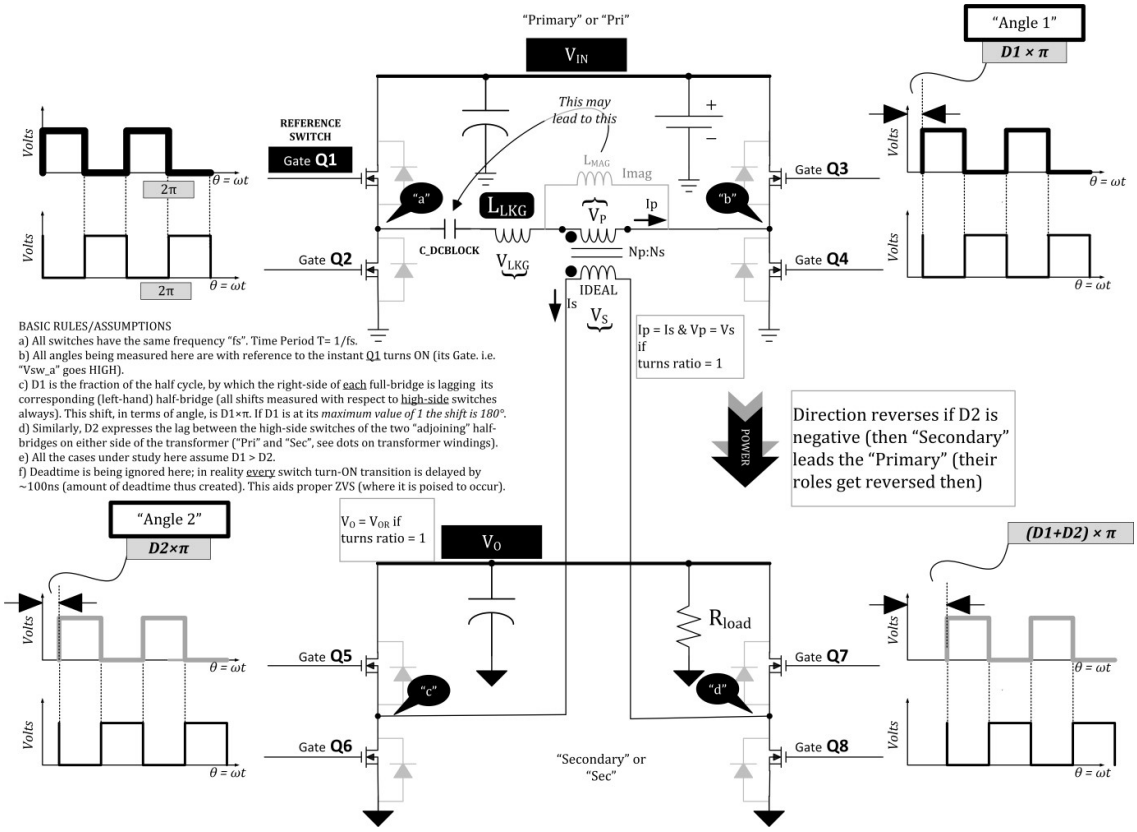


Figure 2.1: The Dual Active Bridge under study

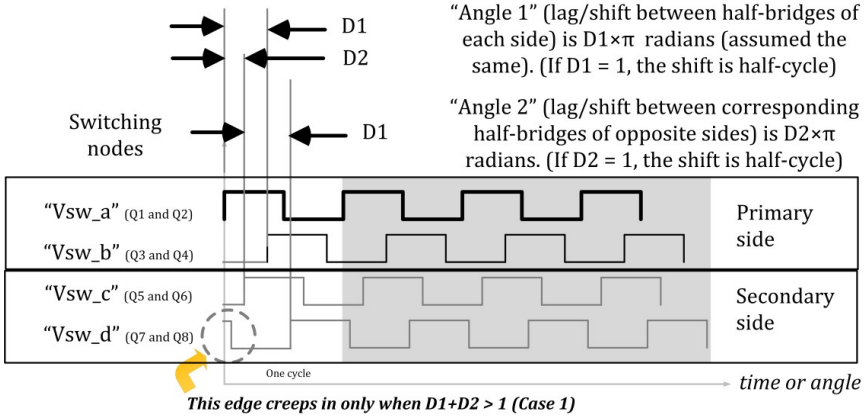


Figure 2.2: The timing diagram and related assumptions

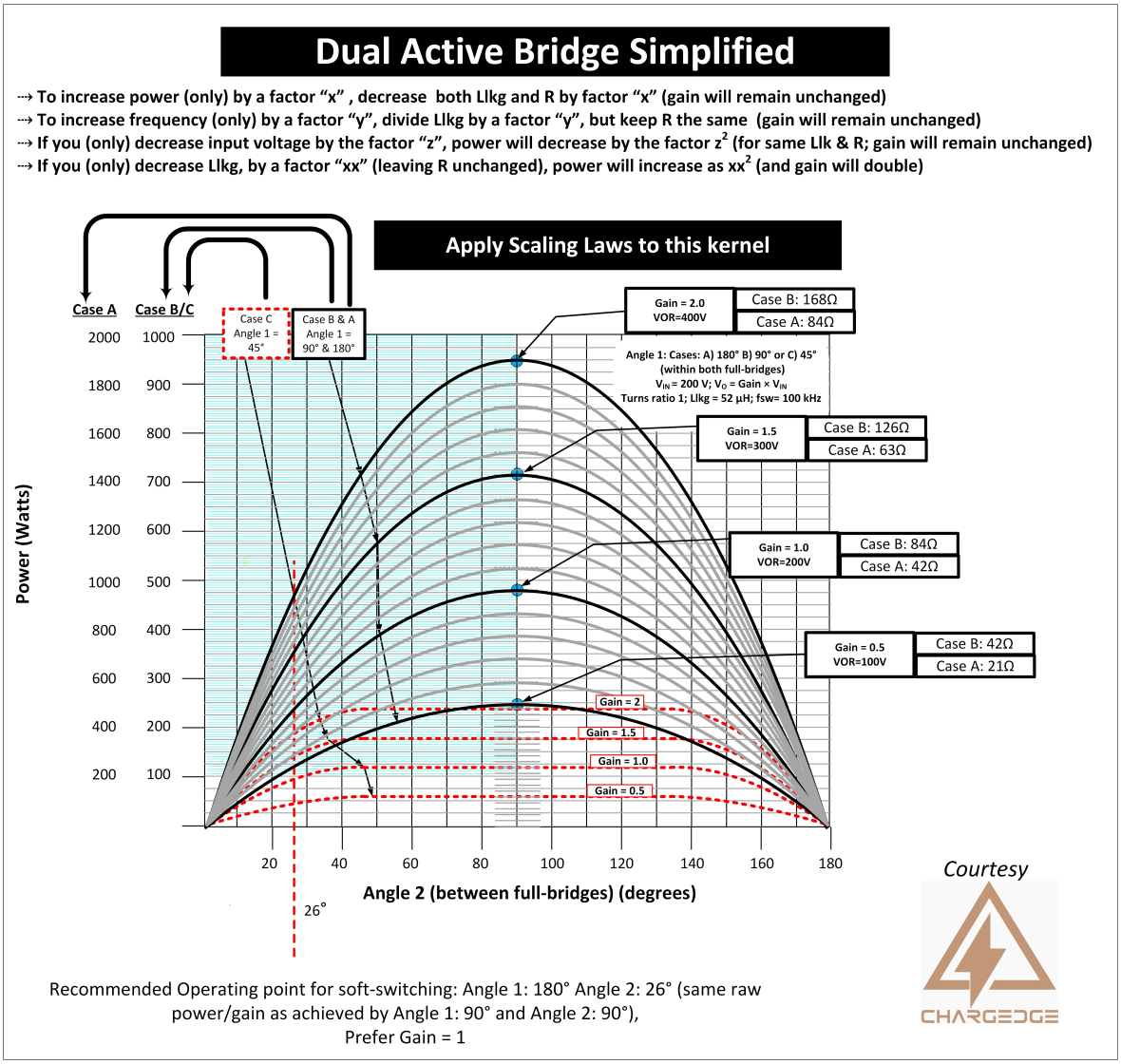


Figure 2.3: The DAB kernel

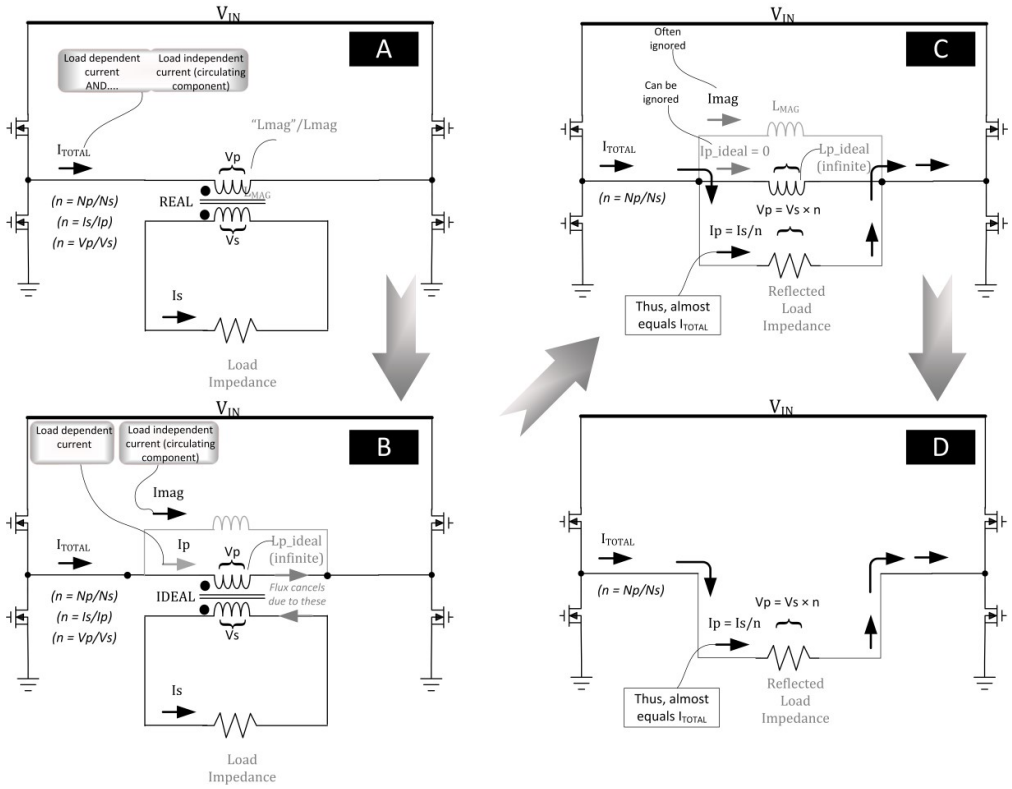


Figure 2.4: Understanding transformer action and simplification for DAB

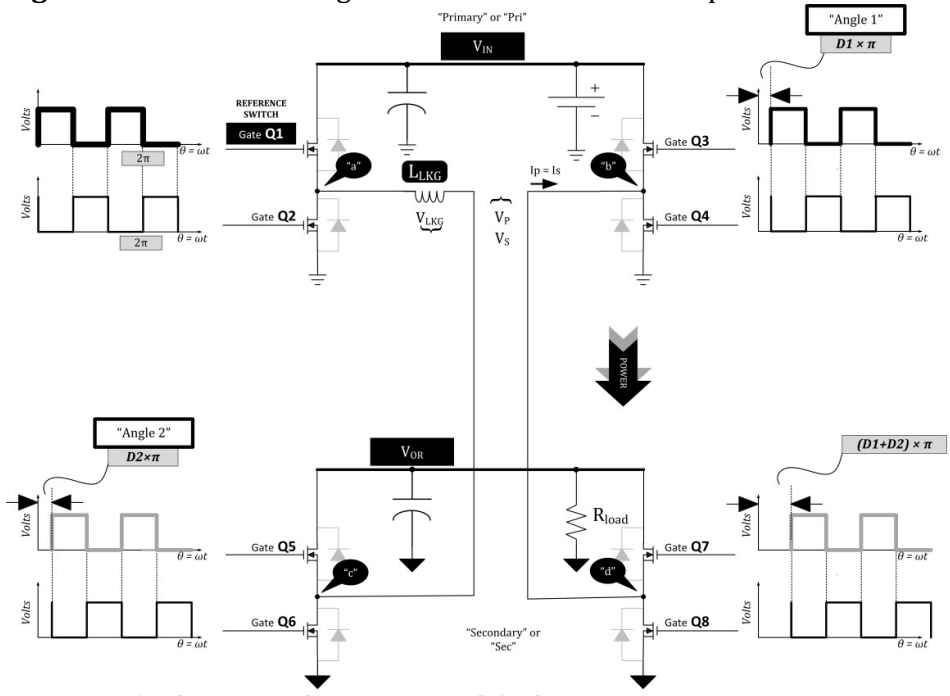


Figure 2.5: The DAB schematic simplified

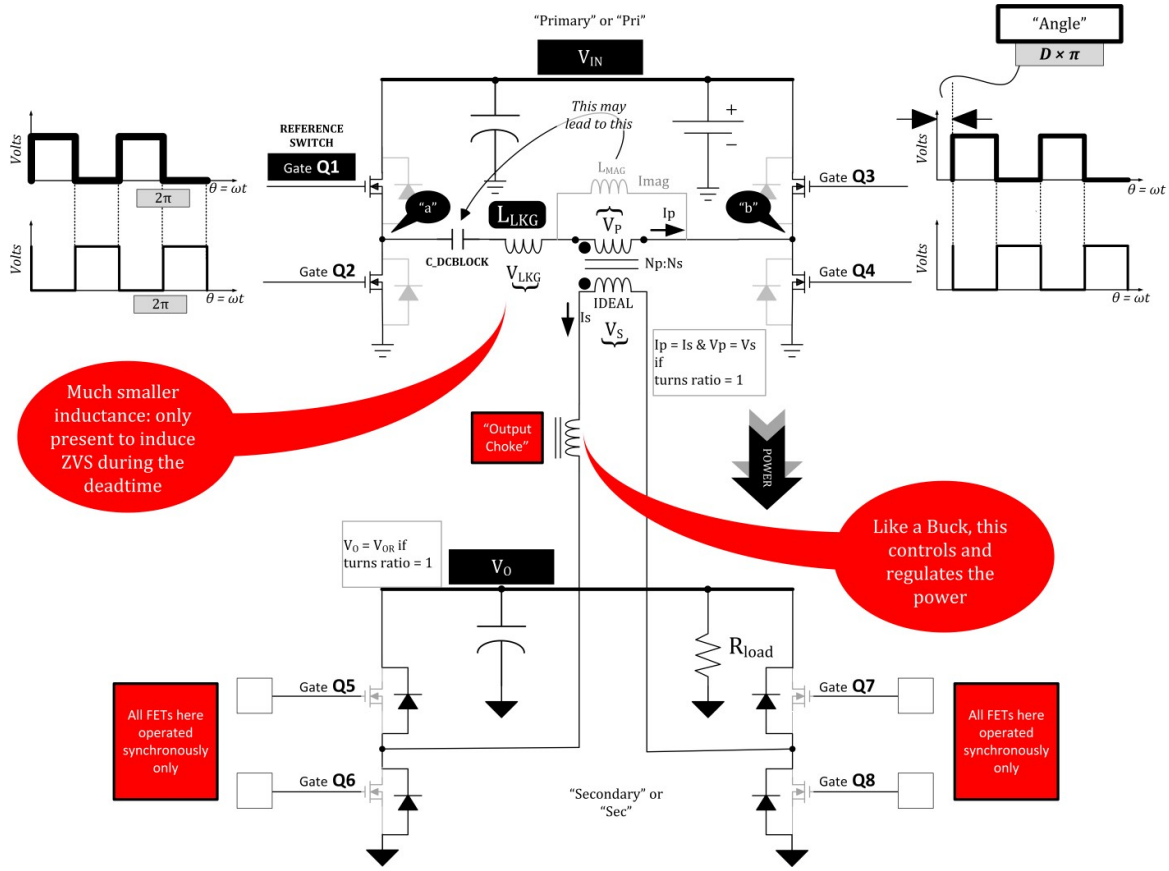


Figure 2.6: The Phase-shifted Full-bridge

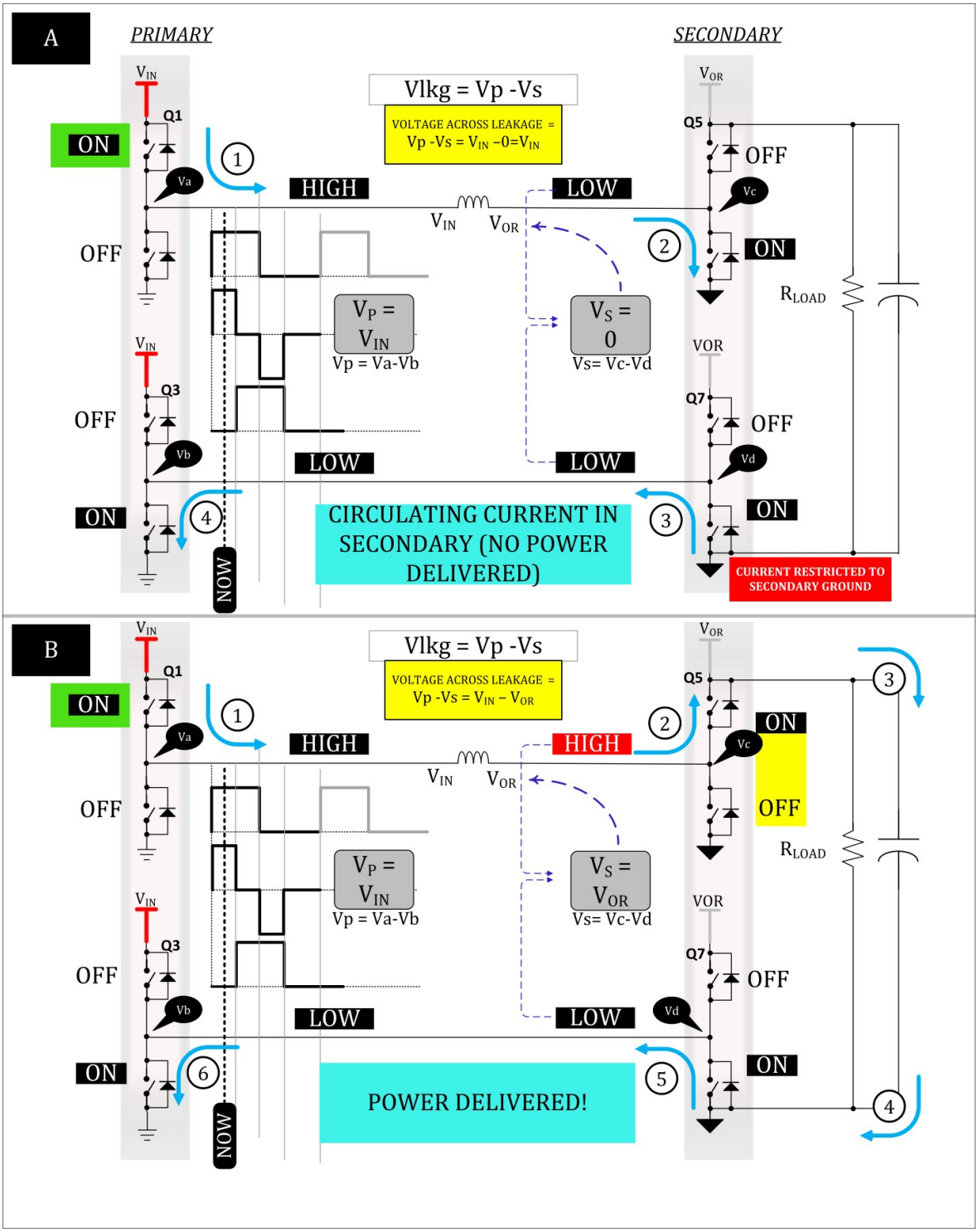


Figure 2.7: Examples of how the DAB works (Part 1)

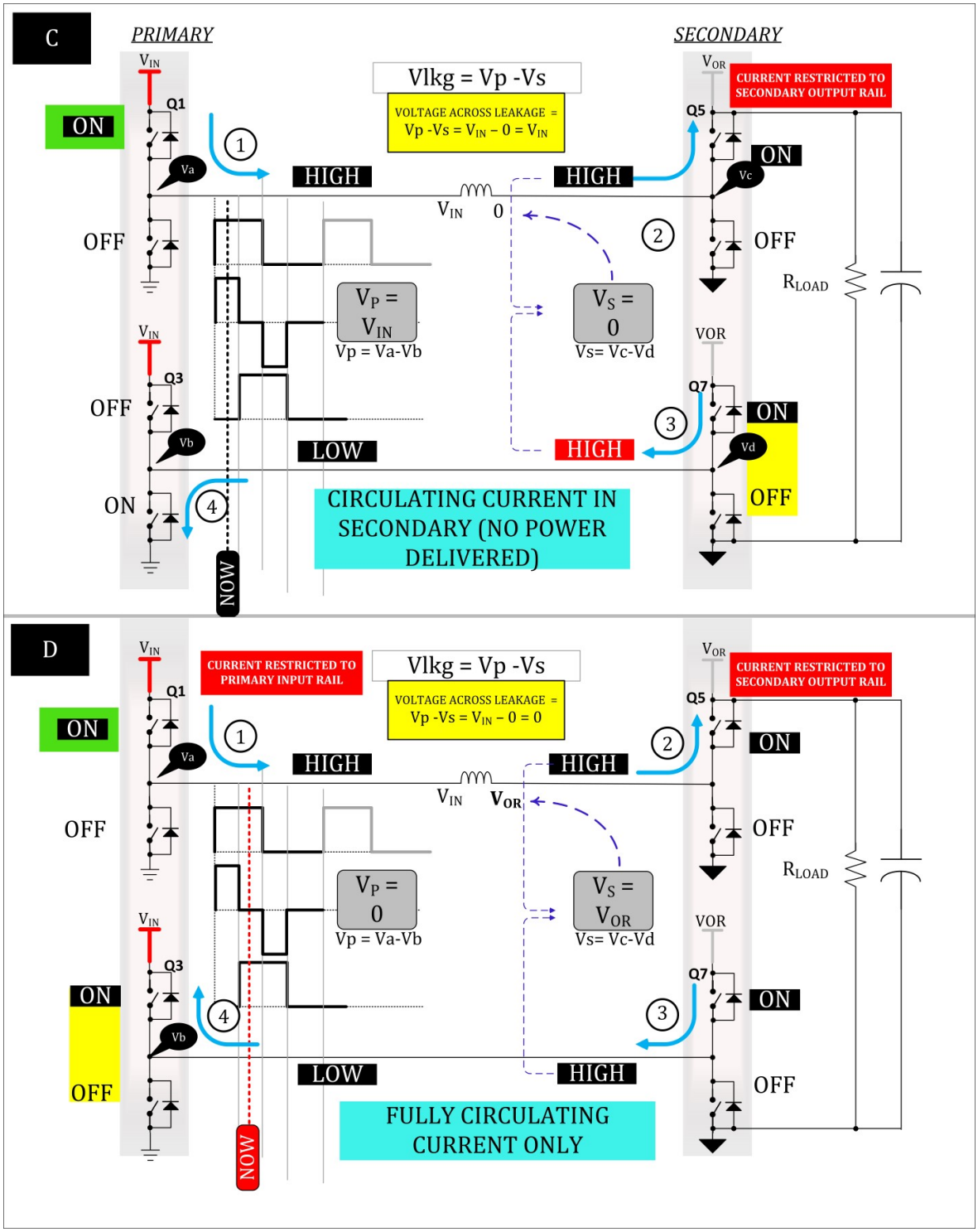


Figure 2.8: Examples of how the DAB works (Part 2)

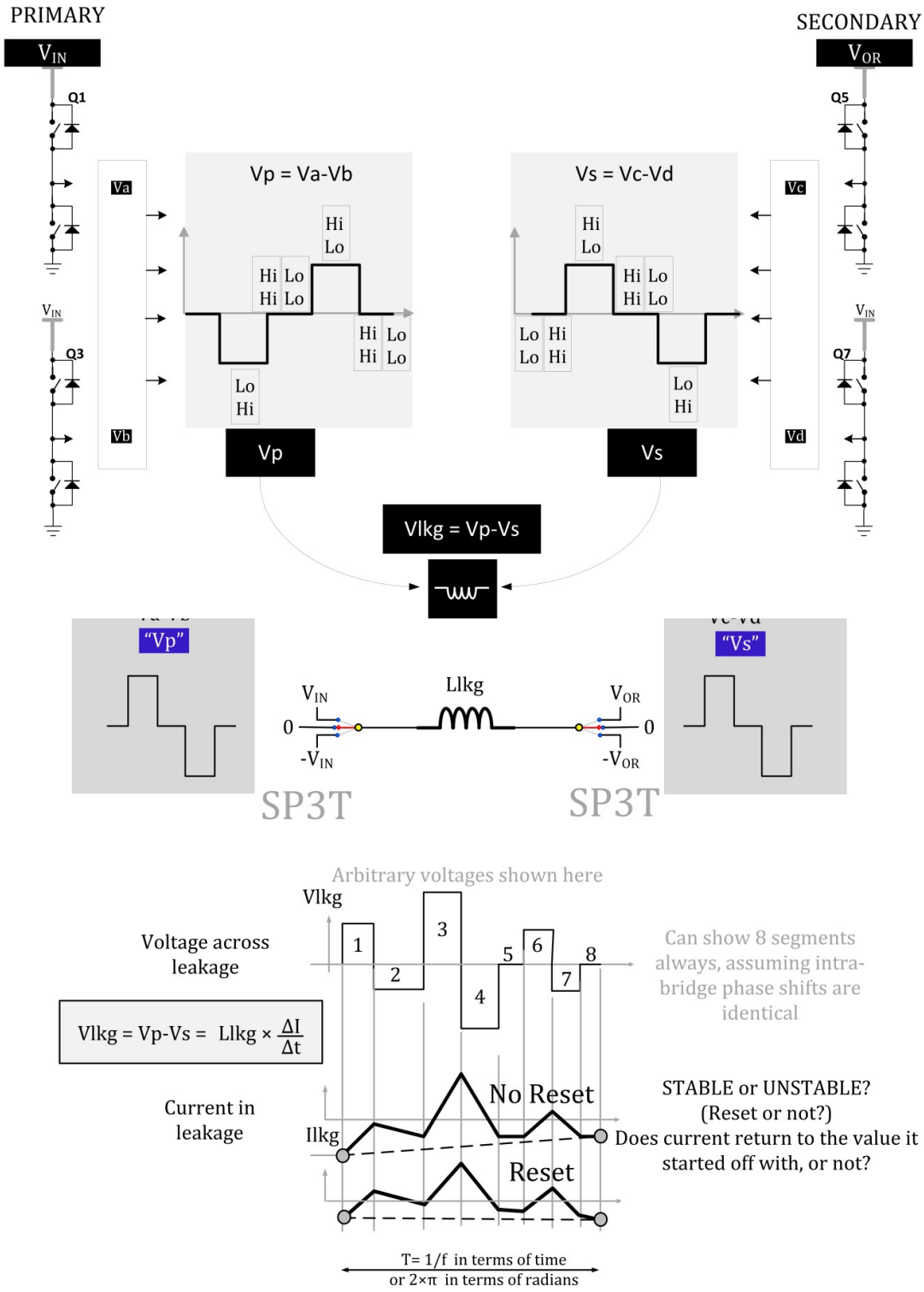


Figure 2.9: How the voltage segments cause changes in slope of current (in principle)

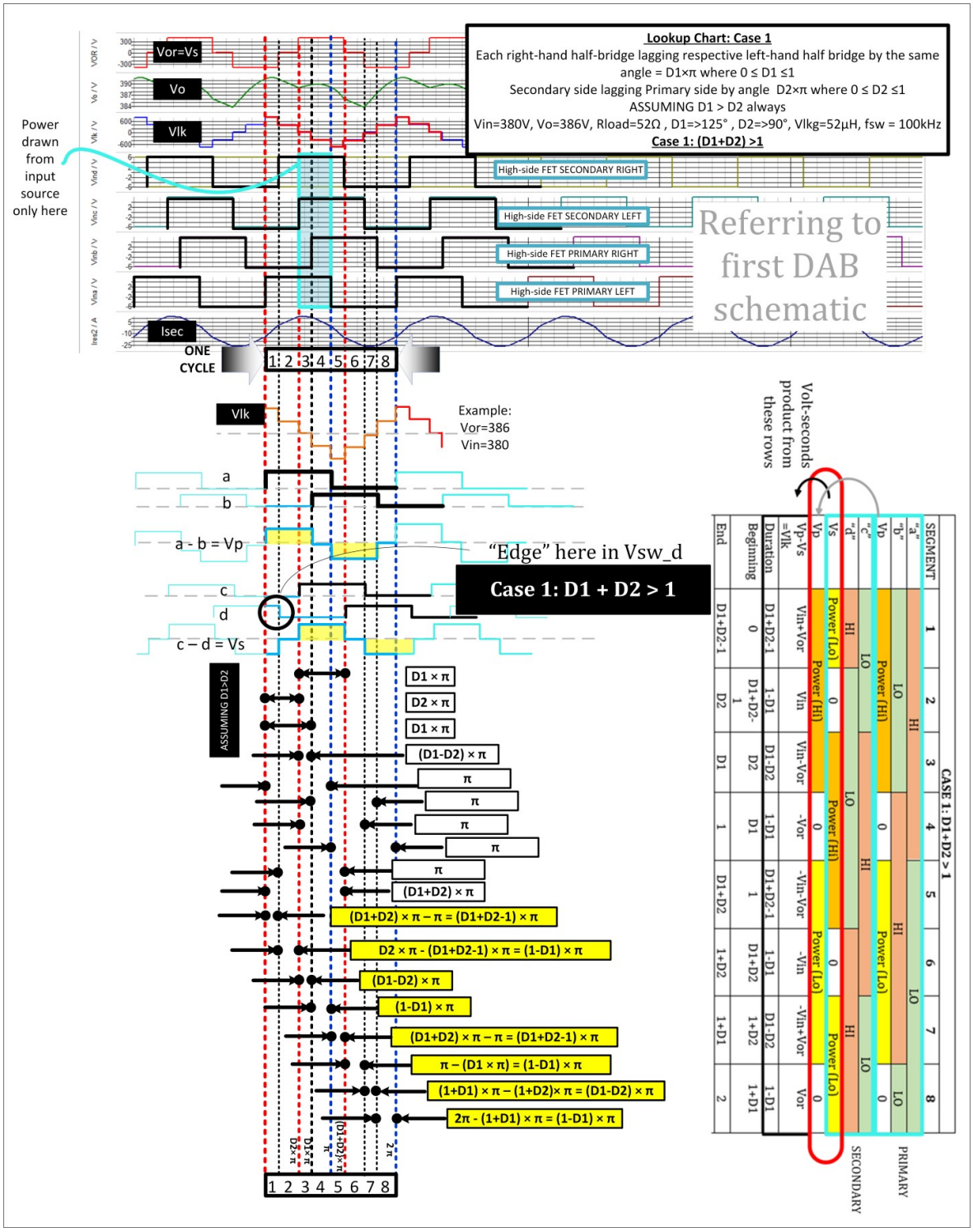


Figure 2.10: The eight segments of Case 1

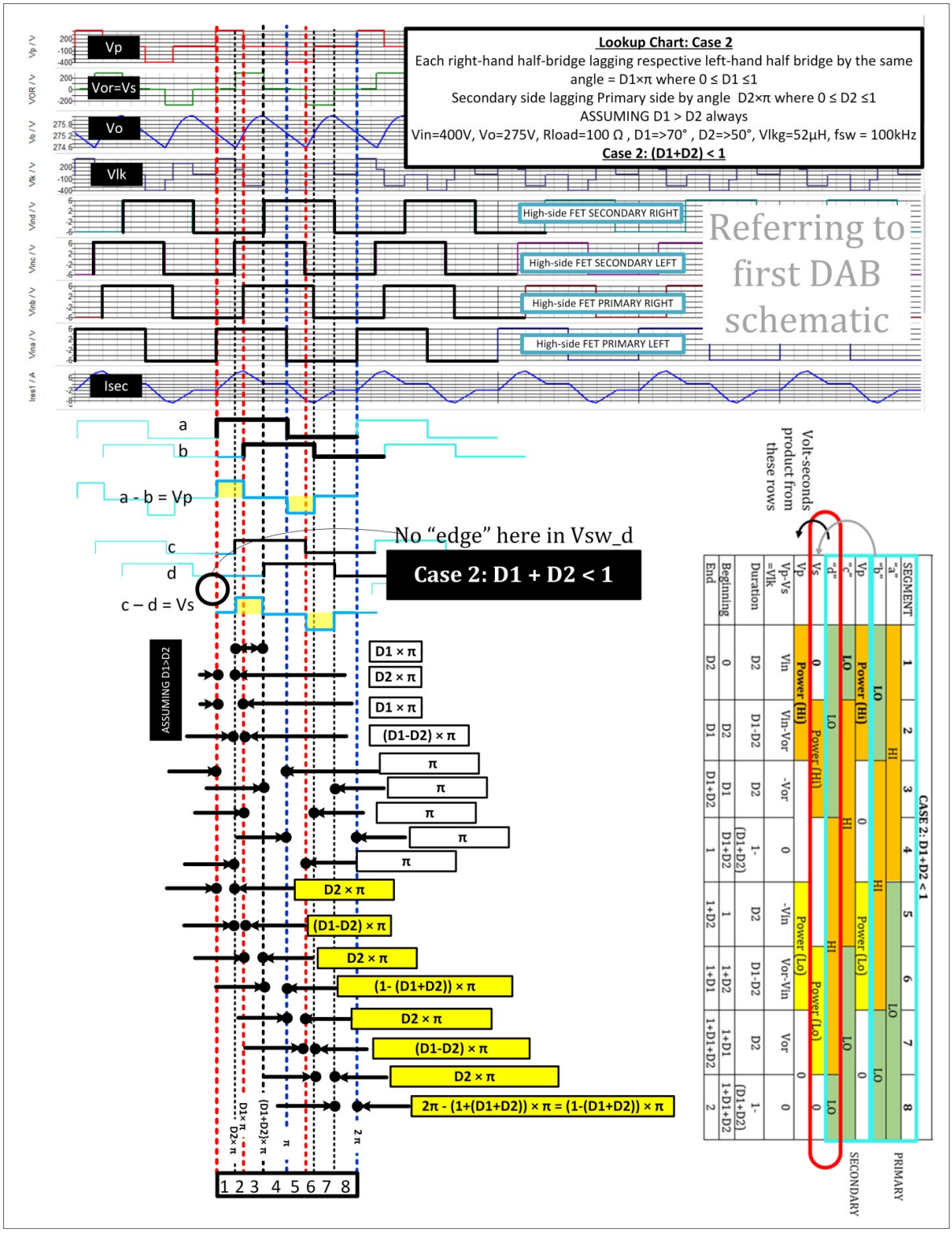


Figure 2.11: The eight segments of Case 2

CASE 1: D1+D2 > 1

SEGMENT	1	2	3	4	5	6	7	8	
"a"	HI				LO				PRIMARY
"b"	LO			HI				LO	
Vp	Power (Hi)			0	Power (Lo)			0	
"c"	LO			HI				LO	SECONDARY
"d"	HI		LO				HI		
Vs	Power (Lo)	0	Power (Hi)			0	Power (Lo)		
Vp	Power (Hi)			0	Power (Lo)			0	
Vp-Vs =Vlk	Vin+Vor	Vin	Vin-Vor	-Vor	-Vin-Vor	-Vin	-Vin+Vor	Vor	
Duration	D1+D2-1	1-D1	D1-D2	1-D1	D1+D2-1	1-D1	D1-D2	1-D1	
Beginning	0	D1+D2-1	D2	D1	1	D1+D2	1+D2	1+D1	
End	D1+D2-1	D2	D1	1	D1+D2	1+D2	1+D1	2	

Volt-seconds product from these rows

CASE 2: D1+D2 < 1

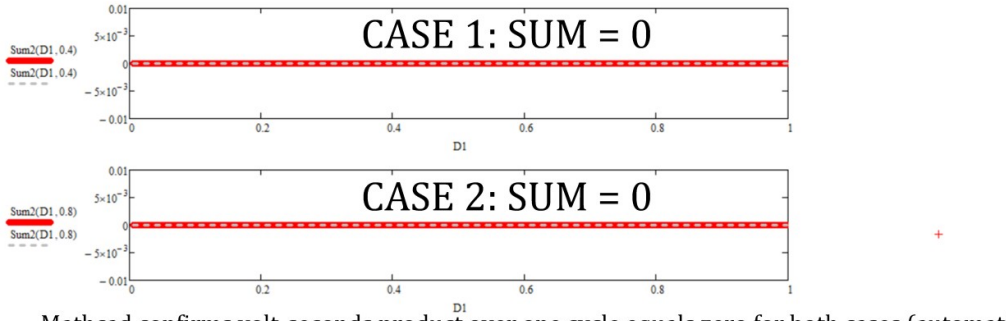
SEGMENT	1	2	3	4	5	6	7	8	
"a"	HI				LO				PRIMARY
"b"	LO			HI			LO		
Vp	Power (Hi)			0	Power (Lo)		0		
"c"	LO		HI			LO			SECONDARY
"d"	LO			HI				LO	
Vs	0	Power (Hi)				Power (Lo)		0	
Vp	Power (Hi)				Power (Lo)		0		
Vp-Vs =Vlk	Vin	Vin-Vor	-Vor	0	-Vin	Vor-Vin	Vor	0	
Duration	D2	D1-D2	D2	1-(D1+D2)	D2	D1-D2	D2	1-(D1+D2)	
Beginning	0	D2	D1	D1+D2	1	1+D2	1+D1	1+D1+D2	
End	D2	D1	D1+D2	1	1+D2	1+D1	1+D1+D2	2	

Volt-seconds product from these rows

Vin := 400 Vor := 35 D2 := 0.9 D1 := 0.005, 0.01 .. 1

$$Sum1(D1,D2) := (Vin + Vor)\cdot(D1 + D2 - 1) + Vin\cdot(1 - D1) + (Vin - Vor)\cdot(D1 - D2) + (-Vor)\cdot(1 - D1) + (-Vin - Vor)\cdot(D1 + D2 - 1) + (-Vin)\cdot(1 - D1) + (-Vin + Vor)\cdot(D1 - D2) + (Vor)\cdot(1 - D1)$$

$$Sum2(D1,D2) := (Vin)\cdot(D2) + (Vin - Vor)\cdot(D1 - D2) + (-Vor)\cdot(D2) + (0)\cdot(1 - D1 - D2) + (-Vin)\cdot(D2) + (Vor - Vin)\cdot(D1 - D2) + (Vor)\cdot(D2) + (0)\cdot(1 - D1 - D2)$$



Mathcad confirms volt-seconds product over one cycle equals zero for both cases (automatic inductor reset will occur, for any V_{IN} and V_{OR}, and of course any load current!)

Figure 2.12: Compiled table for the eight segments, both cases, and a Mathcad "proof" that sum of voltseconds over a full cycle is unconditionally zero

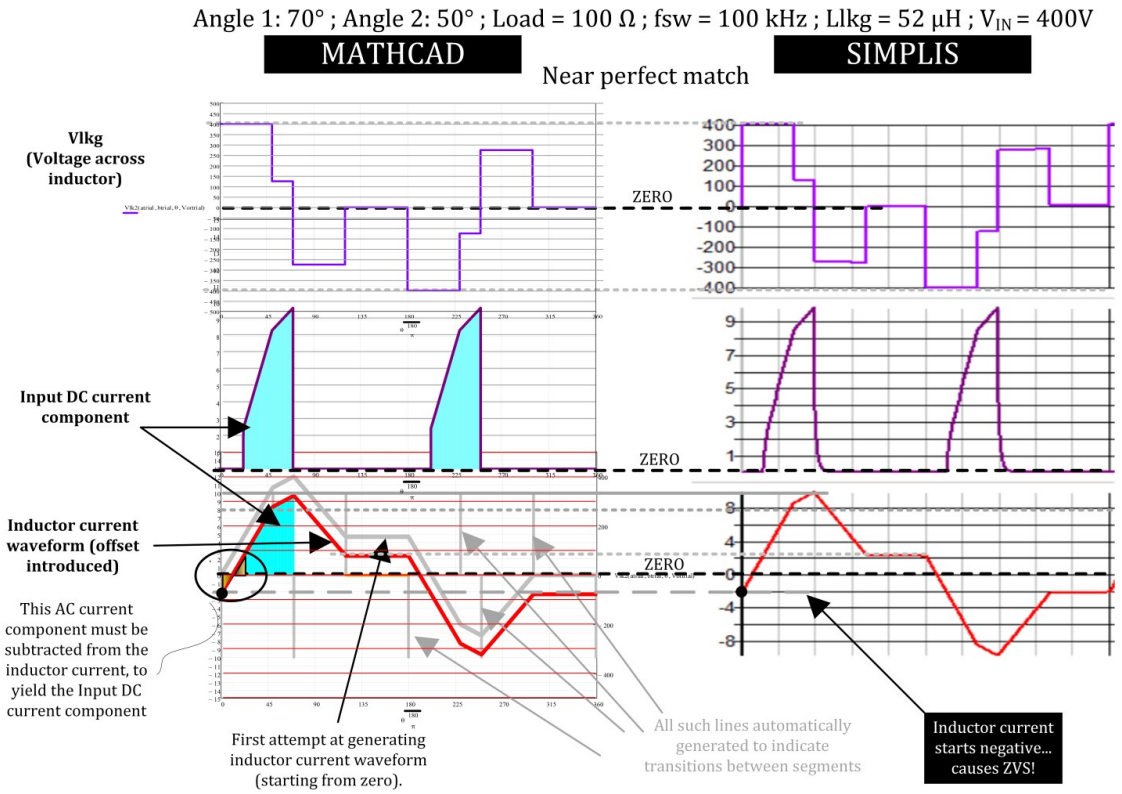


Figure 2.13: Mathcad and SIMPLIS comparison for Case 2 [i.e. (D1 + D2 <1)]

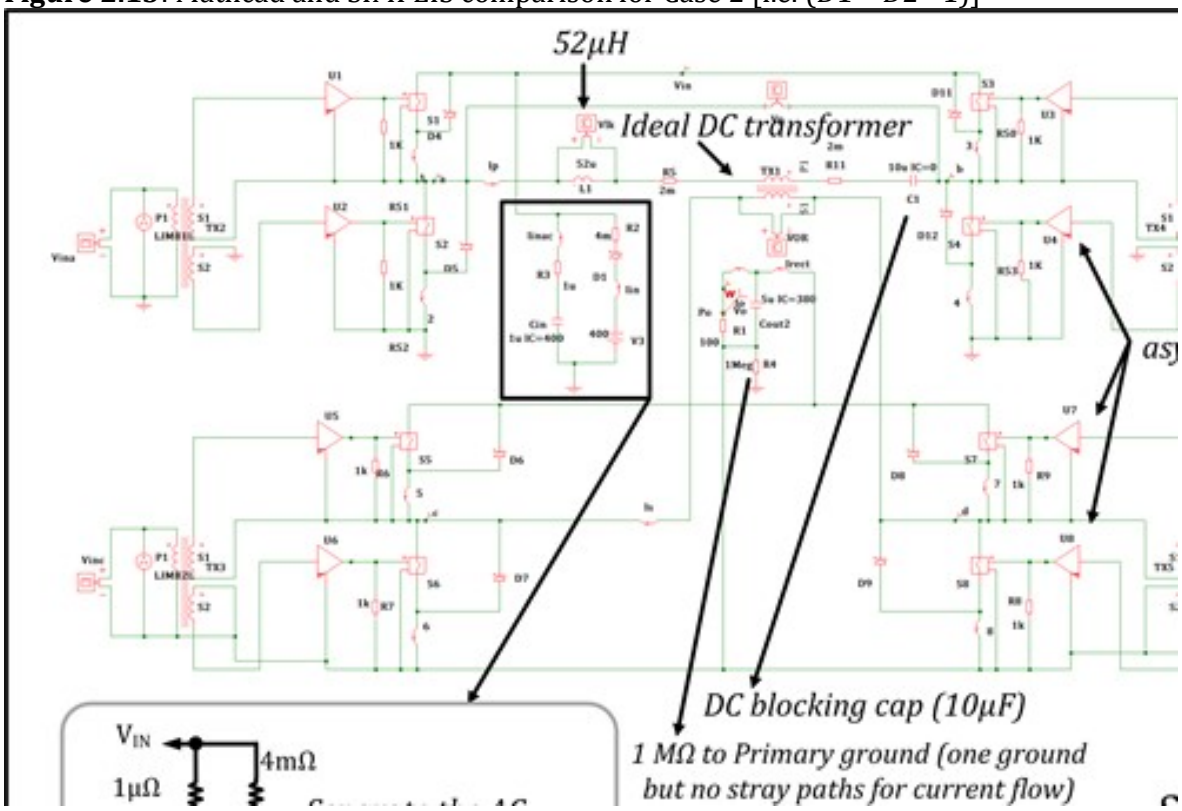


Figure 2.14: The SIMPLIS simulator used, showing how to extract the input DC current component

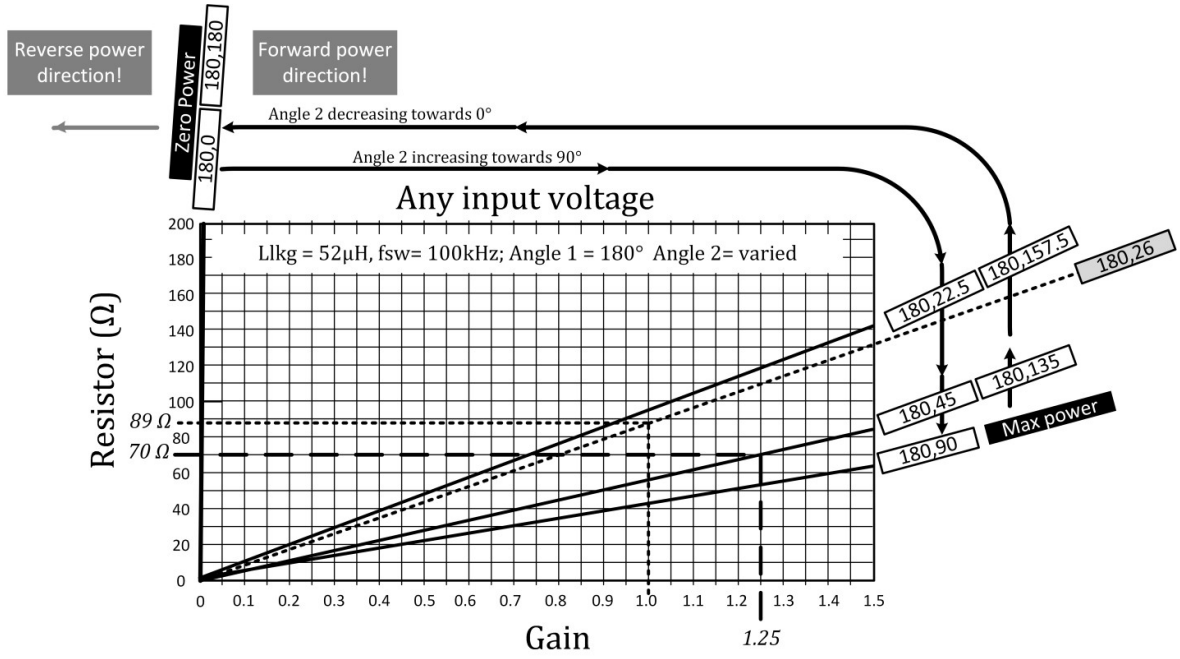
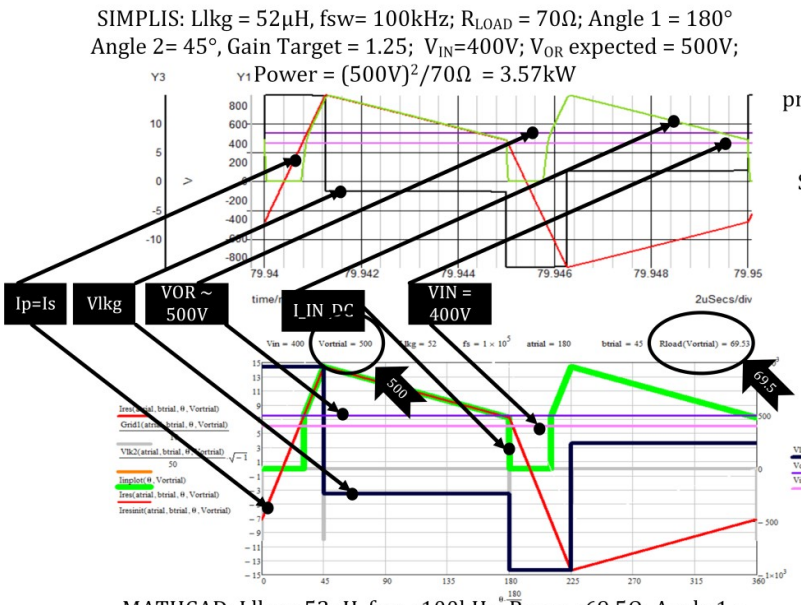


Figure 2.15: Resistor Lookup chart for case of Primary-side half-bridges fully out of phase



MATHCAD: Llkg = 52µH, fsw= 100kHz; R_{LOAD} = 69.5Ω; Angle 1 = 180° Angle 2= 45°, Gain Target = 1.25; V_{IN}=400V; **V_{OR}** expected = 500V; Power = (500V)²/70Ω = 3.57kW

Mathcad-based graphical aid predicted that for a Gain of 1.25 (VOR 500V) we should use 70 Ω. Therefore we used that value in SIMPLIS simulation, and the result was VOR = 505V (very close to 500V)

(Very good match)

Exact Mathcad spreadsheet predicted for a a VOR = 500V (Gain 1.25) use load 69.5 Ω

Figure 2.16: Comparison between Mathcad and SIMPLIS predictions (test case of Figure 2.15)

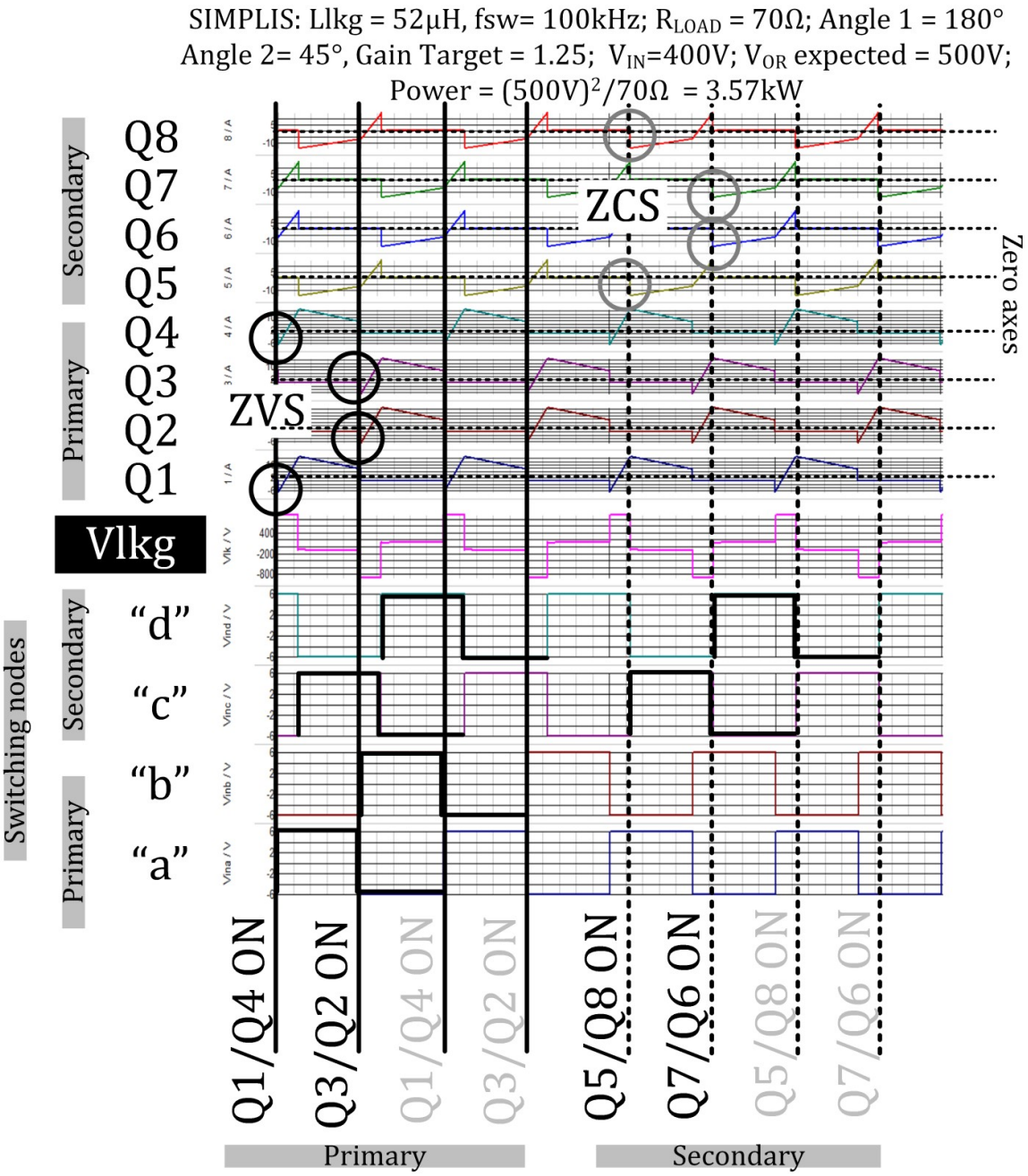


Figure 2.17: Checking for soft-switching in the case presented in the preceding figure

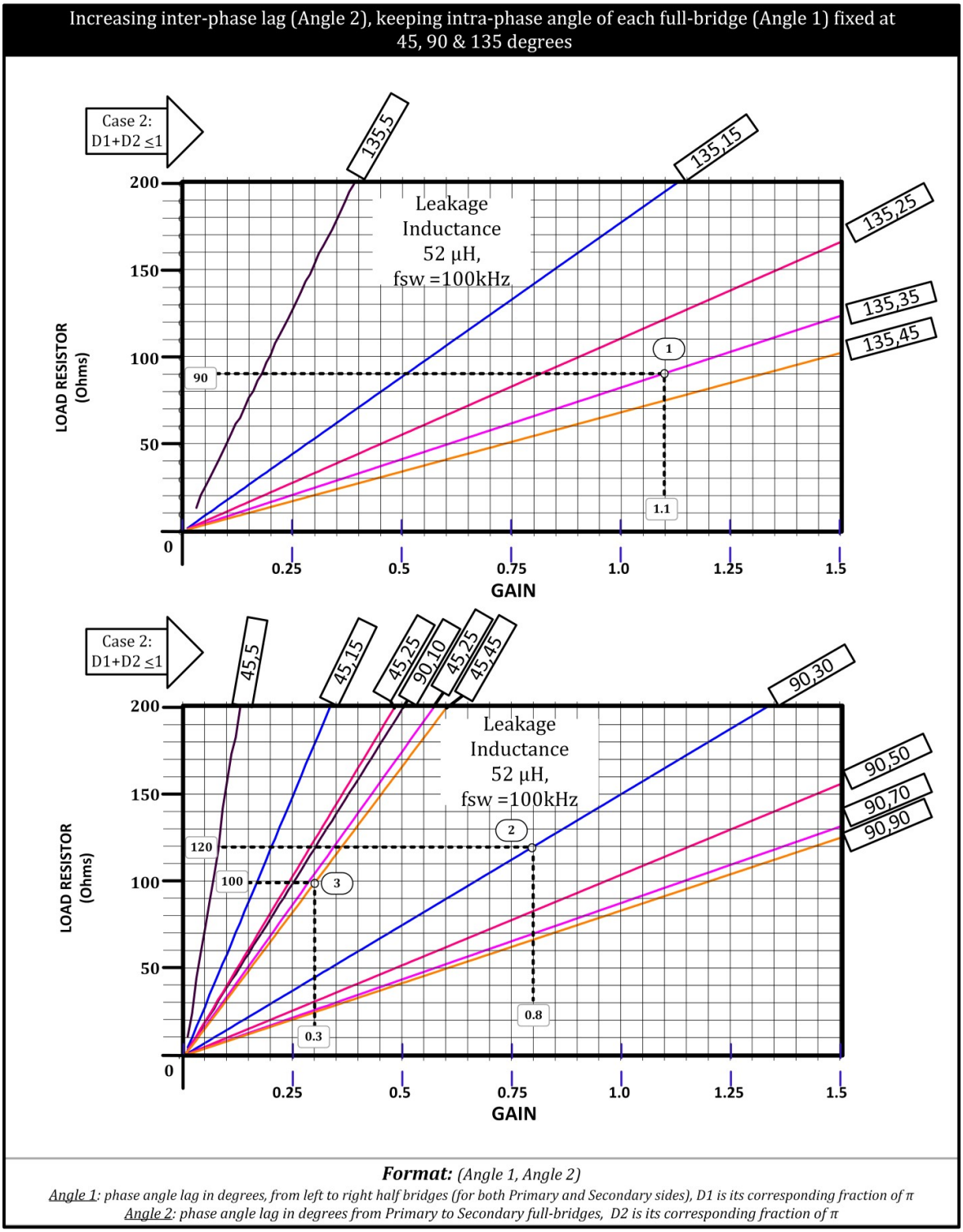


Figure 2.18: More design graphs, focused on Case 2 (D1 + D2 <1)

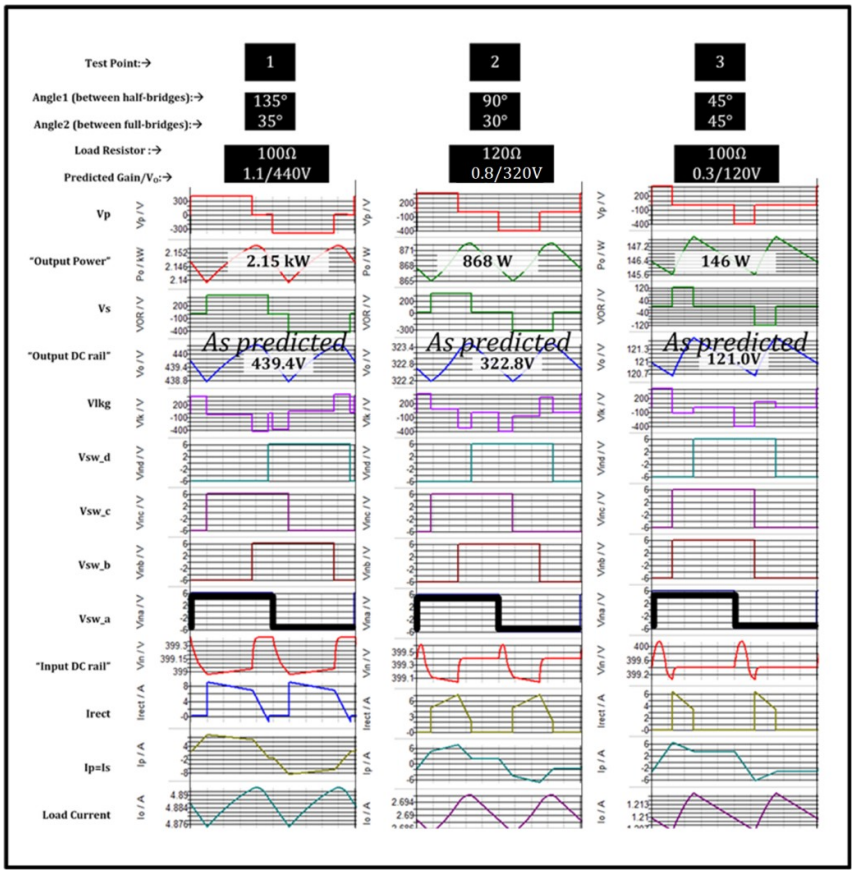


Figure 2.19: Simulation results for three test cases, showing good agreement

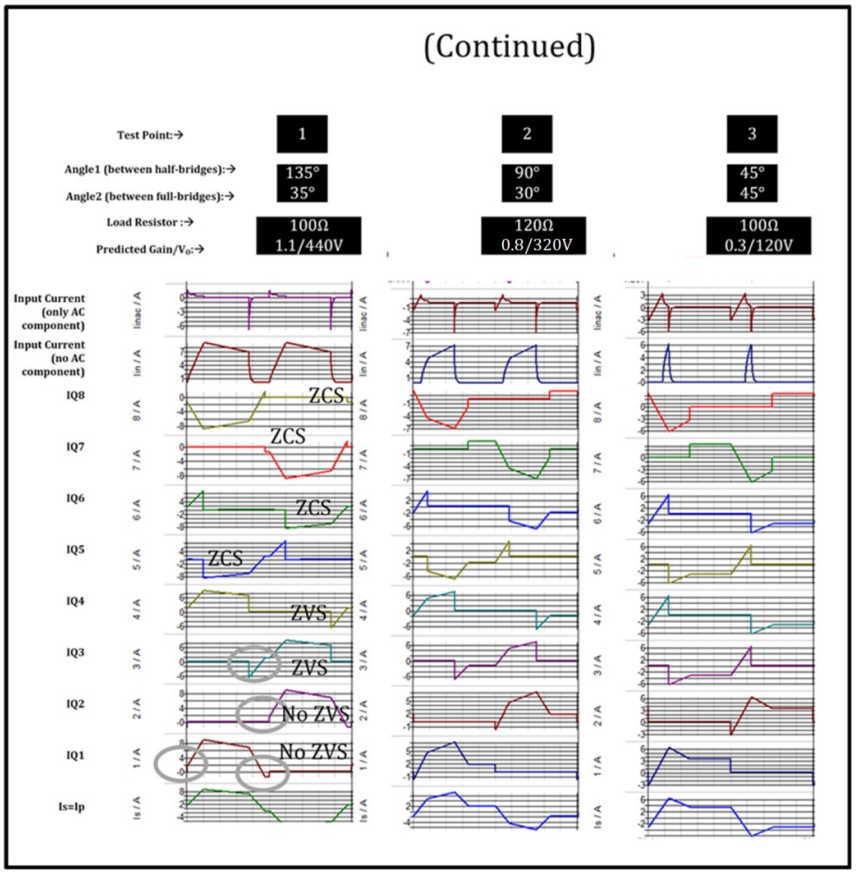


Figure 2.20: Continued simulation results from preceding figure

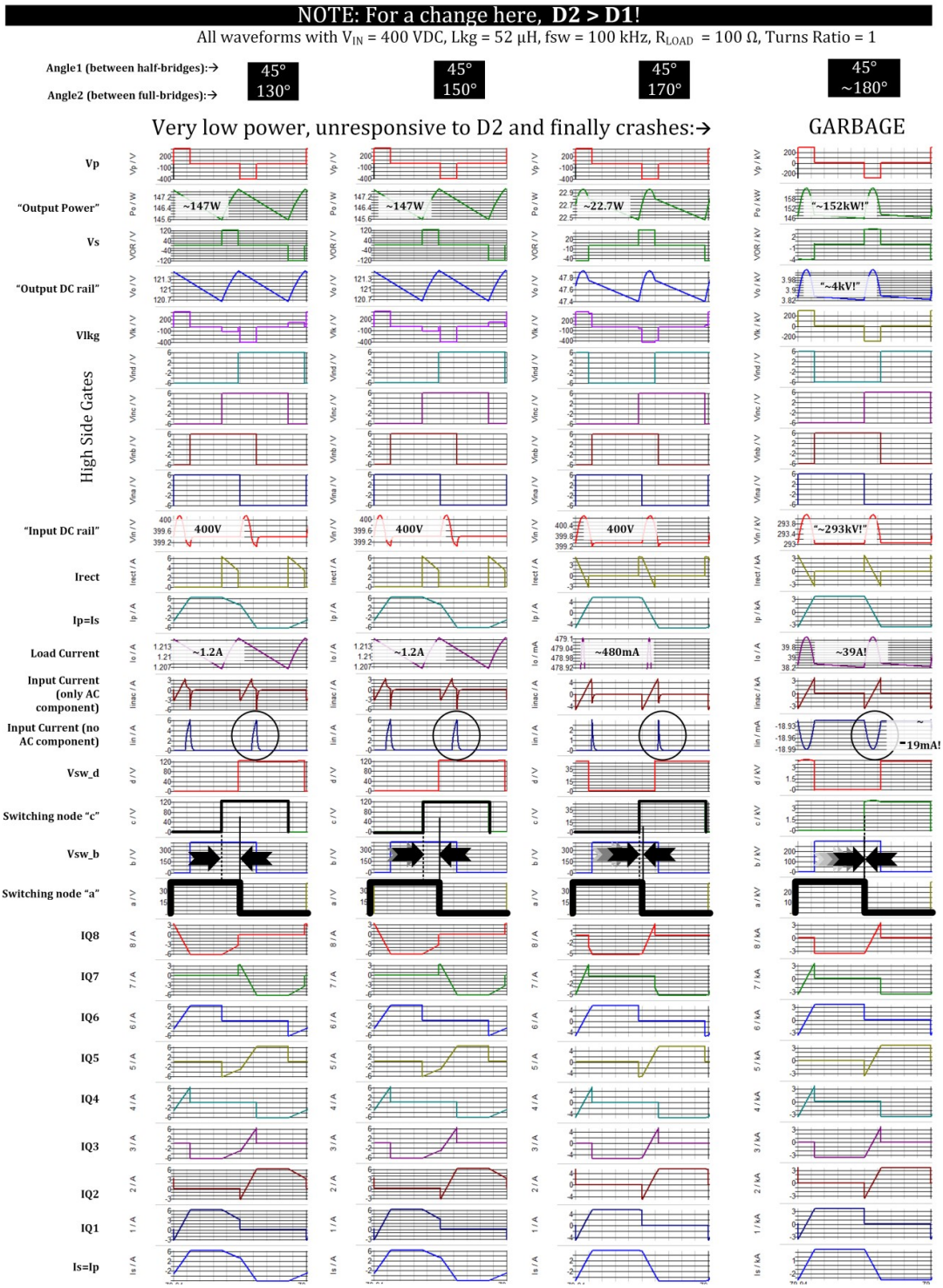


Figure 2.21: What happens if D2 is made greater than D1

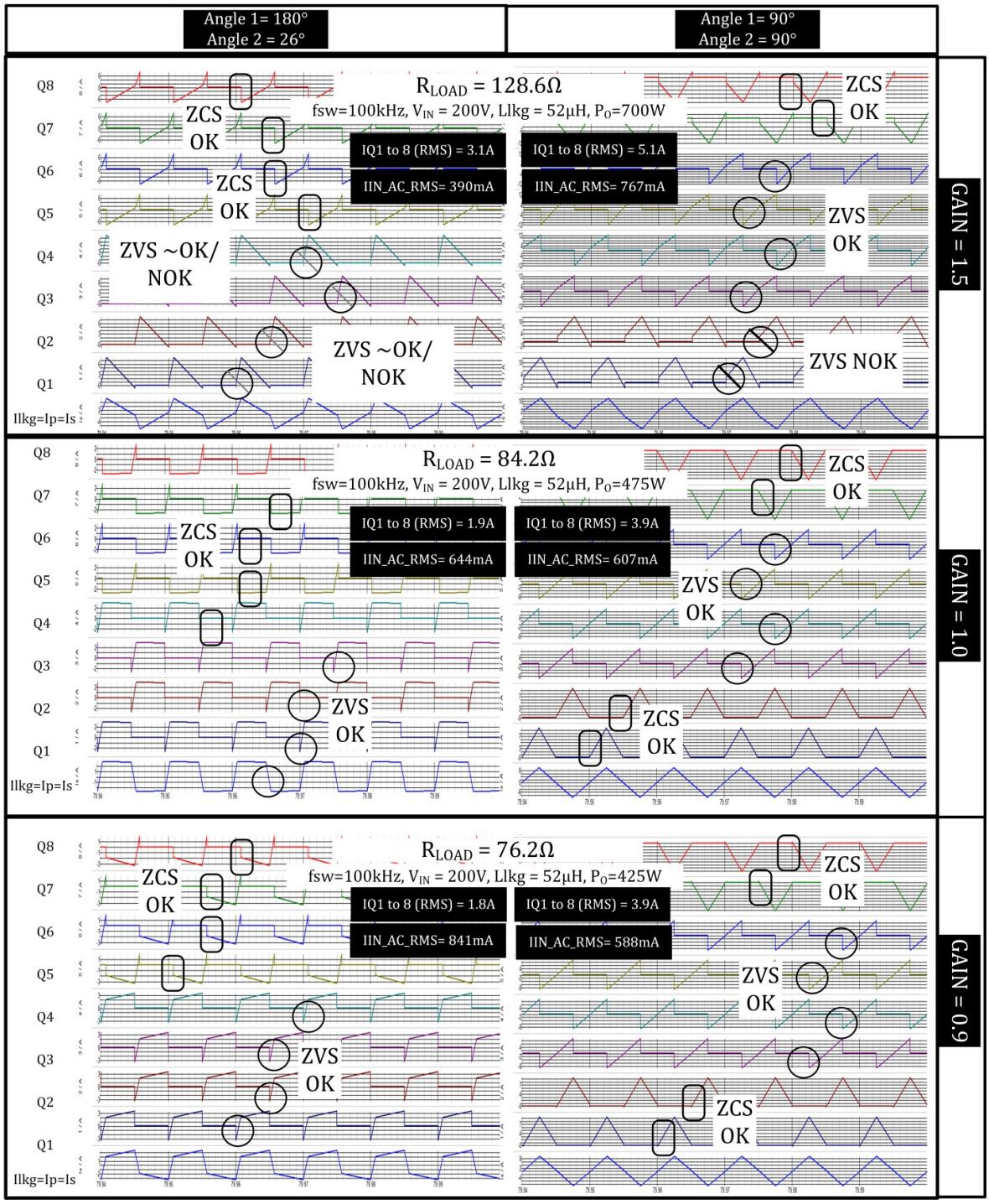


Figure 2.22: Comparing the two possibilities for setting Angle 1

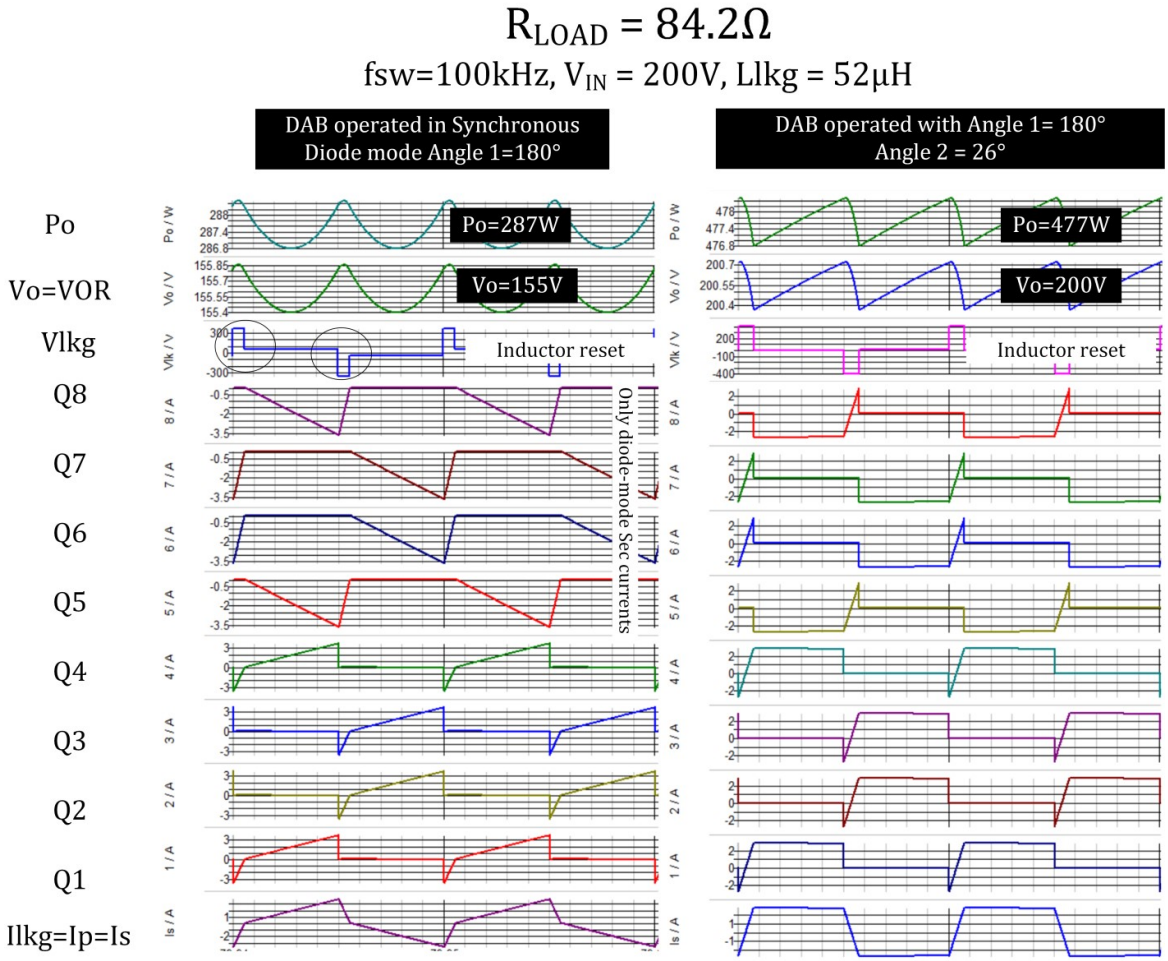


Figure 2.23: Operating the DAB in synchronous mode

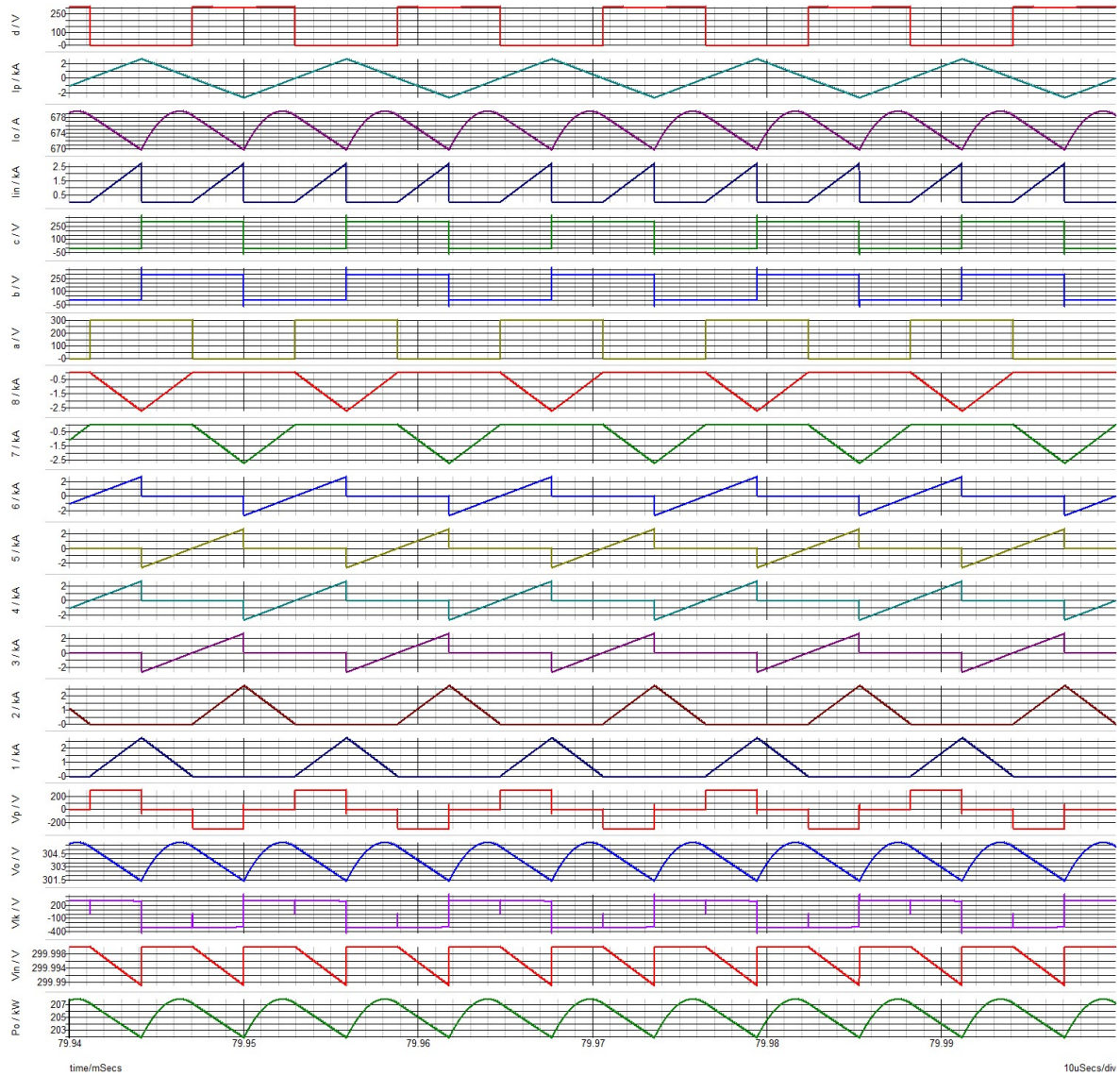


Figure 2.24: SIMPLIS simulation of 200kW converter for Angle 1= Angle 2 = 90°

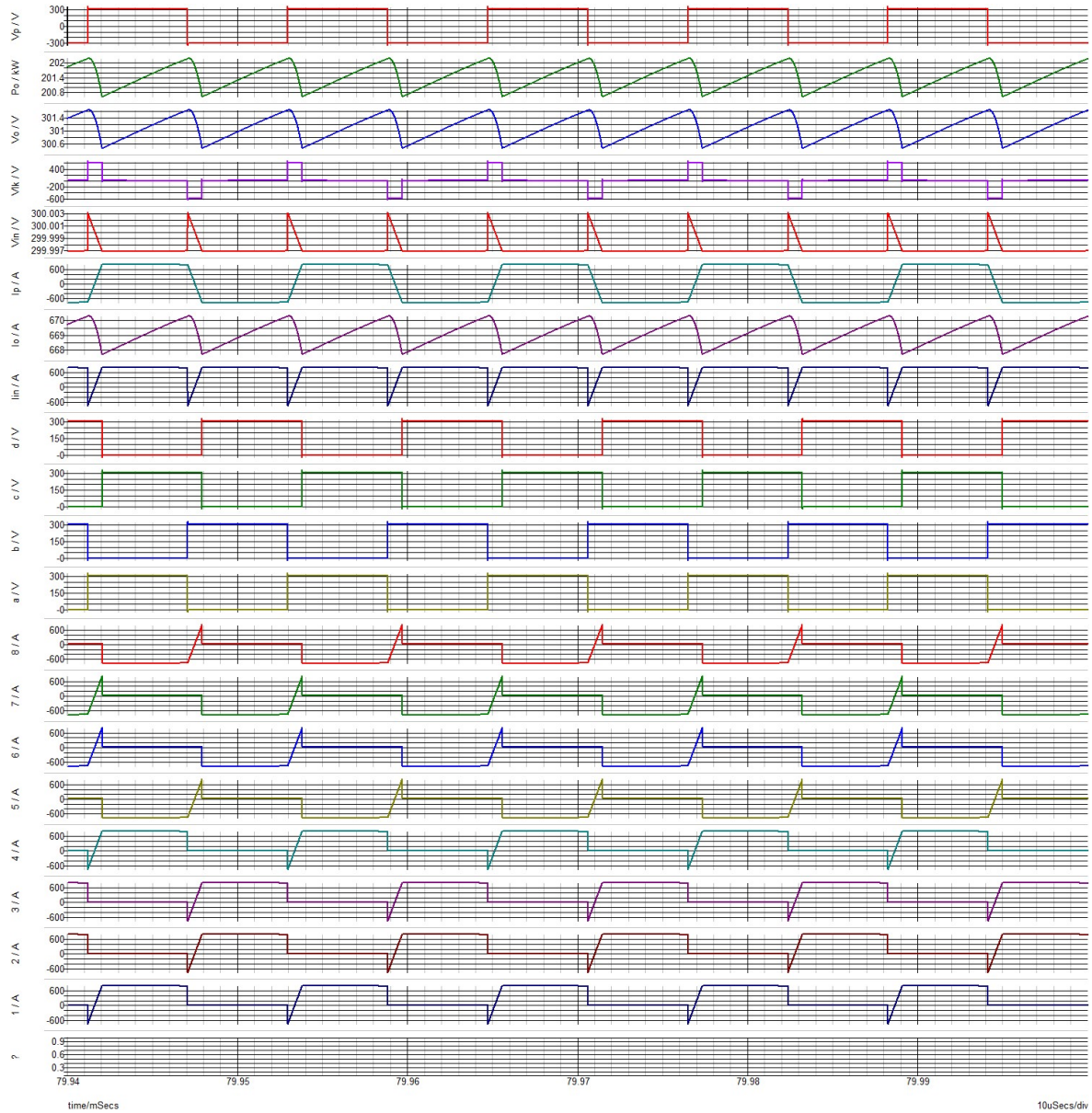


Figure 2.25: SIMPLIS simulation of 200kW converter for Angle 1= 180° and Angle 2=26°

Chapter 3
Small-signal modelling
(Guest Chapter by Nicola Rosano)

Introduction

In 2013 I got my Master's degree in Power Electronics. The same year I became part of Thales Alenia Space power electronics team in Rome, Italy and I was almost sure to be ready designing power systems for satellite applications. None of all the beliefs I had during my professional career was more wrong than that one.

The second working day the boss showed me an A1 foil with something about 500-600 components. It was the schematic of a "simple" Flyback converter used for satellite applications. I realized immediately I was not ready for that job. The scenario becomes worse one week later when the boss assigned to me new specs for a new project. It was always a flyback converter.

At that time, and nowadays as well, one of the most crucial aspects of a power stage was (and is) the control section design. A lot of components linked all together just to get what we call simply "current mode control" or "voltage mode control". Something that by using an electrical simulator we obtain with a small bunch of components. A to Z, Second Edition Chapter 12 focuses on the feedback loop analysis and stability concepts and gives to the reader a basic understanding about two different control strategies: voltage model control (VM) and current mode control (CM). Both assumed in continuous conduction mode (CCM) for sake of simplicity.

Independently from the adopted control strategy, a DC-DC or Offline converter design needs to be compliant to different requirements. Two of them have been fully explained in Chapter 12 and they are the a) Loop gain and phase that have an impact on the stability and the b) audio-susceptibility that 'quantify' the transmission of the noise from the input section to the output section. By the way there are two other important "transfer functions" intrinsically related to a power stage general requirements: they are c) the output regulation related to the output impedance and d) the sensitivity of the power system to input filter interaction related to input impedance. With these four transfer functions we will be able to describe the converter behavior almost completely. This is one of the reasons why introducing a power stage linear model can be very useful to the reader in both directions: understand how the converter fully behaves into the frequency domain AND how to properly design the control loop.

Let's start by focusing on the control to output plant transfer function G(s) assuming Voltage Mode Control in Continuous Conduction Mode (CCM). The intuitive way to get a simplified closed form equation for G(s) is to multiply all terms involved in the loop sketched on Figure 12.10 Page 457 of A to Z, Second Edition.

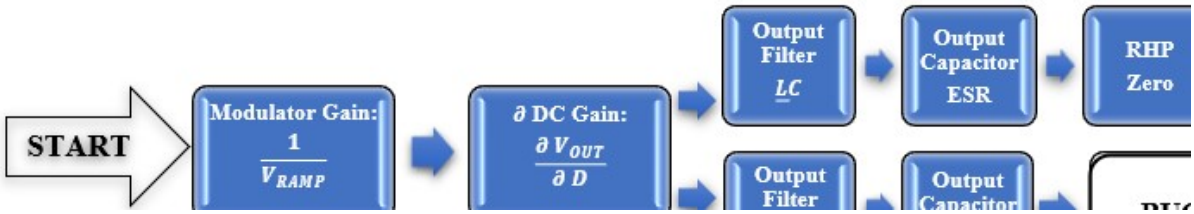


Figure 3.1: Intuitive diagram to identify all elements involved into the control to output TF evaluation for basic topologies (Right Half Plane zero included). Voltage mode CCM assumed.

At this point some questions are spontaneous:

- Why, for "not-buck" topologies, the equivalent inductance for the plant transfer function equation needs to be considered equal to $\underline{L}=L/(1-D)^2$?
- Where does the right half plane zero come from?
- Where do formulas come from?
- Are the formulas correct?
- How can we extract Bode diagrams starting by a non linear circuit in order to verify all previous equations?

All these questions make the VM control theory not so intuitive. Sketching time domain waveforms by using pen and paper is simple but extracting gain, poles and zeros for each topology is another story because power converters are non-linear systems by their nature.

The situation becomes worse if we focus on the Current Mode case in CCM. The current mode control has its pros and cons with respect to the voltage mode control. A detailed study has been performed and presented in Chapter 12 of A to Z, Second Edition. I invite the reader to focus on two sentences of that chapter in particular:

- It can be shown that using this technique (CM) the inductor effectively goes "out of the picture" in the sense that there is no double LC pole anymore.

Is this true? how is it possible to verify this statement? (I ask)!

- Everyone seems to agree the current-mode control alters the poles of the system compared to voltage-mode control, but the zeros of voltage-mode control remain unchanged.

Let's focus on the previous two bullets. If zeros are unchanged it means RHP is present if Current Mode Control (CMC) is used. The RHP by its nature involves the inductance to be calculated. But at the same time the Current Mode approach "pushes the inductor out of the picture"! It seems a paradox.

A to Z, Second Edition, Page 503 reports a table that explains all involved quantities to plot the control to output TF in current mode. Surely the current mode control is less intuitive than the voltage mode control but if we try to sketch the same flow chart as we did for the voltage mode control we get:

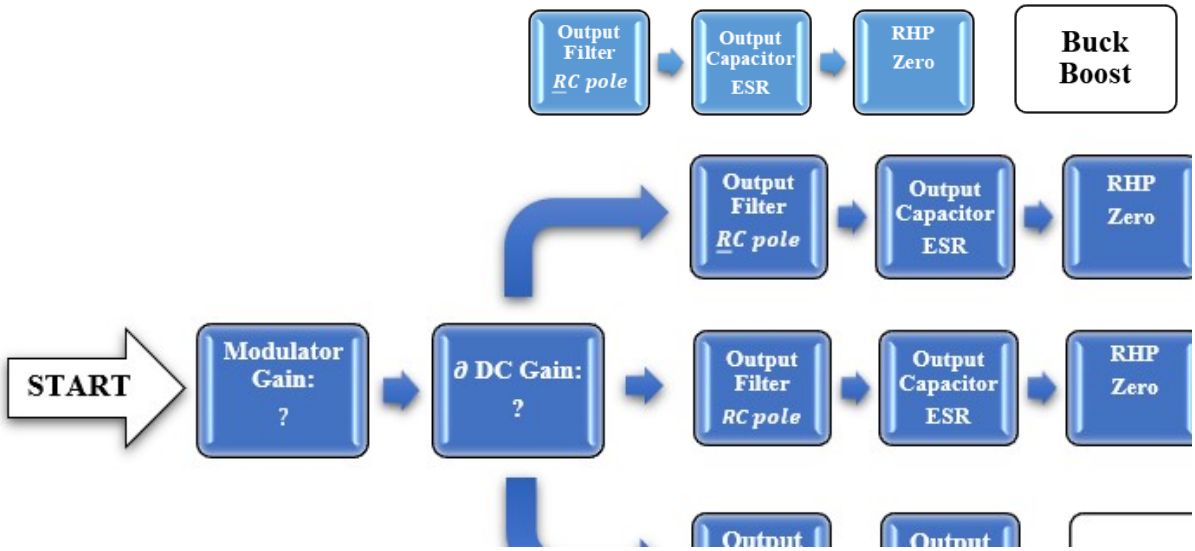


Figure 3.1: Intuitive diagram to identify all elements involved into the control to output TF evaluation for basic topologies (RHP zero included). Peak Current mode CCM assumed.

Now, observing the above diagram added questions complicates the situation a bit more.

- Why there is another difference into the RC pole in the sense that it is equal for Buck and Boost topology and differs for Buck-Boost?
- How the Modulator Gain (G0 on Page 503 of A to Z, Second Edition) has been calculated?
- More than everything: How is it possible that the current mode control time domain waveforms are identical to the one we get using voltage mode control while transfer functions – frequency responses – differ?

Extracting all transfer functions for a specific power stage that uses a specific control method is something not so easy. Linear modeling will be introduced to fix all these questions showing a unique approach valid for all topologies. The only thing we need to do is to map the cycle by cycle non-linear circuit (that involves Mosfets and diodes) into something that uses linear components or controlled sources. Generally speaking we can say linear modeling approach allows one to:

- Get a full understanding about how to design the feedback loop
- Evaluate parasitic elements effect on power stage stability
- Predict the transient response including undershoot and overshoot if a load-step is applied
- "Verify" the plant behavior independently by the control strategy used or the power stage selected
- Simplify the full converter analysis

What is not possible with this method:

- The presented technique doesn't work with variable frequency converters. We will see a linear model automatically locks the switching frequency to be extracted. It is evident that for variable frequency converters we would have "infinite" linear models. For those topologies all transfer functions are calculated by using "*piecewise linear functions".* The predicted transfer function is obtained by exciting the converter at different frequencies spaced linearly storing two info: gain and phase for each frequency point. A feature actually present in almost all modern power system simulators (PSIM, PLECS, SIMPLIS).

Voltage Mode Control PWM switch

Despite we did our job even for the DCM case for both control strategies, in alignment with the entire book, the Voltage Mode linear model will be discussed for CCM case only for sake of simplicity. Three easy steps are needed to get this model:

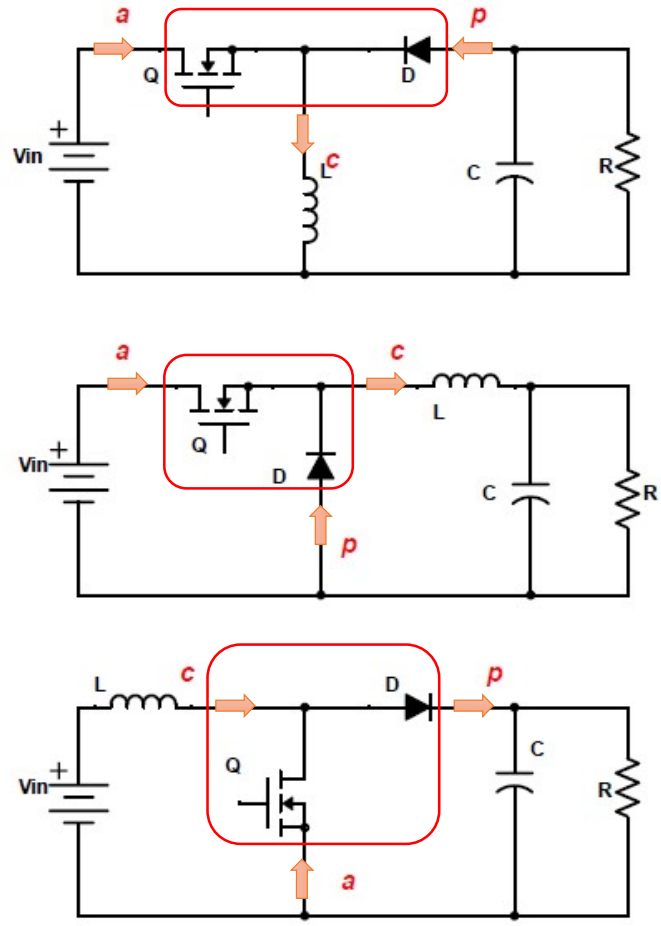
a) Identify non-linear elements and define active, passive, and common terminals

common: is the common terminal between the diode and the power switch
active: is the other end terminal related to the Active Switch (MOS)
passive: is the other end terminal related to the Passive Switch (Diode)

b) Average the current waveforms across the PWM Switch.

Refer to the Table 4.1 (Page 184) of A to Z, Second Edition. The Buck Converter case will be discussed as reference example to extract a linear model. Then the PWM switch "fully invariant" property will be used to manage Boost, Buck Boost, and derived (isolated) topologies.

$$I_{active} = I_a = I_{SWITCH} = I_0 \cdot D = I_c \cdot D$$
$$I_{common} = I_O = I_c$$
$$I_{passive} = I_p = I_{DIODE} = I_O \cdot (1 - D) = I_c \cdot (1 - D)$$



Buck Boost
Current goes into the *a* terminal
Current goes out to the *c* terminal
Current goes into the *p* terminal

Buck
Current goes into the *a* terminal
Current goes out to the *c* terminal
Current goes into to the *p* terminal

Boost
Current goes into the *a* terminal
Current goes into the *c* terminal
Current goes out to the *p* terminal

Figure 3.3: Identifying PWM switch terminals for the three basic topologies

c) Average the voltage waveforms across the PWM Switch and build it

$V_{ap} - V_{tn}$

$V_{cp} = V_{out} = D \cdot V_{in}$ (note the voltage mean value on the inductor is zero)

$V_{ac} = V_{in} - V_{out}$

It will be shown the DC bias point can be "emulated" by using a two-port cell: a current controlled current source on the primary and a voltage controlled voltage source on the secondary. Note also that both controlled sources can be replaced by an ideal transformer with turns-ratio N1:N2=1:D.

The Buck Large signal model is shown on the next page.
If the PWM switch is "fully invariant" it means the large signal model of a cycle by cycle converter can be obtained by replacing the non-linear three port cell (which identifies the couple MOS + DIODE) with a linear invariant three terminal cell represented by an ideal transformer with turn ratio N1:N2 = 1:D. The power of this approach is that it is applicable to all topologies.

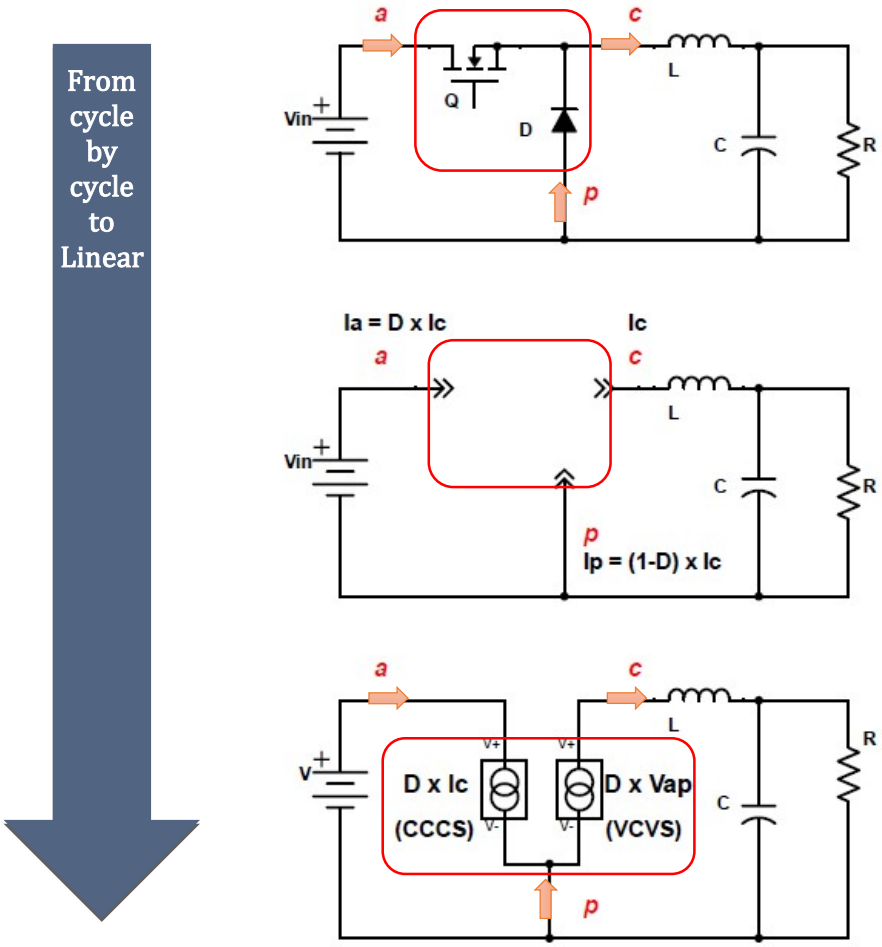


Figure 3.4: From cycle by cycle model to linear model. Case Study: Buck converter in Voltage mode (CCM assumed)

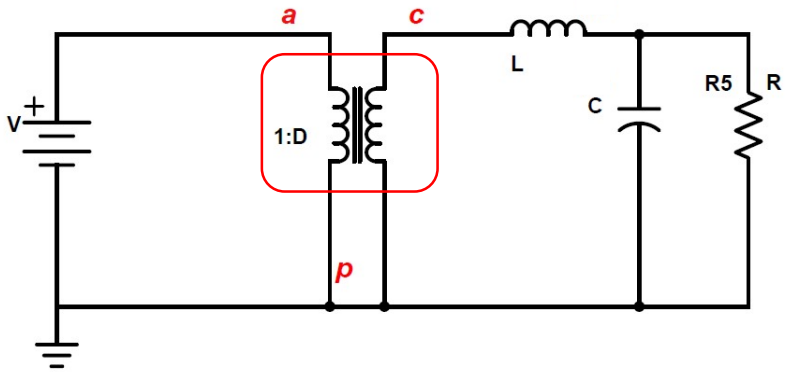


Figure 3.5: Same model we got in **Figure 3.4** by placing the ideal transformer

A small signal version is acquired by linearizing the large signal model. We need to focus on the inserted "controlled sources" only because all the rest of the network is linear per se. One Current Controlled Current Source is $I_a = D \cdot I_c$. It involves two variables: D and I_c. Applying the partial derivatives procedure we get:

$$\hat{I}_a = \frac{\partial}{\partial D}(D \cdot I_c) \cdot \hat{D} + \frac{\partial}{\partial I_c}(D \cdot I_c) \cdot \hat{I}_c \rightarrow \hat{I}_a = I_c \cdot \hat{D} + D \cdot \hat{I}_c$$

Same story for the Voltage Controlled Voltage Source, two variables are involved: D and V_{ap}. So applying the partial derivatives procedure we get:

$$\hat{V}_{cp} = \hat{V}_{out} = \frac{\partial}{\partial D}(D \cdot V_{ap}) \cdot \hat{D} + \frac{\partial}{\partial V_{ap}}(D \cdot V_{ap}) \cdot \hat{V}_{ap} \rightarrow \hat{V}_{cp} = \hat{V}_{out} = V_{ap} \cdot \hat{D} + D \cdot \hat{V}_{ap}$$

Collecting both equations together we get:

$$\begin{cases} \hat{I}_a = I_c \cdot \hat{D} + D \cdot \hat{I}_c \\ \hat{V}_{cp} = V_{ap} \cdot \hat{D} + D \cdot \hat{V}_{ap} \end{cases}$$

From linearization process, two added controlled sources need to be introduced to get the AC model

The AC m same idea into the la

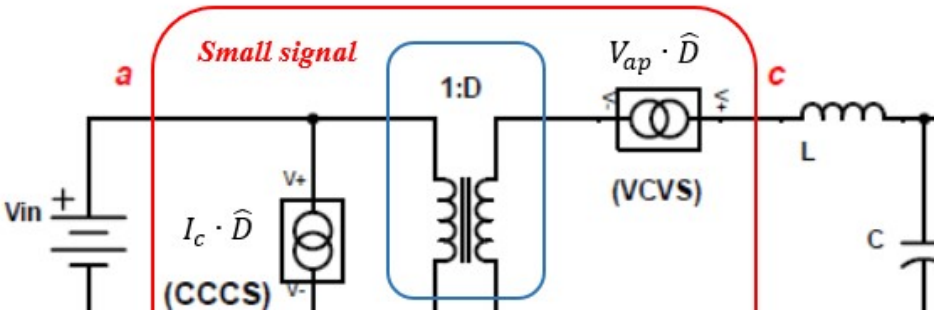


Figure 3.2: AC three terminal invariant PWM switch cell valid for Voltage Mode CCM assumed

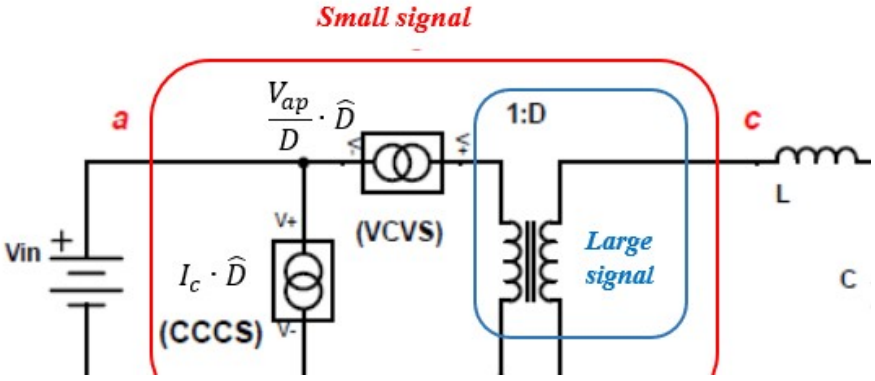


Figure 3.3: Conventional AC three terminal invariant PWM switch cell valid for Voltage Mode CCM assumed. VCVS placed on the primary

PWM switch at work: Voltage Mode CCM assumed (Buck Converter Example)

Input parameters

f_{SW} [Hz] = $100 \cdot 10^3$; Vin [V] = 50; Vout [V]= 30; L [H] = $200 \cdot 10^{-6}$; C [F]=1.25$\cdot 10^{-6}$; R [Ω]=50; V_{ramp} [V]=3

$D = V_{out}/V_{in} = 0.6$; $I_O = V_{out}/R = 0.6$; Mod. gain = $1/V_{ramp} = 333mA$

a) Identify non-linear elements and define active, passive, and common terminals
Please refer to Fig 3

b) Average the current waveforms across the PWM Switch

$I_{active} = I_{SWITCH} = I_O \cdot D = 360mA$

$I_{common} = I_O = 600mA$

$I_{passive} = I_O \cdot (1 - D) = 240mA$

c) Average the voltage waveforms across the PWM Switch and built it (Fig 7)

$V_{ap} = V_{in} = 50$

$V_{cp} = V_{out} = 30$

$V_{ac} = V_{in} \quad V_{out} = 20$

The following simulation's proof of concepts have been performed using TINA-TI Spice-based simulator by DesignSoft company. Its free version is much more user friendly than LTspice (free as well) and powerful if "industrial" version is used. A free TINA-TI version can be downloaded here: *http://www.ti.com/tool/TINA-TI.* The universal three terminal PWM switch in Voltage mode CCM is shown below:

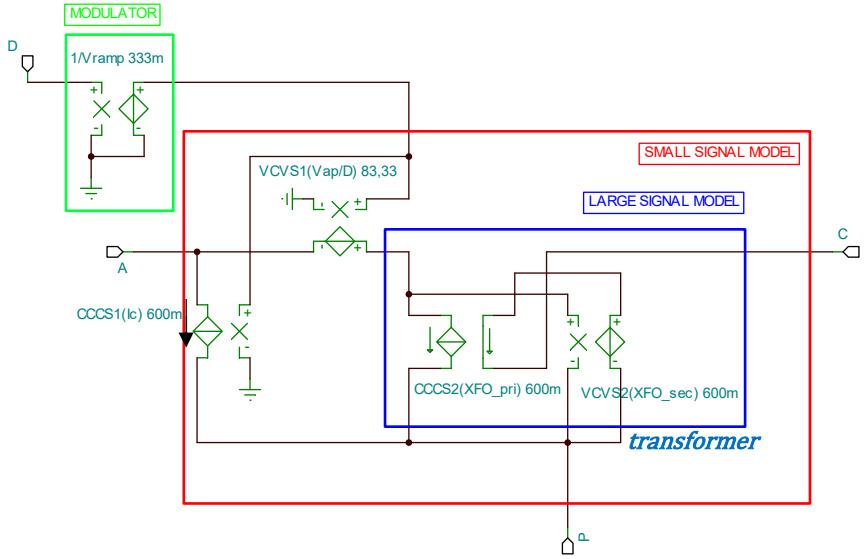


Figure 3.8: Universal AC (and DC?) PWM switch cell structure valid for CCM VM.

Checking the DC bias point we can see a perfect match with calculated values. Further, due to the insertion of the modulator gain, the control voltage is scaled up by $V_{ramp} \cdot$ Duty = $3 \cdot 0.6$ = 1.8. It is set externally.

Note: There are four controlled sources and four terminals: active, passive, common, duty (external). Only controlled sources values change, the "skeleton" is always the same. Further the modulator gain $1/V_{ramp}$ has been included into the model (refer to Page 458 of A to Z, Second Edition)

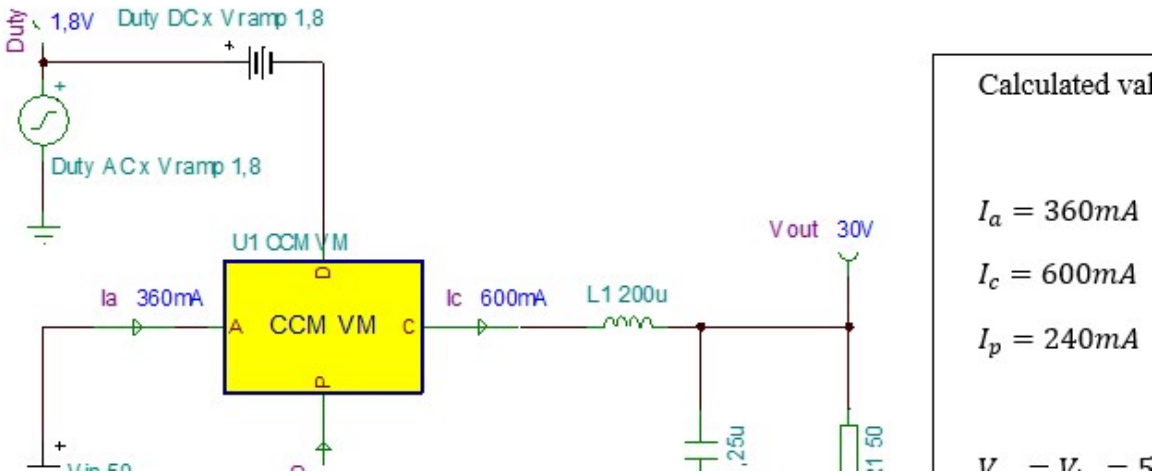


Figure 3.9: Checking DC bias working point

Some considerations are mandatory:

- First, the duty cycle is intrinsically included in the model. Its value is D = 0.6 and is visible in the controlled sources CCCS1, CCCS2 and VCVS2 on Fig. 8. It means the external control variable (terminal "D" on the PWM yellow block Fig. 9) shall be zero (or that point shall be grounded) to properly evaluate the DC bias point. But doing this would be no good for us because when we are called to close the loop we need the "control voltage" terminal to be set externally (it will be the output of the feedback OPAMP) and cannot be grounded.

- The situation is complicated by the fact that the linearization process we performed to get the 3-terminal small signal cell has the following disadvantage: the DC point is lost during the linearization process. So, in theory, the large signal model shall be used to evaluate the DC bias point and the small signal model shall be used to evaluate the AC frequency response. Vorperian AC model – in voltage mode – suffers in storing the DC bias point.

Is it possible to use one circuit to do both? Using different simulation files and different circuits to model the same power stage just to perform DC and AC analysis is annoying, and errors can be produced by using different files for the same power stage (change in parameters, or conditions etc.). The answer is: yes! And the circuit setup is shown on Fig 9. Two external generators needed: one works in DC only, the other one works in both AC and DC. During the DC analysis they cancel each other shutting down VCVS1 and CCCS1 (the transformer acts alone). During the AC analysis all DC values are "not visible" and the complete AC small signal model is fully restored.

PWM Switch flexibility is compared with the control to output TF and the audio susceptibility transfer function ("TF") reported in A to Z, Second Edition Pages 463 to 464 for the Buck Converter. Note we call “Gc2o” the control to output transfer function and “GL2o” line to output transfer function.

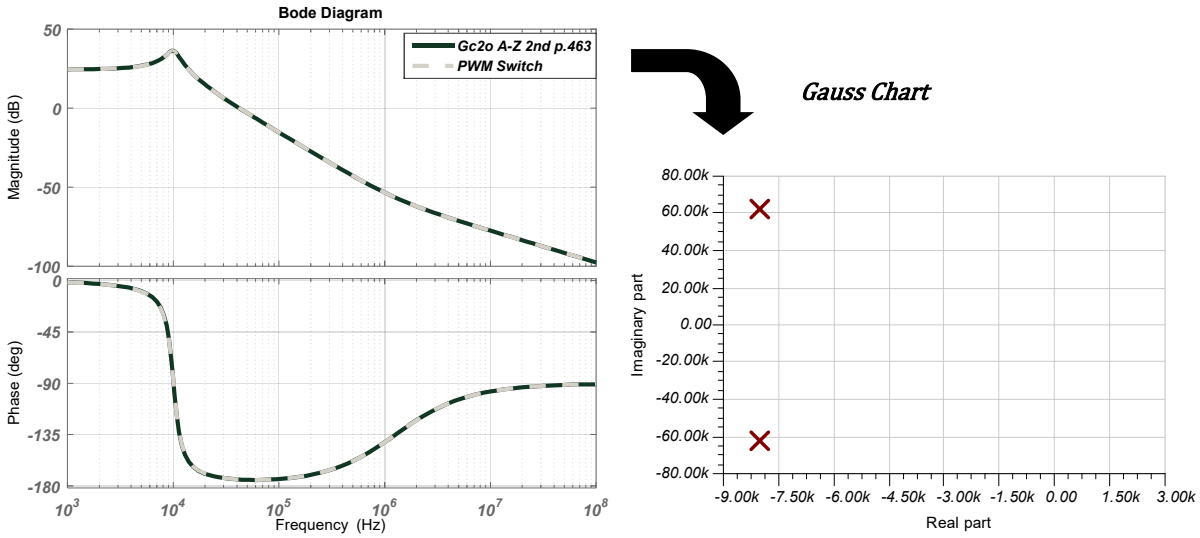
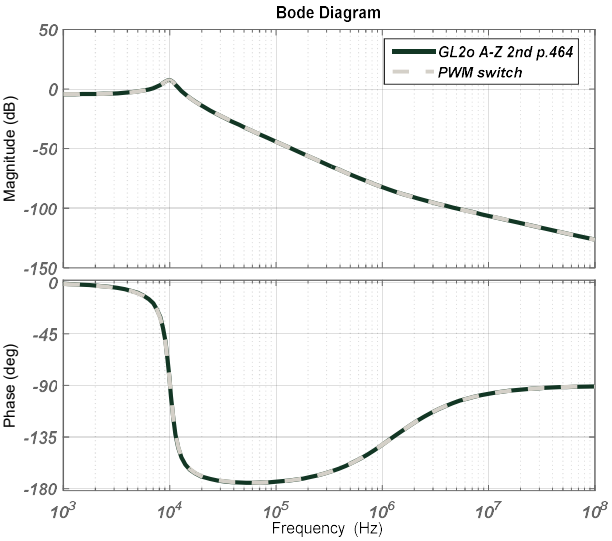


Figure 3.10: Comparing PWM switch Gc2o TF versus Maniktala G(s) Page 463 of A to Z, Second Edition. Curves are perfectly superimposed.



No Gauss Chart needed for the line to output TF because poles and zeros are identical to the control to output TF.
*See **Table 4.1**.*

Figure 3.11: Comparing PWM switch GL2o TF versus Maniktala audio-susceptibility G(s) Page 464 of A to Z, Second Edition. Curves are perfectly superimposed.

Note the presence of complex poles (crosses). No zeros (circles) shown on the Gauss chart for buck topology: Right Half Plane zero (RHP) absent and ESR zero absent. In particular no ESR zero is present because we didn’t include it in **Figure 3.9**. It will be automatically placed on the Gauss chart if included into the simulation (see Boost and Buck Boost Examples below).

Worked Examples

BOOST Example

Input parameters

f_{SW} [Hz] = $100 \cdot 10^3$; Vin [V] = 10; V_{out} [V]= 40; L [H] = $100 \cdot 10^{-6}$; C [F]= $470 \cdot 10^{-6}$; R [Ω]=20; V_{ramp} [V]=5; ESR_C [Ω]=$100 \cdot 10^{-3}$; ESR_L=$70 \cdot 10^{-3}$ [Ω];

$D = 1 - \frac{V_{in}}{V_{out}} = 0.75$; $I_O - \frac{V_{OUT}}{R} - 2$; Mod_gain = $1/V_{ramp}$= 0.2

a) Identify non linear elements and define active, passive, and common terminals. *Pay attention to the reference signs used for the currents coming in or out to the PWM switch cell used for the Buck Converter! They need to be the same to profit of the fully invariant feature*. Refer to **Figure 3.3**.

b) Average the current waveforms across the PWM Switch (Table 4.1 A to Z, Second Edition).

$I_{active} = -I_{Switch} = -I_O \cdot \frac{D}{1-D} = -6$

$I_{common} = -I_L = -\frac{I_O}{1-D} = -8$

$I_{passive} = -I_{Diode} = -I_O = -2$

c) Average the voltage waveforms across the PWM Switch and built it (Fig 7)

$V_{ap} = -V_{OUT} = -40$

$V_{ac} = -V_{IN} = -10$

$V_{cp} - V_{IN} - V_{OUT} - -30$

External Excitation = $D \cdot V_{RAMP} = 3.75$

BUCK-BOOST Example

Input parameters

fsw [Hz] = $100 \cdot 10^3$; Vin [V] = 35; Vout [V]= 100; L[H] =$260 \cdot 10^{-6}$; C [F]= 5 $\cdot 10^{-6}$; R [Ω]=150; Vramp [V]=5; ESR_C [Ω]=$85 \cdot 10^{-3}$; ESR_L=$85 \cdot 10^{-3}$ [Ω];

$D = \frac{V_{out}}{V_{in}+V_{out}} = 0.74$; $I_O - \frac{V_{OUT}}{R}$ –666mA; Mod_gain=$1/V_{ramp}$= 0.2

a) Identify non-linear elements and define active, passive, and common terminals. *Pay attention to the reference signs used for the currents coming in or out to the PWM switch cell used for the Buck Converter! They need to be the same to profit of the fully invariant feature*. Refer to **Figure 3.3**.

b) Average the current waveforms across the PWM Switch (Table 4.1 of A to Z, Second Edition)

$I_{active} = I_{Switch} = I_O \cdot \frac{D}{1-D} = 1.9$

$I_{common} = I_L = \frac{I_O}{1-D} = 2.56$

$I_{passive} = I_{Diode} = I_O = 666m$

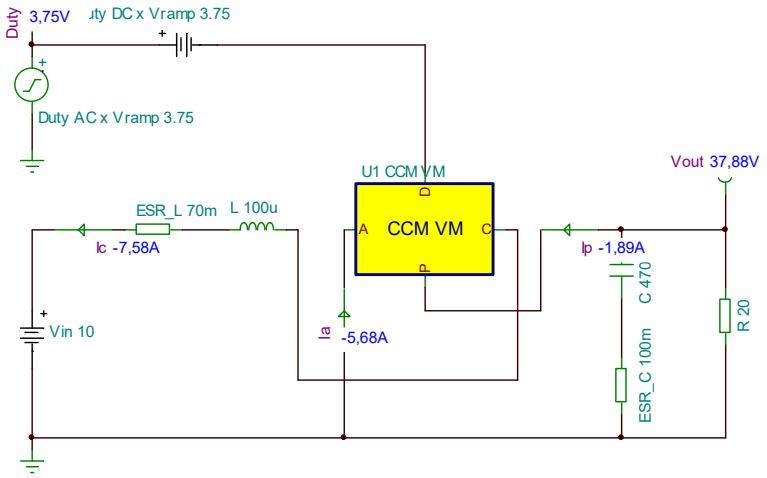
c) Average the voltage waveforms across the PWM Switch and built it (Fig 7)

$V_{ap} = V_{IN} + V_{OUT}\ (Buck\ Boost\ is\ inverting) = 135$

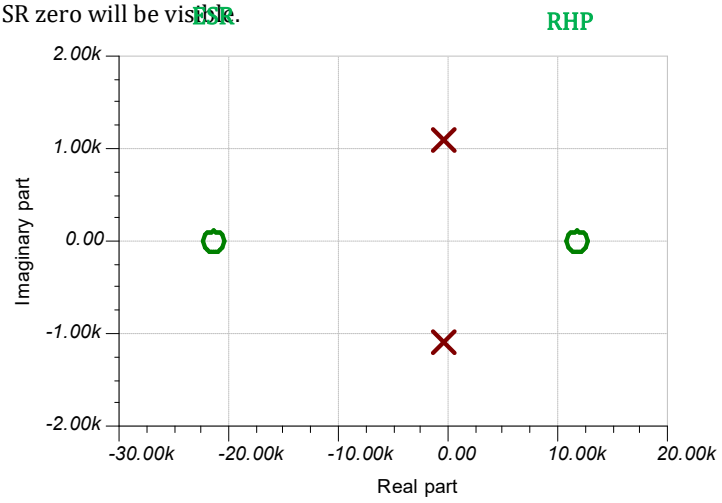
$V_{ac} = V_{IN} = 35$

$V_{cp} = V_{OUT}\ (Buck\ Boost\ is\ inverting) = 100$

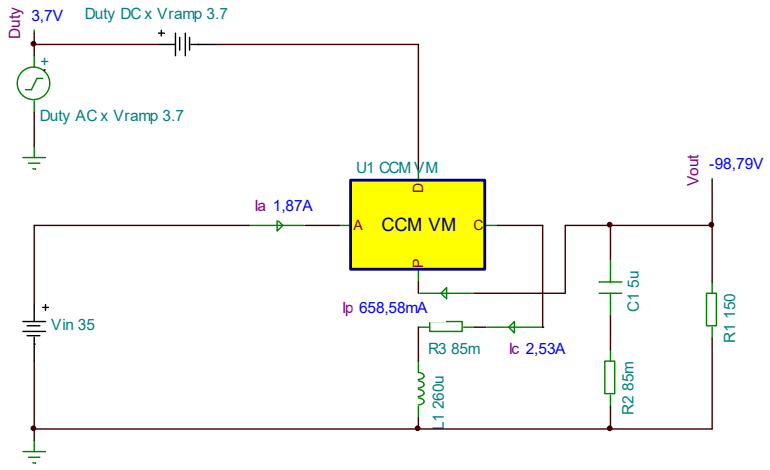
External Excitation = D · V_{RAMP} = 3.7



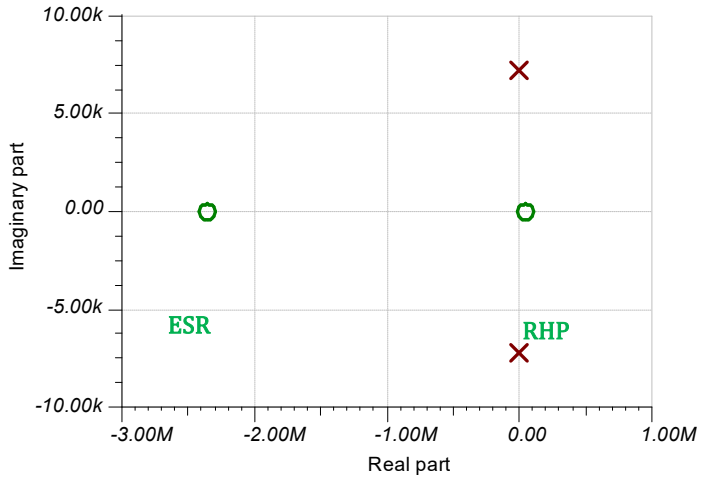
Note1: DC Bias point not lost if AC model is used. It has an "invariant" skeleton, but different values (Fig 8).
Note2: DC Vout almost equal to the ideal value due to parasitics ESR_L. Also the ESR zero will be visible.



RHP zero, ESR zero, complex poles visible on Gauss chart



Note1: DC Bias point not lost if AC model is used. It has an "invariant" skeleton, but different values (Fig 8).
Note2: DC Vout almost equal to the ideal value due to parasitics ESR_L. Also the ESR zero will be visible.



RHP zero, ESR zero, complex poles visible on Gauss chart

Buck	Boost	Buck Boost (inverting)
$\cdot \frac{\left(1+\frac{s}{\omega_{ESR}}\right)}{\left(\frac{s}{\omega_0}\right)^2+\left(\frac{s}{\omega_0\cdot Q}\right)+1}$ $\frac{1}{\sqrt{LC}}$; $\omega_0 \cdot Q = \frac{R}{L}$	$G(s) = \frac{1}{V_{RAMP}} \cdot \frac{V_{in}}{(1-D)^2} \cdot \frac{\left(1-\frac{s}{\omega_{RHP}}\right)\cdot\left(1+\frac{s}{\omega_{ESR}}\right)}{\left(\frac{s}{\omega_0}\right)^2+\left(\frac{s}{\omega_0\cdot Q}\right)+1}$ $\omega_0 = \frac{1}{\sqrt{LC}}$; $\omega_0 \cdot Q = \frac{R}{L}$; $L = \frac{(1-D)^2}{L}$ $\omega_{ESR} = \frac{1}{C \cdot ESR_C}$; $\omega_{RHP} = \frac{R}{L} \cdot (1-D)^2$;	G(s) *same as for Boost with negative sign* ω_0; ω_{ESR}; Q; L; *same as for Boost* $\omega_{RHP} \rightarrow$ differ $\rightarrow \frac{R}{L} \cdot \frac{(1-D)^2}{D}$
$\frac{D \cdot \left(1+\frac{s}{\omega_{ESR}}\right)}{\left(\frac{s}{\omega_0}\right)^2+\left(\frac{s}{\omega_0\cdot Q}\right)+1}$ Q *same as above*	$G(s) = \frac{1}{1-D} \cdot \frac{\left(1+\frac{s}{\omega_{ESR}}\right)}{\left(\frac{s}{\omega_0}\right)^2+\left(\frac{s}{\omega_0\cdot Q}\right)+1}$ ω_0, Q same as above (with L)	$G(s) = \frac{-D}{1-D} \cdot \frac{\left(1+\frac{s}{\omega_{ESR}}\right)}{\left(\frac{s}{\omega_0}\right)^2+\left(\frac{s}{\omega_0\cdot Q}\right)+1}$ ω_0; Q; *same as for Boost*
$\frac{\left(\frac{s}{\omega_0}\right)^2+\left(\frac{s}{\omega_0\cdot Q}\right)+1}{\left(1+\frac{s}{\omega_{RC}}\right)}$ *e as above;* $\omega_{RC} = \frac{1}{RC}$	$Z_{in}(s) = R \cdot (1-D)^2 \cdot \frac{\left(\frac{s}{\omega_0}\right)^2+\left(\frac{s}{\omega_0\cdot Q}\right)+1}{\left(1+\frac{s}{\omega_{RC}}\right)}$ ω_0, Q same as above (with L); $\omega_{RC} = \frac{1}{RC}$	$Z_{in}(s) = \frac{R \cdot (1-D)^2}{D^2} \cdot \frac{\left(\frac{s}{\omega_0}\right)^2+\left(\frac{s}{\omega_0\cdot Q}\right)+1}{\left(1+\frac{s}{\omega_{RC}}\right)}$ ω_0; Q; ω_{RC} *same as for Boost*

Table 4.1: Collecting Closed Form Equations for the VM CCM using PWM Switch model

Comparing Linear Model vs. Cycle-by-Cycle model

Last thing to do before closing this chapter is to have a perception about what happens if we compare the Linear Model VS the Cycle by Cycle model for a case study. Intuitively the first advantage is the speed analysis. We get estimated results in seconds compared to cycle by cycle real model. Second, the linear model should remove the AC high frequency component. We can verify these statements closing the loop for the Buck case study as example.

Referring to A to Z, Second Edition, Second Edition Figure 12.16 Page 480 and Figure 12.19 Page 489 we know the type 3 compensator gives the higher phase boost and introduces 2 poles and 2 zeros (plus one pole at the "origin"). As suggested, we place 2 zeros at LC pole frequency; 1 pole at the ESR zero; 1 pole at $10 \cdot f_{cross}$. Setting f_{cross} equal to 15kHz and assuming ESR_C=100mΩ for the Buck case study we get:

$$f_{p1} \rightarrow \frac{1}{2\pi \cdot ESR_C \cdot C} = 1.27MHz;\ f_{p2} \rightarrow 10 \cdot f_{cross} = 150kHz;\ f_{z1} = f_{z2} \rightarrow \frac{1}{2\pi \cdot \sqrt{LC}} = 10kHz$$

By locking R1=10kΩ we get: C1 = 45nF, R2 = 356Ω, C2 = 352pF, R3 = 666Ω, C3 = 1.6nF.

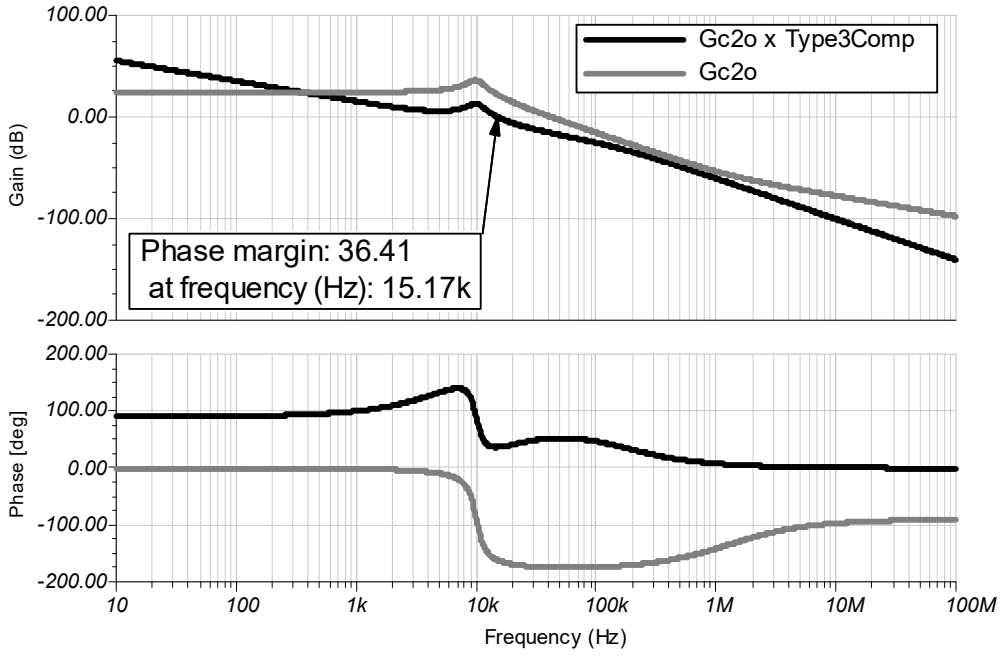
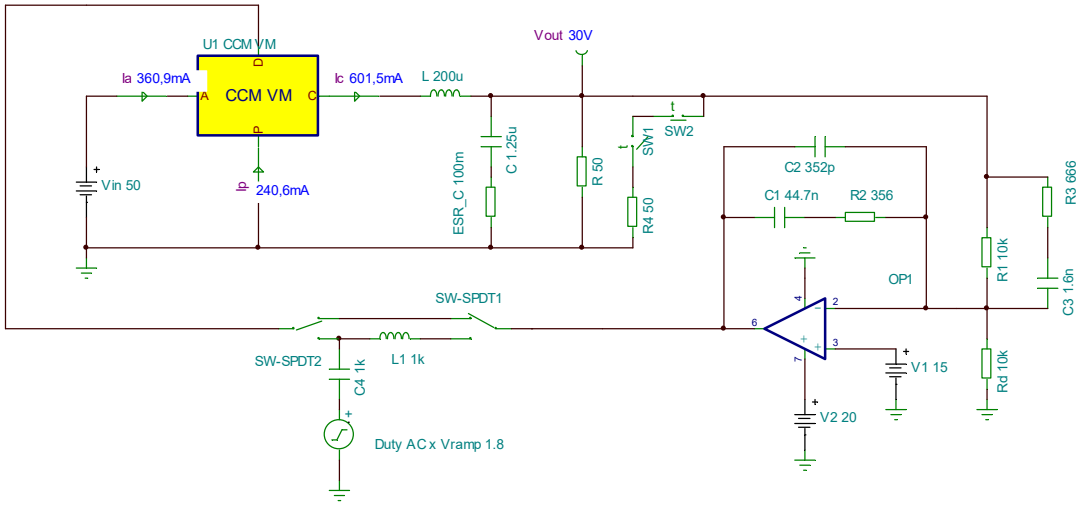


Figure 3.12: AC analysis: Plant (grey) and Plant + Compensator (black)

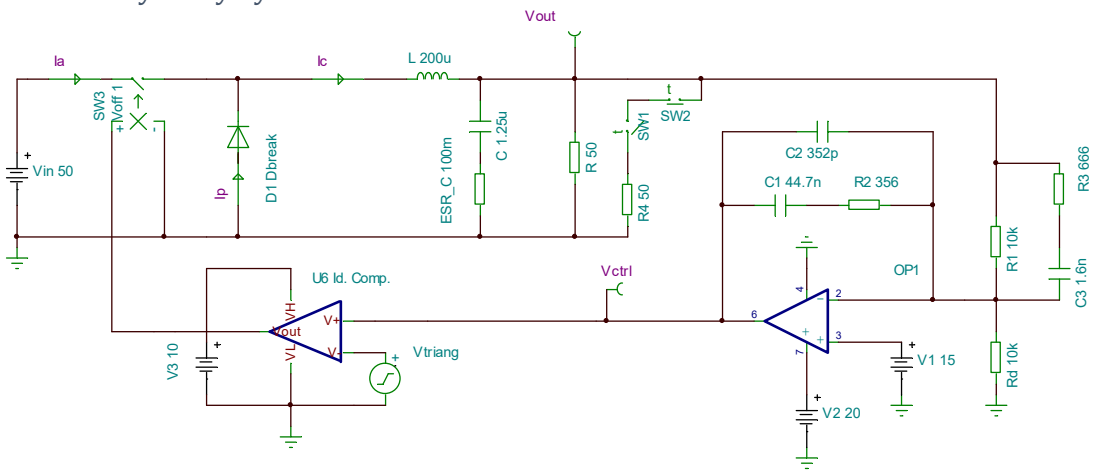
If we remember Bode criteria we can say the system is stable. At the same time we know 36° degrees phase margin (for this particular case) is not so good. Practically speaking a margin lower than 45° (worst case) is not acceptable because the system "rings" too much when a step load is applied and takes a long time to reach its steady state.

Further, the numbers we have used for this example have been selected randomly than for sake of simplicity we will neglect the optimization aspect because we are interested to crosscheck linear vs cycle by cycle behavior. At the same time we are well aware an optimization step is required by practical point of view.

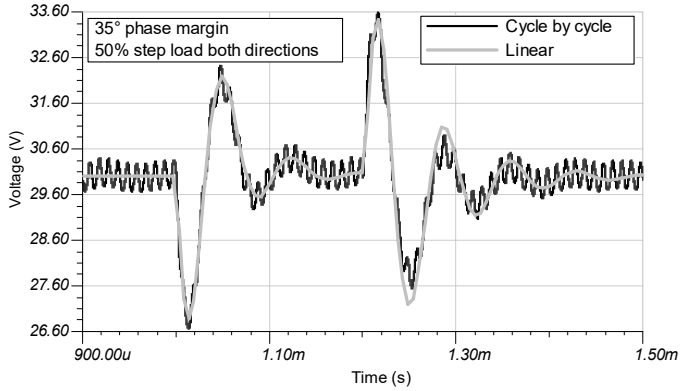


Here, use one circuit to get AC frequency response and linear time domain analysis. Toggle both feedback switches up to get linear transient analysis (loop chain not broken). Toggle both switches down to get AC bode diagrams for plant and plant + compensation (loop chain broken).

Here is the cycle-by-cycle model:



The agreement between them is pretty good as per this overlapped response curve:



Isolated Topologies

Actually, the Vorperian method can be used to get a linear model even for slightly complex topologies such as: Cuk, Zeta, Sepic, Isolated etc. A question now comes spontaneously: What about isolated topologies such as Forward, Flyback, Push Pull etc? Is it possible to use the PWM switch approach? If yes, where is it convenient to place the PWM switch on primary or on secondary? Do main equations differ drastically?

First, it is possible to use Vorperian approach even for isolated (fixed frequency) topologies. Second, it is convenient to place the PWM switch on the primary side simply because it is much easier to simulate a linear model for multi-output converters. If the output is only one, then we can choose arbitrarily where to place the PWM switch: primary or secondary.
Flyback and Forward linear models can be built referring to Figure 3.2 of A to Z, Second Edition, Page 130. Simply by mapping all the secondary side components to the primary side we can cut out the main power transformer getting an equivalent buck (forward case) or an equivalent buck-boost (flyback case). In this way we can get a quick examination of the four main transfer functions by using the same approach used for the three main topologies. In particular, given specific input parameters, this time including primary turns *Np* and secondary turns *Ns* with ***n=Np/Ns***, all we need to do is to map the isolated converter into the equivalent one seen on the primary side. Following equations are valid for both cases: Flyback and Forward. Only the inductor value changes between both topologies.

$$V_{out\ equivalent} = V_{out\ effective} \cdot n;\ C_{eq} = \frac{C_{eff}}{n^2}\ ;\ ESR_{C\ eq} = ESR_{C\ eff} \cdot n^2;\ R_{eq} = R_{eff} \cdot n^2.$$
$$L_{eq\ Buck} = L_{out\ Forward} \cdot n^2 \quad while \quad L_{eq\ Buck\ Boost} = L_{mag\ primary\ Flyback}$$

We can simply use the same Buck (Forward case) or Buck- Boost (Flyback case) assuming the equivalent circuit has been correctly mapped.

Function	***Forward (Buck derived) $n = N_P/N_S$***	***Flyback (Buck-Boost der $n = N_P/N_S$***
Control To Output	$G(s)_{Forward} = \frac{1}{n} \cdot G(s)_{Buck}$	$G(s)_{Flyback} = \frac{1}{n} \cdot G(s)$ *with positive si*
Line To Output	$G(s)_{Forward} = \frac{1}{n} \cdot G(s)_{Buck}$	$G(s)_{Flyback} = \frac{1}{n} \cdot G(s)$ *with positive si*
Input Impedance	$Z_{in}(s) = Z_{in}(s)_{Buck}$ ω_{RC}, ω_0, Q don't change same as Table 1	$Z_{in}(s) = Z_{in}(s)_{Buck}$ L, ω_{RC}, ω_0, Q don't same as Table

Table 4.2: Collecting four main transfer functions for Flyback and Forward

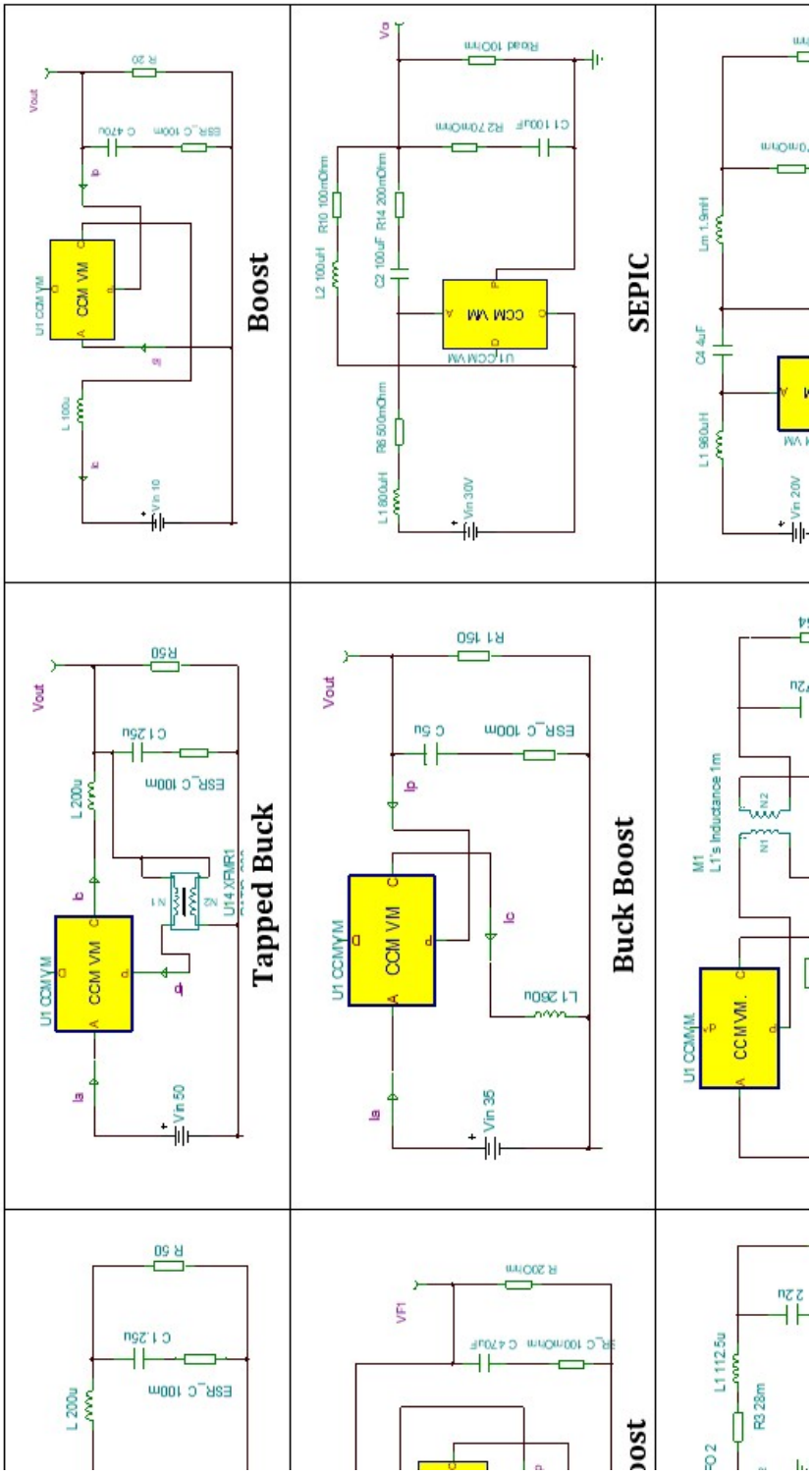


Figure 3.13: Circuit Collection using PWM Switch. The structure is applicable to both cases (Voltage mode and Current Mode, CCM assumed)

Chapter 4
Analog Control Loop Theory

Introduction

Scouring the enticing landscape of control loop theory for enlightenment, we can often stumble and lose our way from seemingly contradictory statements such as: "high gain *reduces* the effect of disturbances; therefore, please *increase* the gain". But in the very same breath: "high gain causes *oscillations*; so *reduce* the gain".

Bewildering!?

Here's another example: "high gain improves load regulation; therefore, you need to *increase* the gain". However, "*voltage positioning* improves load regulation, and for that reason, you need to *reduce* the gain". "Oh, by the way, don't forget to set the phase margin to *greater than 45*°. Worst-case 50°." "Remember, phase margin *really* just needs to be greater than *zero* to avoid oscillations."

Try this one: **https://www.powersystemsdesign.com/articles/interpreting-loop-gain-measurements/18/7041**: "As with voltage-mode control, we want to have a phase margin in excess of 50 degrees. The plot is continued well past crossover to make sure there is sufficient gain margin (what the gain is when the phase margin hits zero.)"

You ask helplessly: "Why 45°, or why 50°? Make up your mind please!"

"In fact, *smaller* the phase margin, snappier the response. So, reduce the phase margin. "Ouch… so why not just 5°?"

"Oh, by the way, *that* what you see on the scope screen has very little to do with loop response, because the calculated loop response is implicitly based on *small*-signal analysis. What you are seeing here is in reality a large-signal response. So, please recalculate your reactive power components. The L and C, dummy! Oh, and don't forget to lower ESL and ESR too".

You ask: "By how much?" "How large is large, how small is small?"

The steady-state or closed loop response/answer to that seems to always be: "Mmmm….". (Take it or leave it).

And just when you think you may have at least got the gist of it all, after struggling through some good University of Colorado books over a pint or two, you have someone advocating "75° or 76° phase margin now— said to be consistent with a "Q of 0.5". **https://www.powerelectronics.com/technologies/power-electronics-systems/article/21853236/transient-response-counts-when-choosing-phase-margin**: "An analytical derivation of the optimum converter phase margin for critically damped response shows it is close to 76 degrees, well above the traditional recommendation of 45 degrees. –"

Incidentally, where did Q suddenly come into the picture, by the way? It already seems futile to ask: what exactly is the relationship between "Q" (quality factor) and phase margin? Not to mention the relationship between the actual measured overshoot to the "Q", or to the phase margin. Is the prevailing, often widely differing industry guidance regarding optimum phase margin, or Q as you prefer to call it now, just based on simulations? OK, maybe it was based on lab results after all, but were they under unknown, unrelated, unstated, uncontrolled, perhaps widely varying conditions? Do they even apply to our case?

We can also ask: "how exactly does this optimum phase angle recommendation change in going from voltage-mode control (VMC) to current mode control (CMC)? How does it depend on the selected L and C_{OUT} of the switcher? Or on its mode of operation: continuous conduction mode (CCM) and discontinuous conduction mode (DCM)? And what is the overall *topology*-dependency? In other words, what happens if we are dealing with a boost or buck-boost, instead of a buck? Is the 'recommended' phase margin always 45°? Or is it 76°? And why?"

And suddenly you also get to hear new terms, barely ever explained, such as "conditional stability." "On what condition?" you may ask. Do not assume though!
Well, clearly the experts seem to be saying: the gain of the closed-loop system is called the "closed-loop gain". *But it is the "open-loop gain" which we are really interested in, for that is what predicts the stability of a closed-loop system*. So, we do a Bode plot measurement—which incidentally, is a plot of the *open-loop gain*. Yes—taken on a system while its loop is literally held closed, not open."

You ask: "So why is it known as the 'open-loop gain'? Why is the open-loop gain so critical to a closed loop system?"

It is endless! How are we ever going to get to the bottom of the seeming quagmire called control loop theory?

That's the rather confusing landscape we all encounter as we take our first baby steps into the mystical world of control loop theory.

The truth is: *control loop theory is a challenging subject indeed*. Even seasoned hardware engineers, so far accustomed to dealing mainly with tangible objects on a lab bench, are expected to grab a paper and pencil and start playing brainy physicist instead! They must hit the ground (plane) running, literally, leaping effortlessly through hoops of imaginary p-planes, s-planes and z-planes, and switching seamlessly between time and frequency domains. If that were not enough, they also need to embrace the fact that the frequencies involved now can not only be negative but imaginary—whatever *that* means! All this can cause the last vestiges of any old-school physical intuition to implode on itself.

It only gets steadily worse—especially when we attempt to apply generic control loop theory to switchers without fully recognizing the fact that many of the concepts we struggled to absorb from all the excellent articles and papers on the subject, need serious reevaluation now. One reason for that is switchers are discrete/digital devices, not continuous as we may have assumed. That is because they have a *discrete* "control effort" *update interval*, related to the discrete pulses coming in at the rate of the switching frequency. The net result, expressed intuitively, is that "error" information is not necessarily sampled and communicated *instantaneously* to be acted upon and corrected—*it is inherently delayed*. That has major consequences. One of which is: **we need to lower the bandwidth of the control loop *to typically less than one-fifth the switching frequency*.** We will try to provide the reasons why, a bit later. But the truth is that if the bandwidth is set higher, we are in grave danger of over-reacting, based on wrong/delayed input information. Of course, were it not for all such reasons, we perhaps would have logically preferred to set the bandwidth *as high as possible*. In principle, or at least seemingly so, we would then attain "instant/ideal correction." But that is out of reach in reality, though incidentally, it is something close to what we get with hysteretic control, often called "bang-bang regulators." More on that later. First, regular closed loop switchers….

Air-conditioners to Closed-loop Switchers?

Traditionally, a control loop system is often explained with reference to a mundane room air-conditioning system. In that example, the setting of the thermostat is the "set point", acting as the input (*to the control loop system*)—referred to as the "IN" node quite obviously. The output of this closed-loop system, the result, is called the "OUT" node. Here it is the temperature of the room. To have OUT approach IN, a thermocouple, or sensor, is obviously present somewhere, to monitor the room temperature OUT, and thus converge to the desired set point IN. An error stage looks at the "error", i.e. the *difference between the set point and the output*. The system incorporates negative feedback as a means of correction, so if the room temperature goes above the set point (room too "hot"), *cold* air gets pumped into the room to try to reduce the error. And so on.

We are further also interested in things like: what is the rate at which cold air gets pumped into the room if the error is say 15°C. What if the error drops to 5°C? Does the *rate* of cold air being pumped fall proportionally too—i.e. by a factor of three? Or does it just keep going at the same rate constantly, and then simply turn off the moment it "thinks" that the error has dropped to zero? In other words, we are asking if the actual "shape" or "profile" of the correction loop is important too. And if so, how?

Keep in mind there could be significant delays involved in the sensing and the resultant response—say for example, based on the non-zero specific heat capacity of the various parts constituting the sensor and blower, not to mention the thermal capacity of objects in the room. If too many delays are present, the room temperature, i.e. the output, could easily undershoot, and go significantly below the set point, before the system even "knows" or realizes it and steps in to correct it. Similarly, the loop may end up overcorrecting too, before the sensor realizes it. Thus, we can get (hopefully diminishing) oscillations around the desired set point before OUT finally converges closer and closer to IN. And in either case, we are interested in minimizing the under- or over-corrections, and thus come up with the best correction profile too.

As indicated, at some stage, the system will probably push in hot air instead of cold, whereas the room was in reality, already too hot at that moment. So, that is how we get an overshoot, or higher discomfort level, though momentarily (hopefully). In addition, in a correctly designed system, there will be a diminishing error over time, and ultimately the room temperature will stabilize very close to the set point. Till the next "disturbance"!

We also must understand that in general, there is, and actually *should be* a small, perhaps even quantifiable "setting error" *at all times*. That residual error between OUT and IN is essentially determined by the "gain profile" of the correction loop. There may be a residual 1 or 2°C error at the end of it all, unless the loop can detect that error progressively, and try to correct it too, perhaps over an extended period of time. But the residual error per se, is essentially mathematical in nature. In other words: *if the system has a high "gain"*, since gain is simply Δ(OUT) divided by Δ(IN), and it is not infinite, there will be a certain Δ(OUT) (or settling error) related to a certain finite, non-zero Δ(IN) (the input "error" or disturbance, which is being corrected by the high-gain loop). The ratio of the two, is based on a finite, *not infinite*, gain. So, for a given Δ(IN), we will be left with a small, but *non-zero* Δ(OUT). For example, if the desired room temperature is 25°C, and the temperature outside is 30°C, the "error" being 5°C, the room may finally settle to 25.2°C. If the outside temperature rises to 35°C, the room may settle down at 25.4°C instead. That is intuitively visualized as based on the "DC gain" of the correction loop. Not related to sudden or immediate responses, which are related to the "AC gain". For example, if the temperature outside *suddenly* changes by 10°C, the room temperature may have temporary, but much bigger overshoots or undershoots before it settles down. And that is a reflection of the "AC gain" (profile) of the correction loop.

Of course the ambient temperature outside does not change "suddenly". But the disturbance could mimic that. For example, if someone opens a window or door temporarily, it is a sudden, applied "disturbance", and the system will rush in to correct the resulting error. In that case, we may be interested in knowing what is the *speed of correction* (the AC response in effect)? Alternatively, we can suddenly change the setpoint too. So we can ask: what happens if we turn the knob of the thermostat, say from 25°C to 28°C? How quickly does the room temperature stabilize now? And what is the "settling time" involved. And so on.

In **Figure 4.1**, we present a basic control loop embodiment of this, as applied to a simple buck switcher for starters. With some differences, as explained below!

The IN node commonly used in general control loop theory is now the REFERENCE or REF node in power conversion. It creates the set point against which the output is compared. The closed loop gain of the system is now Δ(OUT) divided by Δ(REF). The OUT node in general control theory is V_{OUT} in power conversion. Despite some fairly common misinterpretation, the "IN" (of the control loop) is *not* the input to the power stage, or V_{IN}. It is the input to the control system, i.e. REF. Also, once the switcher is on, we don't really wiggle the thermostat/reference around! So the relevance of the oft-quoted closed-loop gain, as visualized and presented in textbook control loop theory, hardly has any significance to switchers. In switchers, the primary "inputs" we are interested in are essentially disturbances, injected at varying points within the closed loop system. Such as line and load variations. So we need to understand at a more general level, how disturbances are attenuated (hopefully not amplified), depending on their point of injection. It is not the same thing as wiggling a thermostat!

The "plant" or "process" in general control theory is now the entire block consisting of three cascaded stages: the PWM comparator, the switching stage, followed by the LC filter. However, the "power stage" of the switcher by definition, traditionally includes only the latter two blocks. So it is the plant *less* the comparator. The comparator, though part of the plant, is considered part of the control section of the switcher since it contains no power components, just signal-level components.

The compensator in general control theory is typically an error amplifier, with all its feedback components present (usually a bunch of small-signal R's and C's). The "sensor" in general control theory (for example the thermocouple in the usual thermostat control loop theory example), is typically the voltage divider in power conversion. But more on its actual effect a little later.

The CONTROL terminal is the same in both representations. But in general control theory it may actually be called the "control effort", whereas in power conversion it is usually called the "control voltage" or "EA OUT" (error amplifier output).

Looking closely at **Figure 4.1**, we see that the entire process of control hinges on the concept of *negative feedback*. So if the output is going up, we try to quickly pull it down! That is why we see different signs around the summation block in the figure.

Note that in related literature, the summation block is often confusingly represented by a *multiplication* sign in the middle of a circle instead of a summation sign. That it is likely done only to make you give up control loop theory altogether? Well!

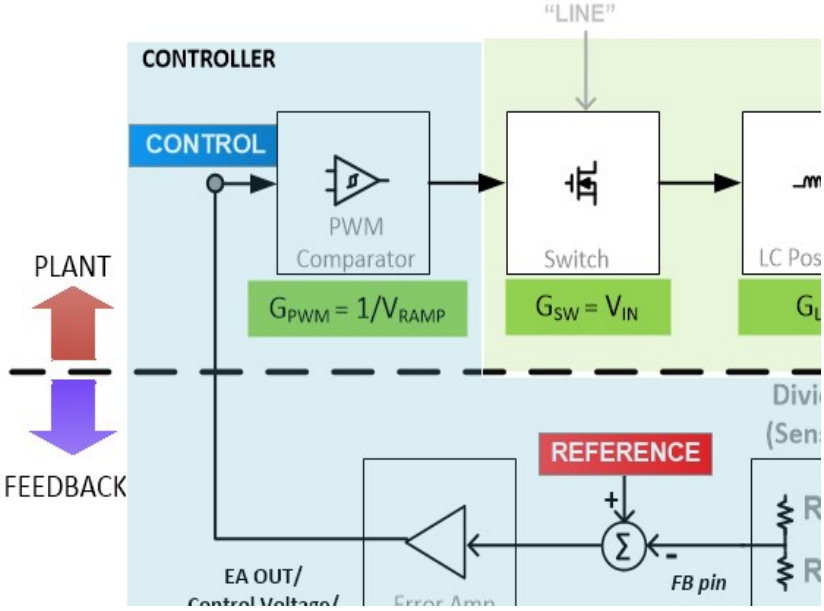


Figure 4.1: Control loop of a switcher

Note that we will often use gain symbol "H" for parts of the feedback section, and "G" for the plant. But in literature it is sometimes *the other way around,* with H being the plant and G the compensator. Beware! Sometimes G is used for all the blocks, in the plant and the compensator. Sometimes K is used, as in older Unitrode App Notes. Or "A" is used for the plant and β for the compensator. And so on. Watch out for a lot of possible confusion as a result of all these terminological variations.

The most important thing to keep in mind is that the representation of **Figure 4.1** assumes we have a multiplicity of *cascaded* gain stages. Which implies the gain of each stage can be quantified as a standalone, and the net gain is then the product of all the individual cascaded gain blocks. But that may turn out to be a pipe dream. For example, the buck topology is the only one where we can actually point to a separate "LC post filter" within the plant. In a boost or buck-boost, even ignoring the relative locations of the L or C, the LC stage is really not "separable" from the rest—because the node between the L and C_{OUT} is connected to the switch/diode—unlike a buck. So we cannot separate the filter from the switch function. Well, at least not easily.

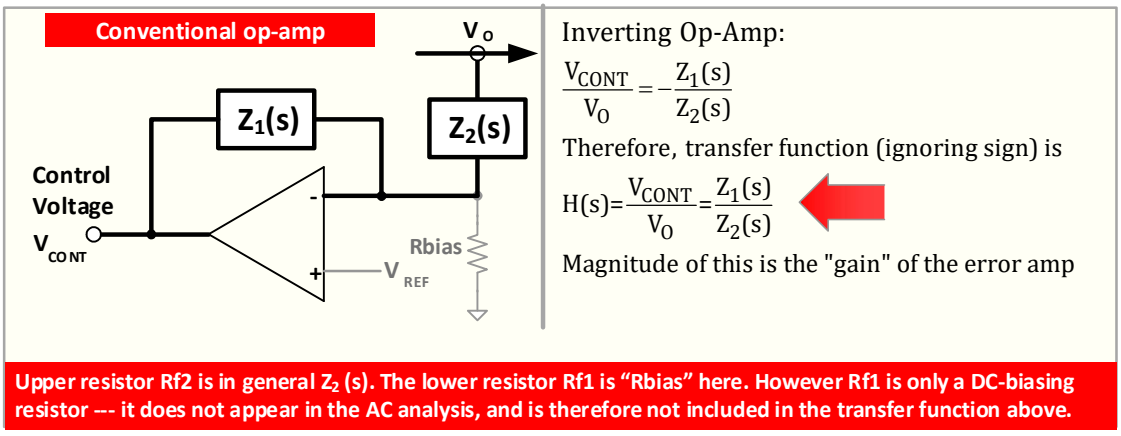
As pointed out, in the canonical model from Middlebrook, we can indeed separate the boost and buck-boost L and C into a separate LC stage, *provided we replace the inductor by an "equivalent inductor"* $\underline{L}$, equal to

$$\underline{L} = \frac{L}{(1-D)^2}$$

The C of this separated LC stage is still the original C_{OUT}.

The voltage divider is also not necessarily separable into a separate gain block. In **Figure 4.2** we show how in a typical error amplifier stage, the lower resistor of the divider goes out of picture. It is just a DC biasing element, not connected with the AC response, which is what we are interested in. In other words, the divider doesn't enter the picture at all from a control loop a perspective. However, if we use a transconductance op-amp, the divider does enter the picture as separable gain block.

This is AC (change) analysis. Therefore V_{REF} is being ignored below as it is a biasing level only
By definition, transfer function ("H(s)") is output/input = V_{CONT}/V_O



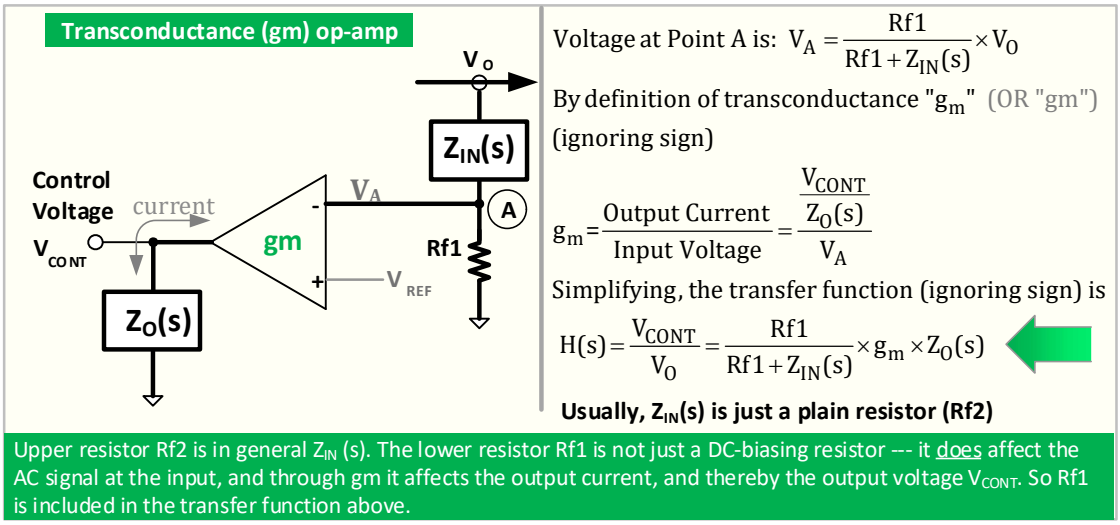


Figure 4.2: Is the voltage divider separable as a distinct gain stage?

The Unitrode Seminar of 1996

To experienced power engineers, some of the nuances mentioned above may seem to be subtle as flying quarter-bricks. It may genuinely surprise them to learn that these "details" are still often missed, or are at least routinely glossed over. But there are some other experts who are not at all surprised. They knew this was going to happen and tried to warn budding power supply engineers since a very long time ago.

The unpolitically correct, or politically incorrect, gist of what some of them have said is: There are many self-proclaimed "experts" who *don't get it*. Listen to the flamboyant Lloyd Dixon of Unitrode (now Texas Instruments). That's what he said, verbatim, while presenting his "Control Loop Cookbook" paper at the Unitrode Power Seminar in Germany. The year was 1996, and this author had the privilege to attend the presentations. Later that day, Bob Mammano, now considered the "father of the PWM IC industry", presented his topic: "Fueling the Megaprocessors - Empowering Dynamic Energy Management".

A short extract from Mr. Dixon's no-holds barred *written*, therefore more "PC" (politically correct), part of the presentation is reproduced below (see **http://encon.fke.utm.my/nikd/Dc_dc_converter/TI-SEM/slup113.pdf**)

"Control Problems your Mother Never Told You About"

That was the actual title of Lloyd Dixon's presentation that day! What he said was this: "A tremendous amount of effort has been put into the development of small-signal techniques and linear models of the various switching power supply topologies. Hundreds, if not thousands of papers have been written over the years. Your academic "mother", whoever "he" may be (note the PC sexual ambiguity), typically focuses on new topologies and/or linear modeling. While not disparaging any of these efforts – far from it, these contributions have been immense and totally necessary – there has been a lack of balance and a tendency to try to force behavior that is uniquely related to switching phenomena into linear equivalent models (with sometimes uncertain results). Many of the major significant problems with switching power supplies do not show up in the frequency domain, or in the time domain using averaged models, unless these problems are anticipated in advance and provided for in the models. Simulation in the time domain using switched models, although slower, reveals these problems that would have been hidden."

Perhaps somewhere along the way, a few relatively inexperienced engineers have gotten a bit carried away with their new-found prowess manipulating Laplace transforms and so on, and have therefore ended up downplaying, if not completely disregarding, some of these "subtle" aspects. Or perhaps their simulations failed to reveal what they perceived to be corner-case" problems—since they had used small-signal averaged or equivalent linear models for the switcher to start with, which then turned out to be a self-fulfilling prophecy: You can't see in the dark, if you didn't realize it may be dark and forget to bring along a flashlight.

Applying Control Loops to Real-world Switchers

The first thing we have to keep in mind is that control loop theory can be applied to a switcher *only when its power stage is considered reasonably optimal*. That should not become the stumbling block. Otherwise, we would be just wasting all our valiant efforts in the frequency domain.

As an example, see **Figure 4.3.** Here we are showing a sample transient waveform from a vendor. See **http://go.intersil.com/lp-stable-power-supply.html**.

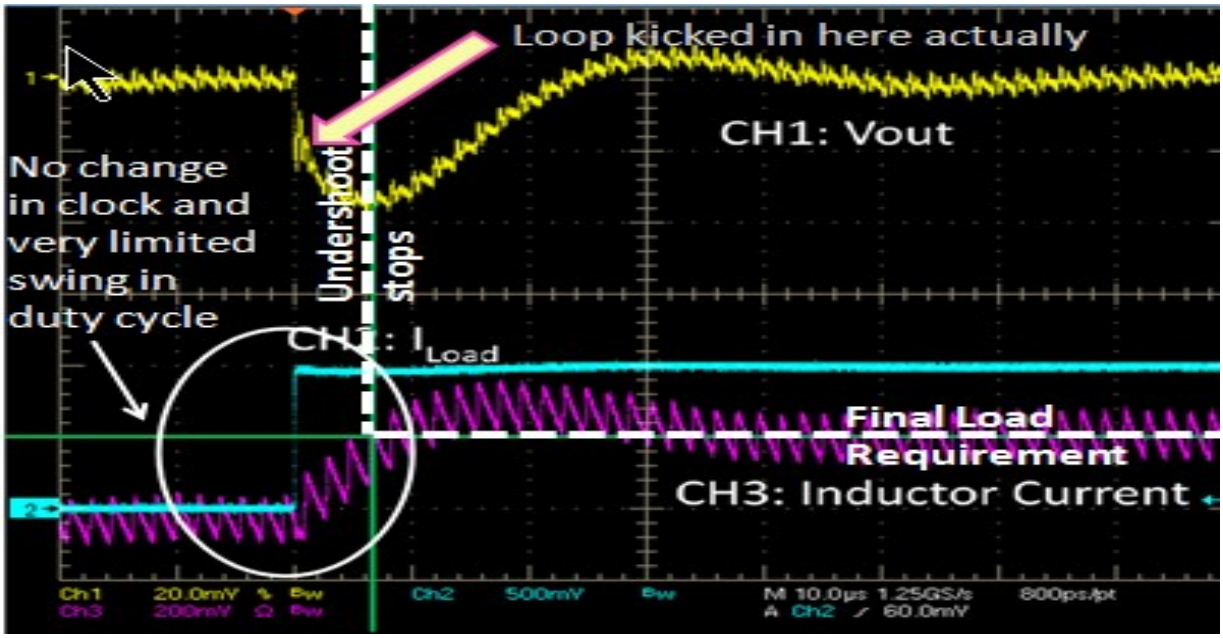


Figure 4.3: Typical vendor plot showing load transient and inductor current (Intersil)

This vendor has helpfully provided the corresponding inductor current waveform too, which is quite unusual. It is quite revealing if you stare at it a bit.

The first interesting thing in **Figure 4.3** is: *the undershoot reverses direction exactly at the point where the inductor current reaches its final* intended *value.* So that seems to be the real gating item. *Not* the control loop. Of course, after that point, being a voltage-mode control system, the inductor current does overshoot a bit, commensurate with the observed output voltage overshoot. In a sense, the inductor current has a certain "momentum". But after a little ringing, the system stabilizes.

However, based merely on the fact that the *undershoot stops exactly at the point where the inductor current reaches its final* intended *value,* the overall response (undershoot) was *power-stage dominated,* not *control-loop dominated*. Which implies that anything more we could do with the control loop may have fallen flat on its face.

In fact the control loop appears to have reacted long before the maximum undershoot (minimum voltage) was recorded. See **Figure 4.3** again. That is actually typical of all well-designed control loops—they all "kick in" after about *three* or more switching cycles.
We see from the same figure that after the control loop kicks in, the control loop was certainly trying to command a correction, but oddly, nothing significant or dramatic happened—at least not immediately. Why? One reason for that could be some sort of an architectural limitation. Indeed, voltage-mode control (VMC) has some inherent deficiencies, based on the "momentum" of its inductor current. But despite that, VMC is nowadays considered a better choice, especially with *input feedforward* included (to be discussed shortly). That is in comparison to current-mode control, with its now-perceived inherent deficiencies such as subharmonic instability, noise sensitivity etc. That seems to be receding in the distance.
Note that hysteretic controllers seem quite promising in this regard, but they have a variable switching frequency, especially during transients. So, we may need to validate their system-level acceptability.

What did we mean about architectural limitations? Well, to optimize any voltage-mode controlled buck switcher, we need to, at a bare minimum, demand that it can get quite close to 100% duty cycle—to rapidly build up the inductor current as shown in **Figure 4.4.**

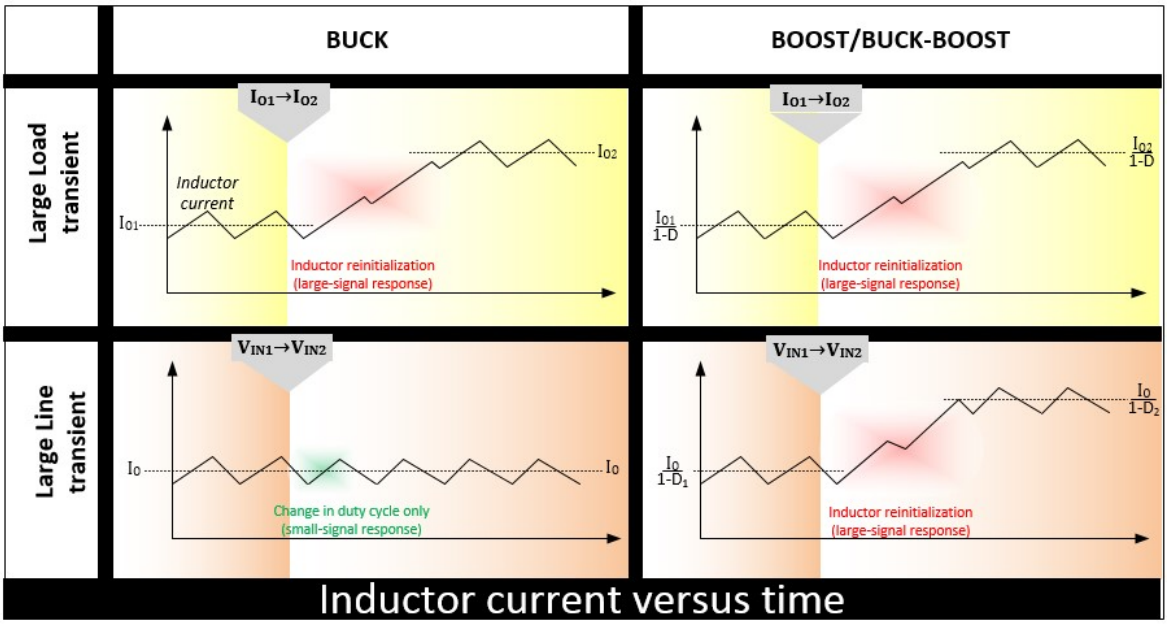


Figure 4.4: Inductor "Re-initialization" problem

Keep in mind that this ~100% duty cycle maximum limit is a very bad idea for a buck-boost or a boost, which actually *depend on the non-zero OFF-time to deliver energy to the output*! If we don't provide that, the output will continue to sag, whatever the current buildup in the inductor. And that is also the intuitive explanation of the oft-mentioned "right half plane (RHP) zero". The RHP zero doesn't exist for buck or buck-derived topologies such as the forward converter, but does enter the picture with the boost and with the buck-boost, including their derivative topologies such as the flyback.

Another thing we can do is change the clock to respond faster to a transient (pulse-on-demand), and if the varying clock frequency is acceptable, it should be fine. That's what happens in a hysteretic controller.

Unfortunately, from **Figure 4.3** we can see that there is no sign, either of on-demand pulses (off-cycle switch turn-ON), or of maximum duty cycle. Which could partially explain the slow curving change in output voltage, despite the control loop having kicked in.

Or as indicated, perhaps the inductance was simply too large.

However, it may also be that the power stage was well-designed and the duty cycle may indeed have been able to "max out" close to 100%, but it *simply wasn't asked* to, at least not quickly enough! Now, that would imply a poorly designed feedback stage, say with poor bandwidth! But we see enough signs that in this case the control loop did kick in after about three switching cycles. Which is about right! So, poor bandwidth doesn't seem to be the culprit here. A bunch of other things to check out instead.

The type of output response we would *like to see* is shown in the lower half of **Figure 4.15**. We don't need to see the inductor current to make some observations. Notice the sharp edge in the output voltage under/overshoot as compared to the waveform in the upper half of the figure. That seems to indicate a *control-loop dominated response*, not power-stage limited. But also keep in mind that to declare victory, this particular sharp-edged output waveform must correspond to a *large*-signal event, such as zero to max load. We certainly cannot pass judgment on the power stage, whether it is "optimal" or "non-optimal", if we are only doing say, an 80% to 100% load test. Because then the small-signal/averaged models Mr. Dixon warned us against, do apply. In effect, we are then no longer dealing with real-world switchers. Just textbooks.

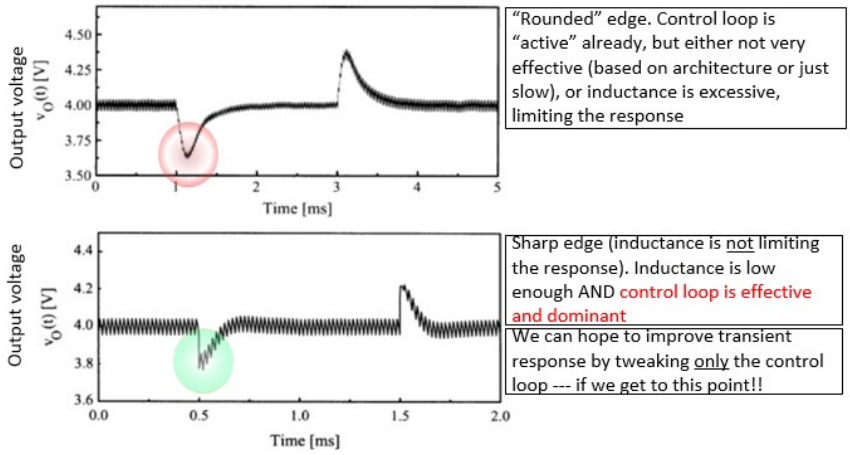


Figure 4.5: The edge of the output response reveals a lot

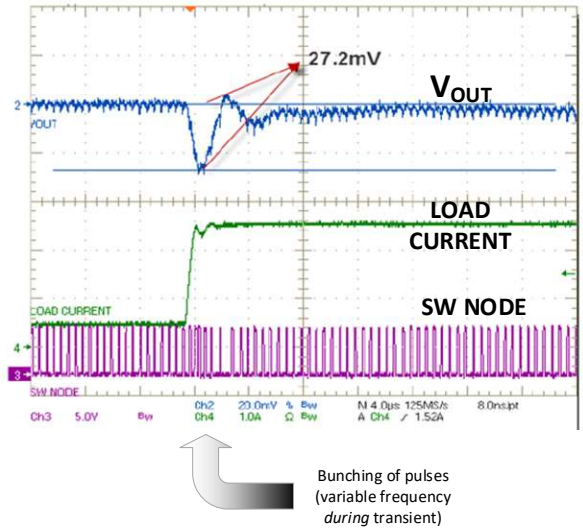


Figure 4.6: Hysteretic controller response

In **Figure 4.6**, we show the exact response of a typical hysteretic controller. Notice the on-demand pulses here, and also the sharp edge as the undershoot starts to head upwards. So, one thing seems quite certain. To get the control loop to make its presence felt during large signal events, we need to optimize the power stage first. One of those steps is to lower the inductance. However, placing the current ripple ratio "*r*" to be close to 0.4 as advocated by this author, is definitely in the ball-park, and there is likely no point reducing the inductance any further. Look for AN-1197 and AN-1246 on the web, originally written by this author in 2001 at National Semiconductor.

What about the output capacitor?

We now hark back to the solved example in Chapter 19 of *Switching Power Supplies A to Z, Second Edition,* to reveal the importance of selecting this vital component too. The results are presented in **Figure 4.7**. Basically, there are three main criteria for capacitor selection. One is the ripple, measured under steady-state max load. It leads to a minimum capacitance requirement of 5.2 µF in this particular example. Another is based on the overshoot which will occur if we just suddenly disconnect the load. This leads to a minimum capacitance requirement of 22 µF. And third, we have a certain *control loop assumption*, which says the output cap must be capable of providing *all* the energy for at least *three* cycles, in case of a large load step. Because during these three cycles, in effect the inductor is not capable of providing most of the energy requirement, as its current slews up in accordance with **Figure 4.4** (discussed in more detail later). This leads to a minimum capacitance requirement of 30 µF in our example. We picked 33 µF as a potential final value. We should however consider tempcos (temperature coefficients) again, and voltage coefficients too.

Note that if in the second calculation, instead of a 2.2 µH inductor, we had used a 4.7 µH inductor, it would have required a minimum capacitance of almost 50 µF, which would have overshadowed the control loop dictated minimum capacitance of 30 µF. *So, we have to be very careful not to select a larger inductance than recommended.* But r = 0.4 should work fine usually.

On the other hand, if the control loop design is sluggish, requiring not 3, but say 6 switching cycles to start acting, we would need to correspondingly increase the minimum capacitance from 30 to 60 µF. That adds to the cost too.

So, hand-in-hand with power stage optimization, we need to optimize the control loop too for best results.

And finally, besides juggling a few things around before deciding the optimal power component selection, let's not forget the basic architecture of the controller either.

Minimum Capacitance and Maximum ESR based on maximum output ripple

Ignoring ESR and ESL, purely based on capacitance, the maximum allowed output ripple determines a minimum output capacitance.

$$C_O \geq \frac{r \times I_O}{8 \times f \times V_{RIPPLE\_MAX}}$$

Including ESR, but assuming C is large and ESL is negligible. The maximum allowed voltage ripple determines a maximum ESR

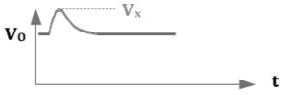
$$V_{RIPPLE} = ESR \times I_O \times r$$

$$ESR \leq \frac{V_{RIPPLE\_MAX}}{I_O \times r}$$

Minimum Capacitance based on maximum overshoot

There is another criterion. In case of a sudden release of load demand, say from max load Io to zero, the inductor energy will all get dumped into the output cap. If we do not want too much of an overshoot (to a new value Vx):

$$\frac{1}{2} \times C\left(V_X^2 - V_O^2\right) = \frac{1}{2} \times L\left(I_O^2\right) \quad \Rightarrow \quad C \geq \frac{L\left(I_O^2\right)}{(V_X + V_O) \times (V_X - V_O)} \approx \frac{L\left(I_O^2\right)}{(2V_O) \times (\Delta V_{overshoot})}$$



$$C_O \geq \frac{L\left(I_O^2\right)}{(2V_O) \times (\Delta V_{overshoot})}$$

(where we have used the approximation Vx + Vo ≈ 2 x Vo. Also, Vx - Vo = ΔV)

Minimum Capacitance based on maximum droop

Typically, with a well-designed loop, it takes about **three switching cycles for the loop to react** and start correcting the output to meet a sudden load demand. During that time we do not want the output cap to fall more than a certain value Vdroop. Thus, using I = C dV/dt, we get

$$I = C\frac{\Delta V}{\Delta t} \quad \Rightarrow \quad C \geq \frac{I \times \Delta t}{\Delta V} = \frac{I \times 3T}{\Delta V_{droop}} = \frac{I \times 3}{\Delta V_{droop} \times f}$$

Here the droop is actually related to the extra load demand, since the normal load requirement is being met every cycle without any droop. So the current here is actually the load increase.

$$C_O \geq \frac{3 \times \Delta I_O}{\Delta V_{droop} \times f}$$

$$C_{O\_MIN\_1} = \frac{r_{VINMAX} \times I_O}{8 \times f \times V_{O\_RIPPLE\_MAX}} = \frac{0.4147 \times 5}{8 \times 10^6 \times 0.05} = 5.1834 \times 10^{-6}\ \text{F}$$ 5.2 µF #1

$$C_{O\_MIN\_3} = \frac{L \times I_O^2}{2 \times V_O \times \Delta V_{OVERSHOOT}} = \frac{2.2 \times 10^{-6} \times 5^2}{2 \times 5 \times 0.25} = 2.2 \times 10^{-5}$$ 22 µF #2

$$C_{O\_MIN\_2} = \frac{3 \times (I_O/2)}{\Delta V_{DROOP} \times f} = \frac{3 \times (5/2)}{0.25 \times 10^6} = 3 \times 10^{-5}$$ 30 µF #3

We picked 33µF standard value

(but need to watch out for tolerance, temperature, voltage coefficients etc...may need to oversize as much as by a factor of 2)

Note: If Inductance was 3.3uH, not 2.2uH we would get 33 uF not 22uF for condition 2 above. It will therefore start dominating the selection!
Ensure inductance is NOT excessive

Figure 4.7: Worked example for output capacitor selection criteria

Open and Closed Loop Gains

Referring to Lloyd Dixon's presentation again, he says: "*The open-loop gain, T, is defined as the total gain around the entire feedback loop (whether the loop is actually open, for purpose of measurement, or closed, in normal operation).*" So in our terminology, T = GH, where G and H themselves may be the product of cascaded stages as we can see from **Figure 4.1**.

Similarly, Dixon says: "*Closed-loop gain defines the output vs. control input relationship, with the loop closed*".

Actually Mr. Dixon is calling the reference as the "control" node here. Which is a bit misleading. Besides that, the reference in his case is placed between the voltage divider and the error amplifier. Which actually need not be the case always, even assuming the divider can be taken out as a separated gain stage, which in reality may not be so either, as mentioned previously.

Which is why it is very important to truly understand how a disturbance gets attenuated on account of the closed loop control system, compared to the case of no feedback (open control loop). And this also depends on the *point of injection of the disturbance*, in this case the "wiggle" in the reference (which as mentioned is actually of no real significance in a switcher either!).

Just to resolve the widespread confusion, let's see the form of the open-loop gain function and get comfortable with our understanding.

Refer to **Figure 4.8**, where we compare what happens in two cases, depending on the location of the reference wiggle, whose effect on the output is what closed loop gain is all about.

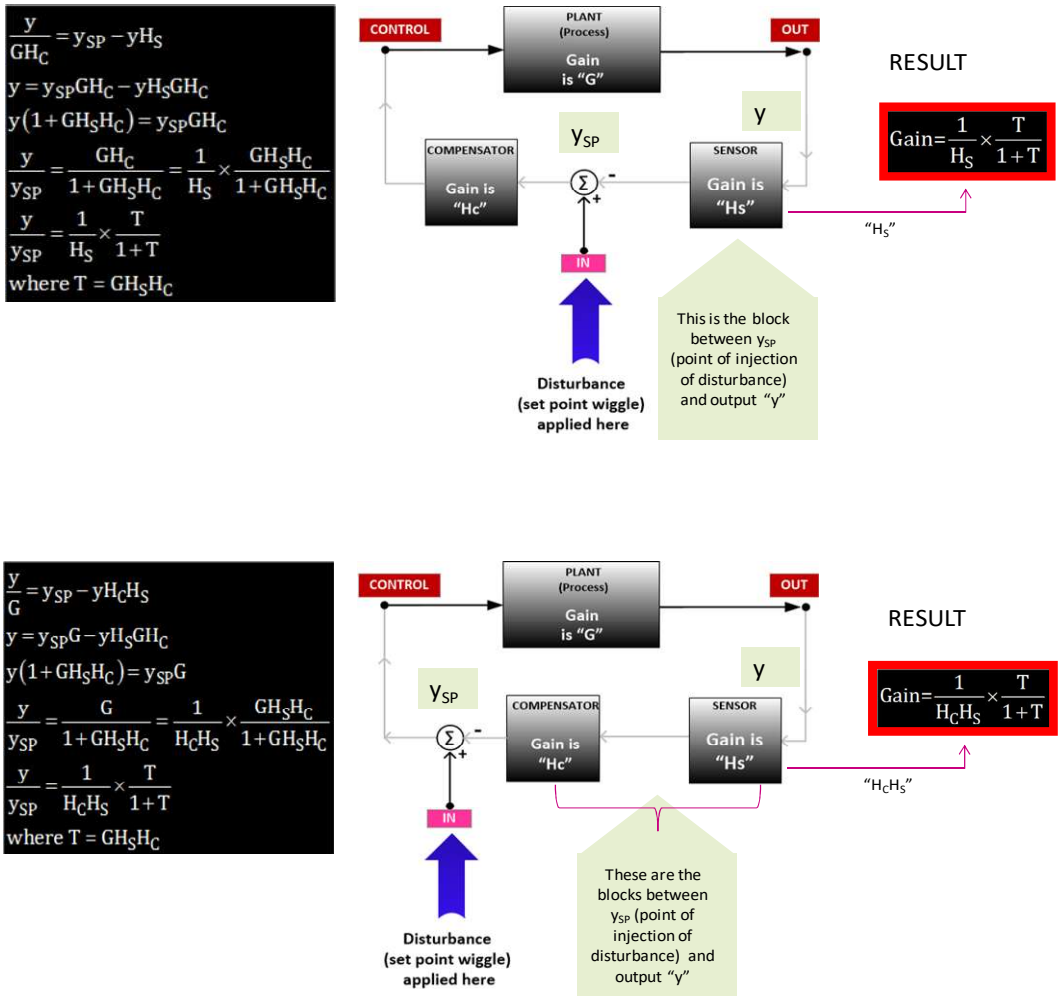


Figure 4.8: Point of injection of disturbance and corresponding closed loop gain

In the first (top) case we first realize that any change in the output, "y", propagates clockwise in the close-loop system. So starting from the output, we retrace its path *backwards* through the plant G and the compensator H_C (anticlockwise), and the signal at the output of the summation block must be $y/(G \times H_C)$. Now, going clockwise from the output rail instead, y becomes $y \times H_S$ after passing through the sensor. After the summation block it is therefore $y_{SP} - yH_S$. But this must equal $y/(G \times H_C)$. Equating the two, we get the expression evaluated within the figure.

Similarly, we can go clockwise and also anticlockwise with the reference re-positioned as shown in the lower half of **Figure 4.8**, and we get the expression evaluated within the figure. We see that in both cases of the figure, we end up with a form

$$\text{Open Loop Gain} = \frac{y}{y_{sp}} = \frac{1}{H_x} \times \frac{T}{1+T}$$

where H_x is the net gain of all the gain blocks between y and y_{sp} in the forward (clockwise) direction.

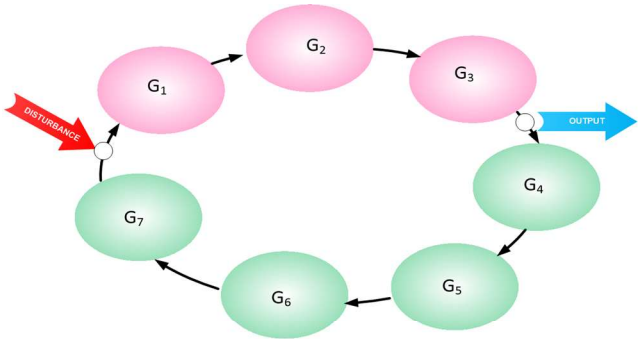
Note that T is simply the net gain of all the blocks involved, be they considered part of the plant (G) or of the feedback network (H). And that leads us to a very general derivation of the closed loop gain of several cascaded stages, where we can just use one symbol for all of them, say G as in **Figure 4.9**.

What this figure is saying is that:

$$\text{Gain with feedback} = \text{Gain with no feedback} \times \frac{1}{1+T}$$

In terms of conventional labels, "*Gain with feedback" is the closed loop gain*. In fact "Gain with no feedback" should actually be considered as the open-loop gain, since that expresses the actual change in the output for a change in REF, with no feedback present. If that is so, *then T should more correctly be simply called the "loop gain"* since it is the cascaded product of all the gain stages of a closed loop system.

1/(1+T) is the "correction factor" which tells us that the effect of closing the loop reduces (hopefully) the effect of the disturbance on the output, by the factor 1/(1+T).



In general, IN can be any input disturbance (line, reference etc.) injected from outside into the closed loop of the plant and feedback.

Its effect on the output OUT is:

$$\left.\frac{OUT}{IN}\right|_{no-feedback} = G_1G_2G_3$$

Gain with no feedback

$$\left.\frac{OUT}{IN}\right|_{with-feedback} = \left(\frac{1}{G_X}\right)\times\left(\frac{T}{1+T}\right) = \left(\frac{1}{G_4G_5G_6G_7}\right)\times\left(\frac{G_1G_2G_3G_4G_5G_6G_7}{1+T}\right) = G_1G_2G_3\left(\frac{1}{1+T}\right)$$

Where T= "loop gain" (i.e. $T = G_1 \times G_2 \times G_3 \times \ldots \times G_N$)

Correction factor

Figure 4.9: General form of closed loop gain for arbitrary point of injection of disturbance

We can do a more detailed analysis of the correction factor:

> For DC, or rather near-DC, where there is no or little associated phase shift, T is just a real number (no imaginary component). And if it is very large, then $1/(1+T) \approx 1/T$. *So it reduces the effect of the disturbance on the output by the factor 1/T, compared to the case where there was no feedback present.* At least for low frequencies, this is easy to visualize.

For example, if in a buck, the input is 10V and output is 1V, the output is $V_{OUT} = V_{IN} \times D$. So if the input doubles to 20V, so will the output (with no closed loop in place). In terms of dB, the input doubled ($20 \times \log 2 = 20 \times 0.3 = 6$ dB), and the output also went up by 6 dB. However in AC analysis, gain is the ratio of the *change* in output to a *change* in input. The correct way to proceed is shown in **Figure 4.10**.

$$\text{Input\_Output\_Gain}_{\text{with\_no\_feedback}} = 0.1$$

So if input jumps by 10V (to 20V), the output will jump by 10V $\times 0.1 = 1$V (to 2V

$$\text{Input\_Output\_Gain}_{\text{with\_no\_feedback\_dB}} = 20 \times \log(0.1) = -20\text{dB}$$

$$\text{Correction Factor} = \frac{1}{1+T} = \frac{1}{201} = 5 \times 10^{-3}$$

$$\text{Correction Factor}_{\text{dB}} = 20 \times \log(5 \times 10^{-3}) = -46\text{dB}$$

$$\text{Input\_Output\_Gain}_{\text{with\_feedback\_dB}} = \text{Input\_Output\_Gain}_{\text{with\_no\_feedback\_dB}} +$$

$$\text{Input\_Output\_Gain}_{\text{with\_feedback\_dB}} = -46 - 20 = -66\text{dB}$$

$$\text{Input\_Output\_Gain}_{\text{with\_feedback}} = 10^{\text{dB}/20} = 10^{-66/20} = 5 \times 10^{-4} = 0.5\,\text{m}$$

Figure 4.10: Sample calculation of DC settling error

Note that when we express gain in terms of decibels, multiplication (say of cascaded gain stages) is reduced to a summation (in decibels). *Also, here we are referring to the "input" as literally the input rail, not the reference.* We show how the expected 1V shift in the output on account of the input doubling, with no closed loop correction, is reduced to just 5 millivolts on account of the high DC gain of the system.

- At high frequencies, if T equals -1, the denominator will "explode". Which means that we will in effect have sustained oscillations, because there are always limiting parasitics present, which will not allow the output to really rise to "infinity".

Still, we need to understand how T can equal -1. Very simply that means, in terms of magnitudes and phase expressed in polar notation, i.e. in the form of r ∠ θ, we can write T = 1∠-180°. That is a magnitude of 1 and opposite phase.

So we have full-blown instability if the loop gain T equals 1, and the corresponding phase is -180°. Why is that a problem? Because combined with the intrinsic -180° associated with negative feedback (the signs around the summation block where the reference is introduced in **Figure 4.1**, or the fact that the output is fed to the *inverting* pin of the op-amp in **Figure 4.2**), we get a total phase lag of -360°. This means: "in-phase". The disturbance is reinforcing itself.

Criterion of instability and "safety margin"

We thus arrive at the criterion for instability of a closed-loop system: It all depends on T, which as we are now realizing is better referred to as the "loop gain", not really "open-loop gain", which is rather misleading. So at the particular frequency at which T = -1, the disturbance has in effect gone around the closed loop and returned to the point of injection with exactly the same magnitude and phase it started off with. So, it is going to reinforce *and* sustain itself!

The frequency at which ||T|| = 1 (same as 0dB axis) is called the crossover frequency. If at the crossover frequency, the phase of T is exactly -180°, then the system will be unstable!

Note that in the 1996 Unitrode presentation, Mr. Dixon went to great pains to explain why there is a possibility of oscillation *only at the crossover frequency*. Why is it that the signal can actually return in phase with a gain far greater than 1, and the system still be considered stable? Mr. Dixon confessed to spending "sleepless nights" thinking about this, and ultimately explained it as a vector formation that just can't "close"…and thus can't exist, at least not for long.

But it does throw up the possibility of "conditional stability" which people in the industry seem to have widely differing views on. Because one problem is, if during a sudden large load transient, the inductor needs time to build up current, or the error amplifier "rails", then in effect the gain collapses, and it could at some point reach the self-sustaining condition expressed as $T = 1\angle -180°$.

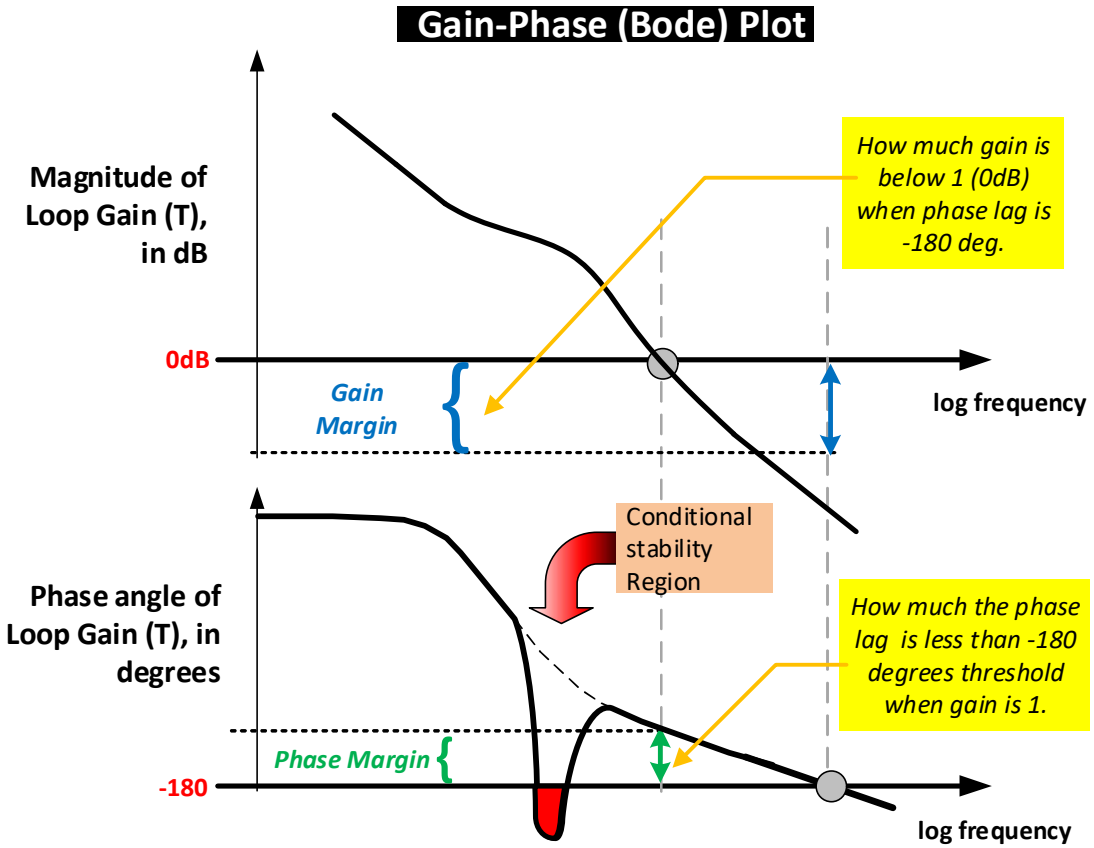
This author will at a later stage describe how in fact this "conditional stability", which incidentally, even Ray Ridley underplays completely, can actually be a major contributor to the "ringing" we see on the output during load transients, *and thus should be carefully weeded out as far as possible*.

Lloyd Dixon also cautioned quite a bit against conditional stability, but more in relation to the gain collapsing for various reasons, and the resulting propensity for sustained oscillations. He didn't relate it to any improved transient response, which we will do later in this book.
In an effort to create a safety margin from instability, the terms "phase margin" and "gain margin" were coined. See **Figure 4.11**. This is called the Bode plot and it tells us everything about "T", the loop gain (previously called open loop gain). That is all we need to know for ensuring stability (Nyquist's criterion, as Mr. Dixon pointed out, is useful, but truly necessary only in cases where there are multiple f_{CROSS}).

The limitations of Bode plots are: they really tell us nothing about what *will* happen in the *time domain*—in terms of amplitude or frequency of overshoots or undershoots as a result of any line or load transients, or even reference-voltage wiggles. That is why the "connection" to any optimum phase margin remains nebulous.

In **Figure 4.11,** we also see what conditional stability can look like, but it is not clear from the Bode plot what all are the unintended consequences of this—whether it should be subdued in some way, or if it truly represents a "rugged" system as Ridley opines here:
http://www.ridleyengineering.com/loop-stability-requirements.html?showall=&start=2

Gain and phase margin are inter-related based on the fact that we typically aim for the loop gain T to *drop at the rate of "-1"*. That is a slope of -20 dB/decade. Note that gain expressed in decibels (dB) is $\text{Gain}_{dB} = 20 \log ||\text{Gain}||$. So this means the gain is falling at the rate of 10× (i.e. 20 dB) for a 10× (i.e. decade) shift in frequency. This simply means we have set the loop gain to be *inversely proportional to the frequency* past some break-point (in this case the break-point is close to 0 Hz). That is the most common and easily-handled profile for T, because it corresponds to a "first order filter" (involving only one reactive component combined with a resistance). Such filters can produce only 90° of phase shift, so that leaves us with a comfortable (stable) phase margin of around 180-90 = 90°. Though with some reactive parasitics added to it, we may end up with a lower phase margin. Or we may deliberately try to lower the phase margin, say by placing a pole at a frequency close to the crossover frequency, and so on. But a second-order filter profile for T would have produced a 180° of phase shift right off the bat, rendering it unusable, because the phase margin would then be zero! So a "-1' slope is what we aim for.



With an inherent 180 degrees phase lag on account of negative feedback, if the the total phase lag reached 360 degrees, with a gain of unity, a sustained oscillation becomes possible

Figure 4.11: Gain and phase margins, along with conditional stability

Phase Margin (Degree)	Gain Margin (dB)	
20	3	Great ringing, absolutely bad value
30	5	Slight ringing, slightly bad value
45	7	Borderline damping, best response time
60	10	Generally appropriate value

Figure 4.12: Relationship between gain and phase margin (typical)

In **Figure 4.12**, we show a typical table of the relationship between phase and gain margin, based on the "-1" slope profile we desire. However, some reactive parasitics could cause the phase to veer upwards as shown in **Figure 4.13**, and so there may be no way, practically speaking, to define a gain margin. In that case, ensuring phase margin should usually suffice.

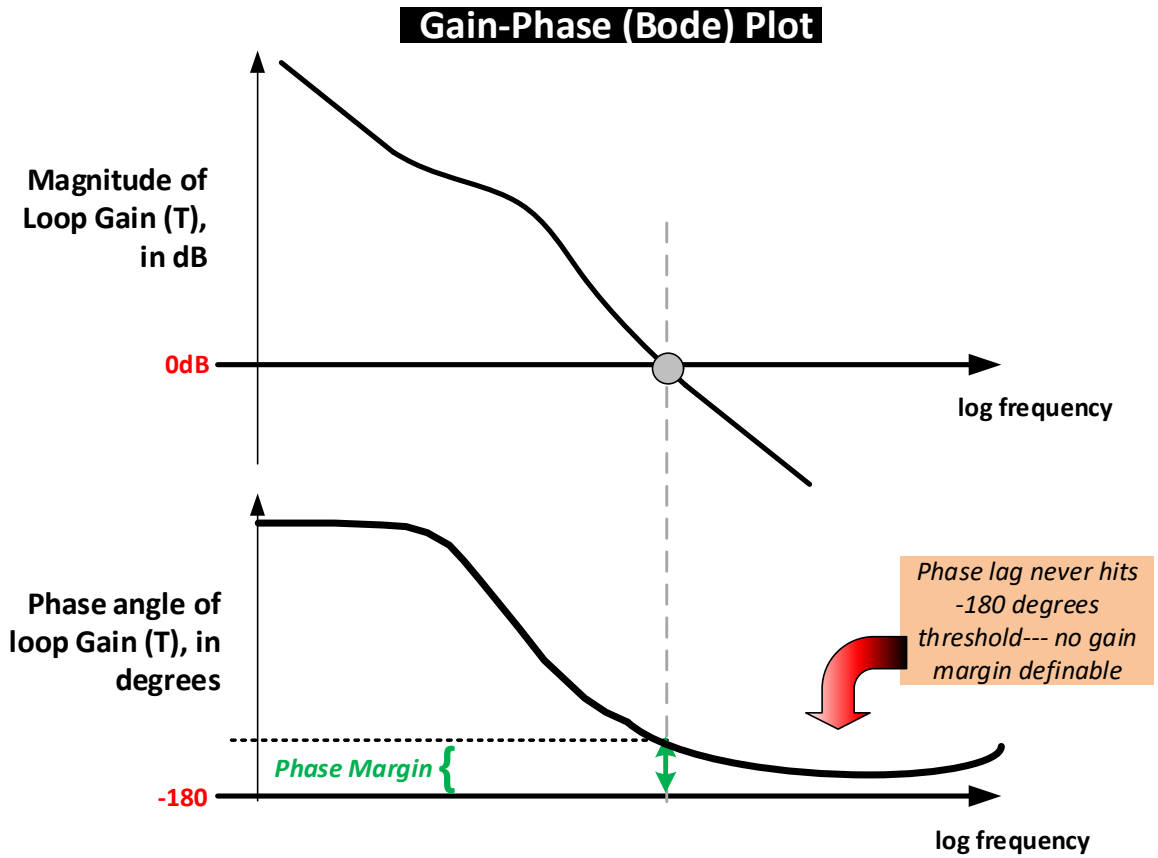


Figure 4.13: Gain margin may not be definable

DC Gain and Settling Error

We typically pick a high DC gain for T and then roll it off at high frequencies to avoid phase-shift based reinforcement of disturbances. That defines the AC response. But it is reassuring to confirm how a high DC gain helps in bringing the DC rail close to its intended set point value (reference).

In **Figure 4.14**, we show how this works, putting some numbers to the test to feel comfortable. It reveals that *a certain settling error is inevitable*, since DC gain is practically never infinite.

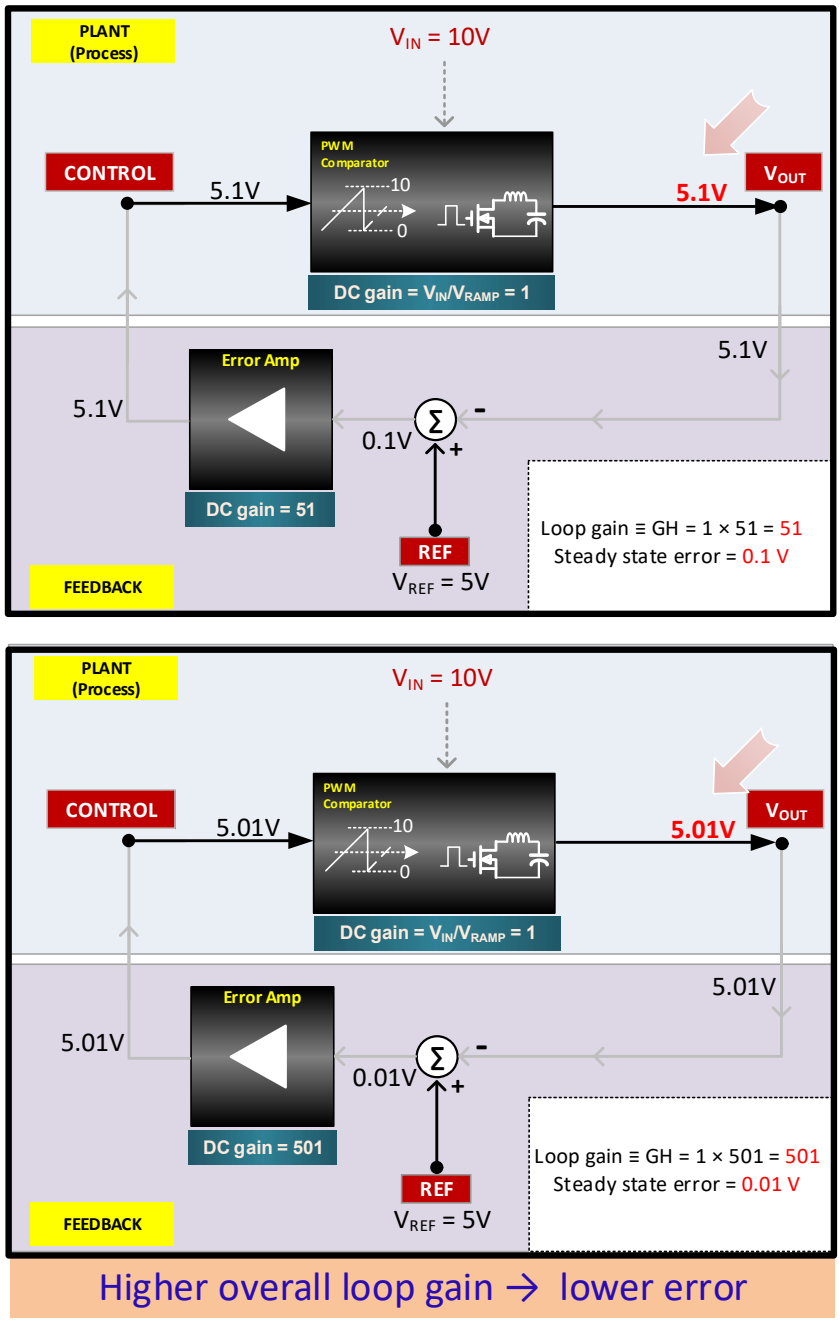


Figure 4.14: DC gain and settling error example

This leads to the AC and DC analysis shown in **Figure 4.15**, which serves to emphasize what we are really trying to eventually do. It is good to keep our key goal in mind at all times, lest we lose our way.

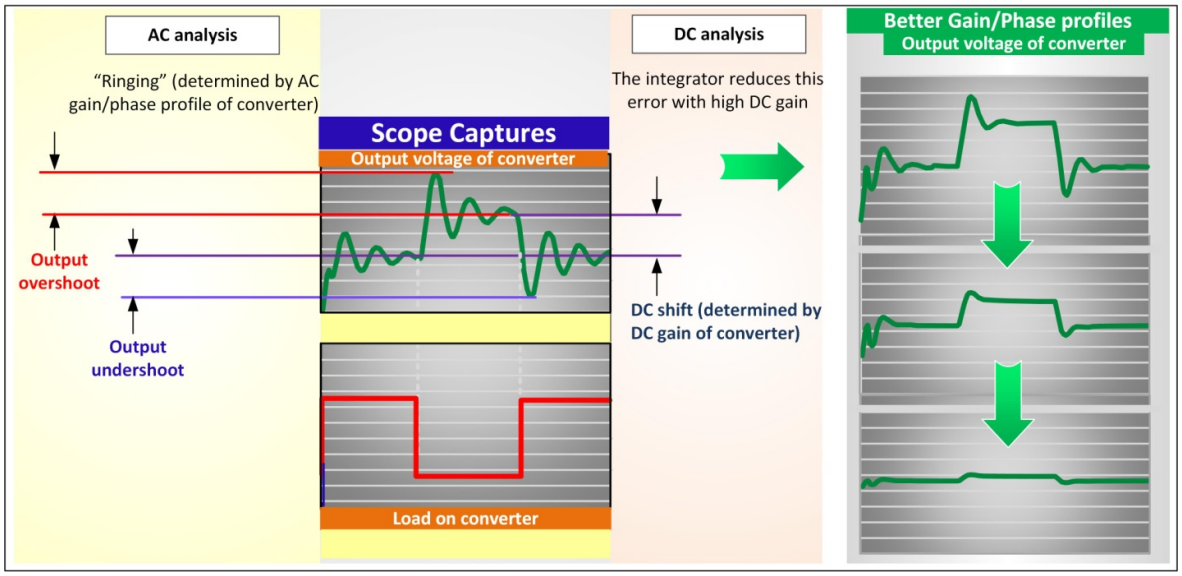


Figure 4.15: AC and DC analysis of a transient waveform and ultimate goal

Voltage Positioning

There is a technique to improve AC response, by trading off some DC gain in the process. Though that leads to a deterioration in the settling accuracy, as we see from **Figure 4.16**, it helps restrict the transient response to within a certain acceptable window.

This technique essentially allows the output rail to collapse a little bit as the load is increased. This can be done by introducing a small resistance after the point of regulation and between the load. That would be called passive voltage positioning. Since this is bound to be slightly dissipative, the modern technique actually varies the set point as a function of load. It is called *active voltage positioning.*

Either way, this positions the output voltage at the lower end of the acceptable window, so now if we suddenly remove the load, the output rail will tend to fly up as expected. *But since it was positioned further down to start with, it now has a larger available overshoot (excursion) before it exceeds the upper threshold of the allowed window.*

Some have expressed the view that this allows the output capacitance to be reduced somewhat. See **www.linear.com/docs/5600**. Indeed, but only if the control loop is the dominant criterion for selection of the output capacitance, as described in **Figure 4.17**.

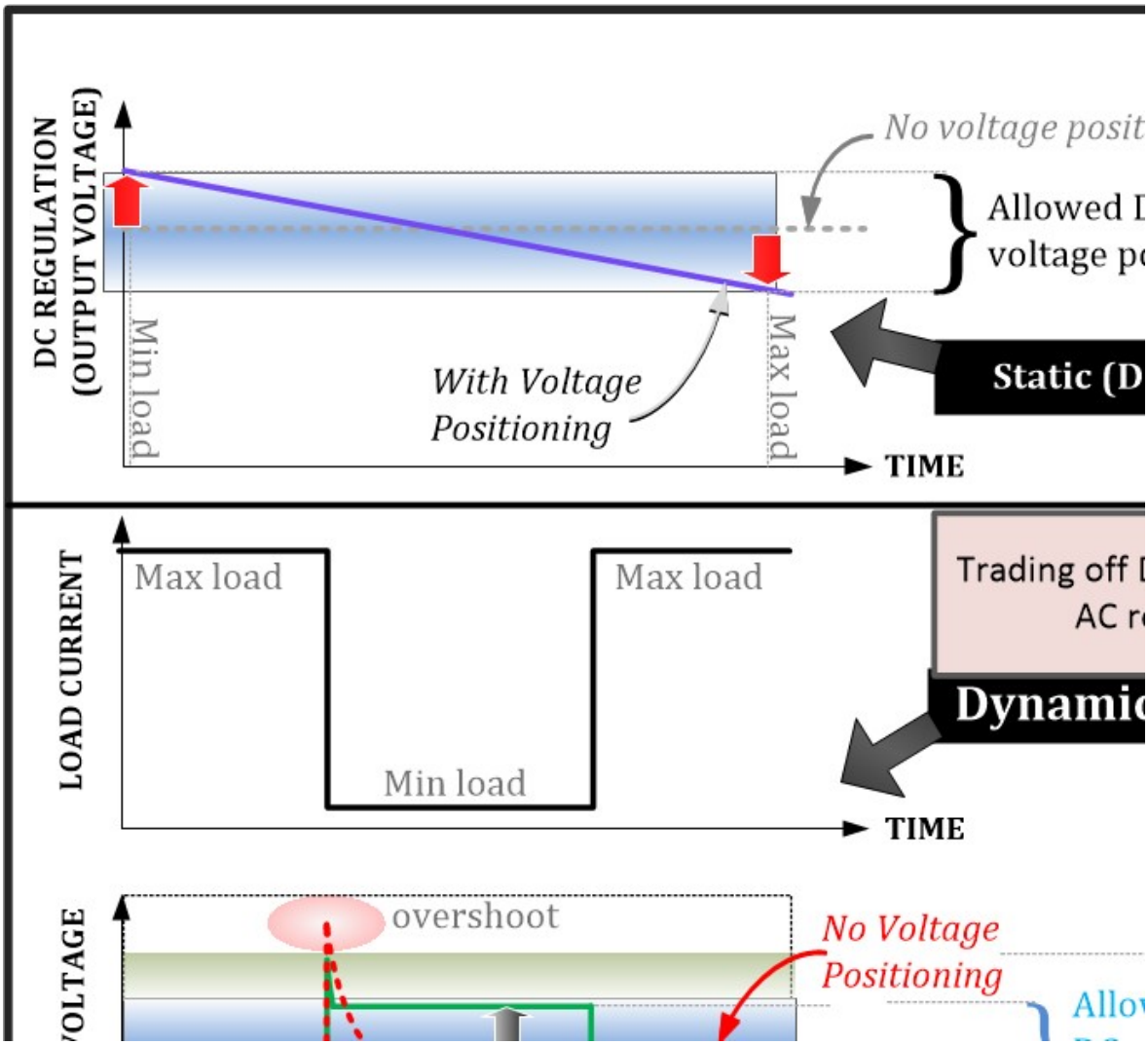


Figure 4.16: Voltage positioning

Input Ripple Rejection

Returning to **Figure 4.10**, we featured a simple numerical example using a regular buck converter to show how a high DC gain helps reduce the effect of input variations on the output. Now we want to extend the same argument to an AC-DC power supply and show how the crossover frequency attenuates the low-frequency input voltage ripple from appearing on the output.

Note that here, "*input" once again refers literally to the input rail, not the reference*. And admittedly, it is better to call that the "line" instead, to avoid confusion.

Let us take the case of an AC-DC power supply with a certain input ripple at 100Hz (full-wave rectified input of 50 Hz). Assuming it is a forward converter with buck-like characteristics, and its duty cycle is 30%, the input-to-output transfer function will provide a dc attenuation of |20xlog(D)| = 10.5 dB, because D is the factor that connects the input rail to the output rail, as in **Figure 4.10**.

But this may receive a further attenuation due to the turns ratio, which may be N_{PRI}:N_{SEC} equal to 20:1. That gives us 20×log (20) = 26 dB. So we have a net attenuation, without feedback, of 10.5 + 26 = 36.5 dB. In terms of actual factors, this is equal to an attenuation of

$$\text{Gain\_attenuation} = 10^{\text{dB}/20} = 10^{36.5/20} = 66.8$$

This means that if we have an input ripple of 10V, the output would have seen a corresponding ripple component of 10/66.8 = 150 mV.

But now let's introduce closed-loop correction. Suppose the entire loop gain (T) is such that it falls at roughly -1 slope, and crosses over at 50 kHz. We ask: what is the loop gain at 100 Hz? —the frequency of our interest here. Since a -1 slope simply indicates inverse proportionality,

$$\frac{\text{Loop\_gain}_{100\text{Hz}}}{\text{Loop\_gain}_{\text{fcross}}} = \frac{\text{f}_{\text{CROSS}}}{100\text{Hz}} \text{ (since -1 slope implies inverse proportionality)}$$

$$\text{Loop\_gain}_{100\text{Hz}} = \frac{50000}{100} = 500 \text{ (since Loop\_gain}_{\text{fcross}} = 1 \text{ by definition)}$$

Expressed in dB, this is

$$20 \times \log\left(\text{Loop\_gain}_{100\text{Hz}}\right) = 20 \times \log(500) = 54 \text{ dB}$$

So since the correction factor is 1/(1+T) ≈ 1/T, this is equivalent to an additional attenuation of 54 dB. So now the net attenuation is 54 + 36.5 = 90.5 dB. In terms of factors

$$\text{Gain\_attenuation} = 10^{\text{dB}/20} = 10^{90.5/20} = 33.5\text{k}$$

This means that if we have an input ripple of 10V, the output will see a corresponding ripple component of 10V/33.5k = 0.3mV.

This is a major improvement over the 150mV without feedback.
Of course, to get the actual output ripple, we have to add the contribution from the output filter stage etc. This is just the additional low frequency modulation that will be superimposed on that high-frequency ripple.

Gain of Entire Plant

The PWM comparator is a key gain block of the closed loop system, as shown in **Figure 4.1**. Its gain (transfer function) has an input which is the control voltage, and an output which is the duty cycle.

The PWM comparator basically superimposes the control voltage against a ramp, and picks a duty cycle based on the intersection, as shown in **Figure 4.17**. Since the control voltage is the "in", and D is the "out" for this gain block, we can see from the figure that the gain is simply $1/V_{RAMP}$. Smaller the ramp, higher the gain. Also, this gain is not a frequency-dependent block. It applies to all frequencies extending up to the switching frequency and beyond. There is no "associated" phase shift either. This is just "DC".

Coming to the switch,

$$V_O = D \times V_{IN} \quad \textit{(buck)}$$

Therefore, differentiating

$$\frac{dV_0}{dD} = V_{IN}$$

So, in very simple terms, the required transfer function of the intermediate "duty-cycle-to-output stage" (i.e. the switch) is equal to V_{IN} for a buck. And it is not frequency-dependent either. Just a DC block.
All the frequency dependent response of the plant comes from its LC post filter.
Finally, the control-to-output (plant) transfer function is the product of the three (cascaded) transfer functions, i.e. it becomes, using $s = j\omega$, and $j = \sqrt{(-1)}$:

$$G(s) = \frac{1}{V_{RAMP}} \times V_{IN} \times \frac{1/LC}{s^2 + s\left(1/RC\right) + 1/LC} \qquad \textit{(buck: plant transfer function)}$$

The LC post filter (third term) above will be discussed in more detail later. Here L is the buck inductor, C the output capacitor, and R the load resistor across the output terminals of the buck. This is an approximation so far, because we are ignoring the ESR (equivalent series resistance) of the output capacitor, and the DCR (DC resistance) of the inductor. Alternatively, this simplified plant transfer function can be written out as:

$$G(s) = \frac{1}{V_{RAMP}} \times V_{IN} \times \frac{1}{\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q}\left(\frac{s}{\omega_0}\right) + 1} \qquad \textit{(buck: plant transfer function)}$$

where $\omega_0 = 1/\sqrt{(LC)}$ is the resonant (break) frequency of the LC post filter, and $\omega_0 Q = R/L$. Or equivalently, $Q = R\sqrt{(C/L)}$.

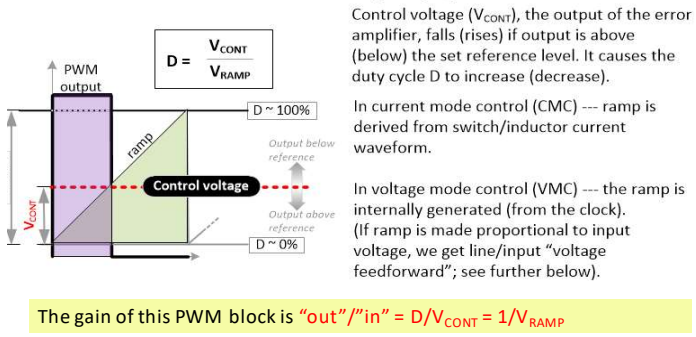
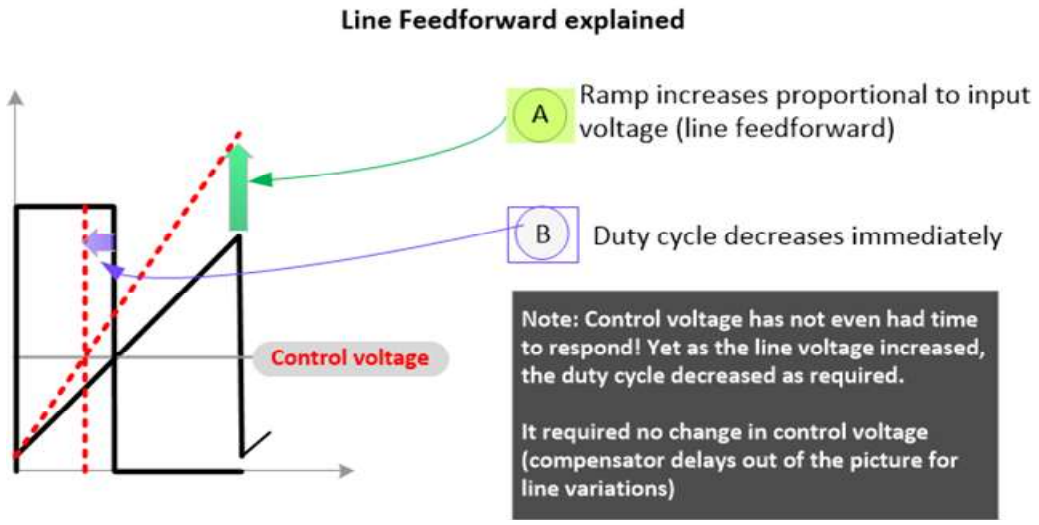


Figure 4.17: Gain of PWM stage

Input Line Voltage Feedforward

A way to speed up the response of a buck or buck-derived switcher (such as a forward converter) is to introduce a way of sensing the input voltage and changing its slope *proportionally*, as shown in **Figure 4.18**. So if the input increases by a certain factor, the duty cycle automatically decreases by the same factor—as is desirable based on the simple buck duty cycle equation $D = V_{OUT}/V_{IN}$, which basically implies that for a given output, the duty cycle is inversely proportional to the input.

With this feedforward technique, there is now "input/line rejection" because the correction to the line transient (any frequency or rate) is almost instantaneous. It barely depends on the control loop to act (belatedly), through all its inherent resistor-capacitor combinations and other delays. Indeed, the control loop will eventually fine-tune the correction, but the bulk of the correction is already complete, through this feature.



VMC with line feedforward is nowadays considered superior to CMC

Figure 4.18: Input voltage feedforward

Input Correction in Current-mode Controlled Buck

One of the oft-touted historical advantages of CMC, is *inherent line rejection*. Let us look at that a bit.

Figure 4.17 implies that the PWM ramp is created artificially from the fixed internal clock of the switcher. That is voltage mode control (VMC) of course. In current mode control (CMC), the PWM ramp is an appropriately amplified version of the switch/inductor current.
Though the line feedforward technique described in **Figure 4.18** is applicable only to VMC, the original inspiration behind the idea does come from *current mode control*— in which the PWM ramp, generated from the inductor current, automatically increases if the line voltage increases. That partly explains why current mode control seemed to respond so much "faster" to line disturbances than traditional voltage mode control at the time.
However, once **Figure 4.18** has been implemented, VMC has effectively imbibed the key advantage of CMC. One question remains: how good is the "built-in" automatic line feedforward of CMC, compared to VMC with line feedforward? *It turns out the latter is better*. Because in a buck topology, the *slope* of the inductor current up-ramp is equal to $(V_{IN}-V_O)/L$. So if we double the input voltage, we do *not* end up doubling the slope of the inductor current or the PWM ramp as desired. That means the duty cycle does not halve exactly, as we want it to, based on $D = V_{OUT}/V_{IN}$. However, in the case of VMC with line feedfoward, it does exactly that as explained in **Figure 4.18**.

In other words, voltage mode control with proportional line feedforward control, though inspired by current-mode control, provides *better* line rejection than current mode control (for a buck).

Transfer Functions of Other Topologies

For a boost topology, using its duty cycle equation, we can similarly derive the gain of the switch stage.

$$V_O = \frac{V_{IN}}{1-D}$$

$$\frac{dV_O}{dD} = \frac{V_{IN}}{(1-D)^2}$$

So, the plant transfer function is a product of three transfer functions: $G(s) = \frac{1}{V_{RAMP}} \times \frac{V_{IN}}{(1-D)^2} \times \frac{\frac{1}{\underline{L}C} \times \left(1 - s\left(\frac{\underline{L}}{R}\right)\right)}{s^2 + s\left(\frac{1}{RC}\right) + \frac{1}{\underline{L}C}}$ *(boost: plant transfer function)*

where **$\underline{L}$ = L/(1-D)^2** as discussed previously. It is *the inductor in the "equivalent post-LC filter" of the canonical model.* Also note that C remains unchanged. It is just C_{OUT}.
Alternatively, the above transfer function can be written as

$$G(s) = \frac{1}{V_{RAMP}} \times \frac{V_{IN}}{(1-D)^2} \times \frac{\left(1 - \frac{s}{\omega_{RHP}}\right)}{\left(\frac{s}{\omega_0}\right)^2 + \frac{s}{\omega_0 Q} + 1}$$ *(boost: plant transfer function)*

where $\omega_0 = 1/\sqrt{(\underline{L}C)}$, and $\omega_0 Q = R/\underline{L}$.

We have included a surprise term in the numerator, the RHP zero, which can be shown to be present in both the boost and buck-boost, after detailed modeling. Its location is

$$f_{RHP} = \frac{R \times (1-D)^2}{2\pi L}$$ *(boost)*

Similarly, for a buck-boost we get:

$$V_O = \frac{V_{IN} \times D}{1-D}$$

$$\frac{dV_O}{dD} = \frac{V_{IN}}{(1-D)^2}$$

(Yes, it is an interesting coincidence --- the slope of 1/(1-D) calculated for the boost, is the same as the slope of D/(1-D)) calculated for the buck-boost!). So the control-to-output transfer function is

$$G(s)=\frac{1}{V_{RAMP}}\times\frac{V_{IN}}{(1-D)^2}\times\frac{\frac{1}{\underline{L}C}\times\left(1-s\left(\frac{\underline{L}D}{R}\right)\right)}{s^2+s\left(\frac{1}{RC}\right)+\frac{1}{\underline{L}C}}$$ *(buck-boost: plant transfer function)*

where **$\underline{L}$ = L/(1-D)^2** is the inductor in the *equivalent* post-LC filter.
Alternatively, this can be written as

$$G(s)=\frac{1}{V_{RAMP}}\times\frac{V_{IN}}{(1-D)^2}\times\frac{\left(1-\frac{s}{\omega_{RHP}}\right)}{\left(\frac{s}{\omega_0}\right)^2+\frac{s}{\omega_0 Q}+1}$$ *(buck-boost: plant transfer function)*

where $\omega_0=1/\sqrt{(\underline{L}C)}$, and $\omega_0 Q=R/\underline{L}$.

Note that, as for the boost, we have included the RHP zero term in the numerator (in gray). Its location is similarly calculated to be

$$f_{RHP}=\frac{R\times(1-D)^2}{2\pi L\times D}$$ *(buck-boost)*

Enter: The Laplace Transform

Some of the rather unexpected situations described above, and others we will encounter, can be equally unexpectedly visualized more elegantly, in terms of the "dreaded" Laplace transform technique. We need to feel comfortable with it, and for that reason we will discuss it a bit here. We will also take this opportunity to restate, summarize or emphasize some of our key "lessons learned".

The best way to quell our fear of the Laplace transform is to understand that we are simply moving to an alternative *mathematical* domain to simplify computation. We have been doing the same thing for years using log math. See **Figure 4.19** and **Figure 4.20**. Complex multiplications or divisions of very large numbers get reduced to easy addition and subtraction instead. Of course, we rely on tables previously created, to go in and out of this logarithmic plane. So in a sense, the spade work was already done once and for all, by creating the log and antilog tables. Because, to return to the normal, non-logarithmic plane, we have to use antilog or inverse-log tables.

Figure 4.19: Using logarithms to simplify multiplication of large number

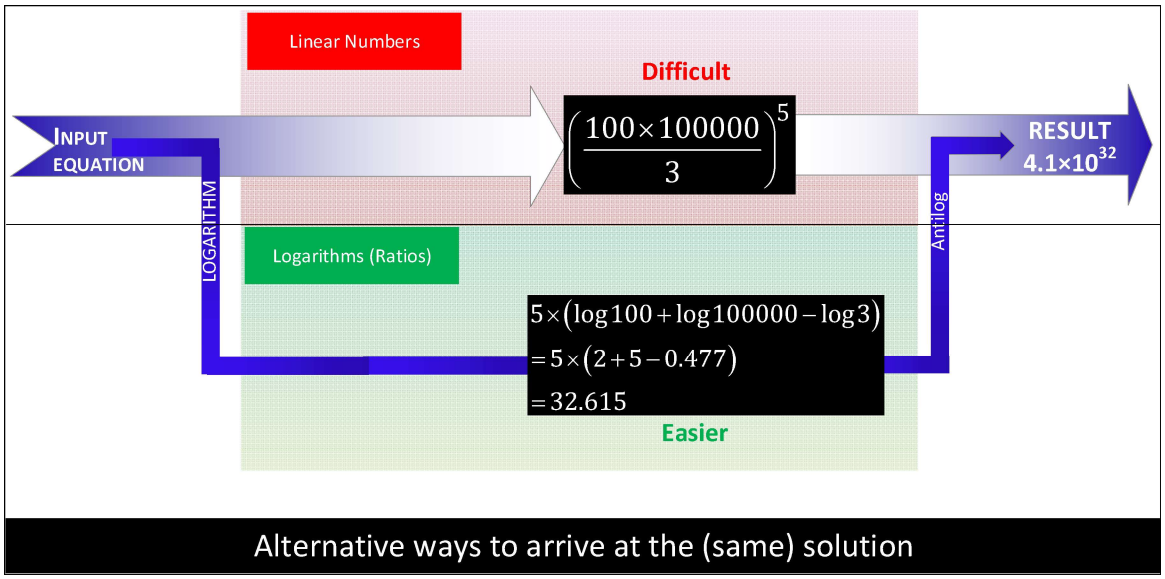


Figure 4.20: Visualizing the logarithmic plane

See **Figure 4.21** now, for the Laplace transform technique. This figure does *not* intend to show a closed-loop control system. Think of it for example, as just a simple filter stage, consisting of various capacitors, resistors and inductors. We apply an arbitrary input signal or impulse, and we are interested in seeing what happens at the output of this network. We discover that the differential equations to solve this problem (in the normal "time domain") become very complicated.

It also turns out that using the alternative "frequency domain" or "s-plane", i.e. the Laplace transform, the math becomes simpler. But once again we rely on a bunch of readily available tables, with the help of which we can move in and out of the new *mathematical computation domain*.

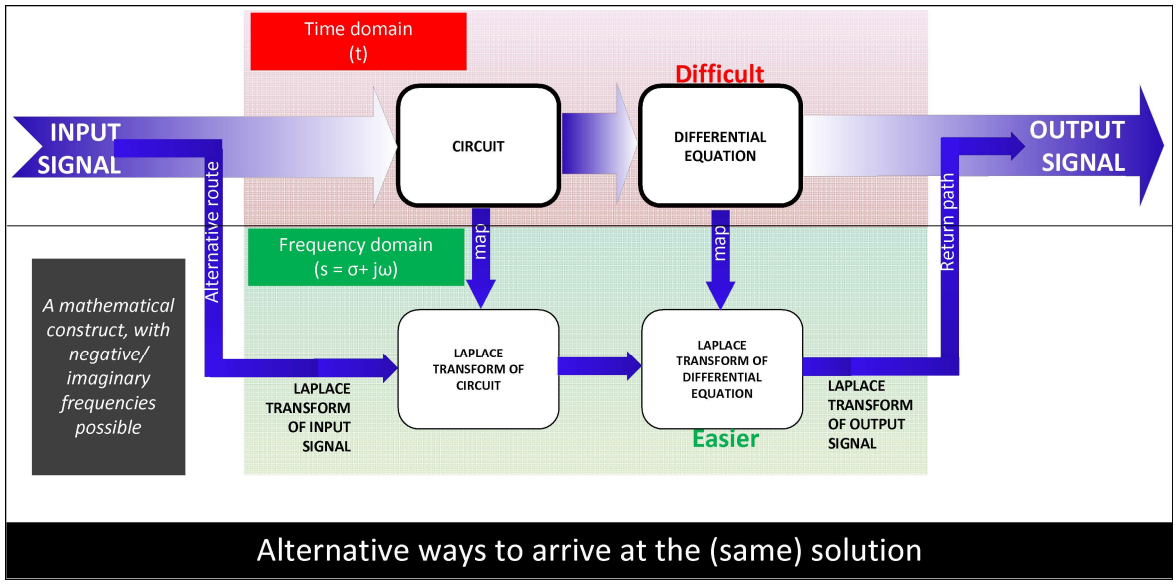


Figure 4.21: The Laplace transform technique (frequency domain analysis)

What exactly are we achieving by the Laplace transform? Essentially, we are breaking up an arbitrary non-repetitive, time-varying signal or impulse (the "disturbance"), into a *continuous* spectrum of both positive and negative frequency components (i.e. in the frequency domain). This is akin to the well-known Fourier analysis technique used for decomposing *repetitive* waveforms into *discrete* (positive) harmonics. Note that decomposition approaches are easier for the same reason that we routinely break up "force" in classical mechanics, into its "independent" x, y and z components, do the math for each "independent" axis separately, compute the x, y and z components of the acceleration, and finally sum them up vectorially to give us the final acceleration vector: i.e. its magnitude and direction. That is what we are doing in the Laplace transform method too, quite similar to what we did in high-school with the Fourier series, except that the decomposed frequencies are now a continuous spectrum. We will come to this in a little more detail in the next chapter.

Summarizing, think of **Figure 4.21** as a simple 2-port network, say a combination of several resistors and capacitors, with an input time-varying excitation (voltage or current), and we are trying to deduce the output waveform. Usually, we will need to set up complicated differential equations to solve the problem. However, the Laplace transform method allows us to take the Laplace transform of *both the circuit and the excitation*. As a result, this becomes a simple algebraic problem where we can sum over (integrate in our case) the result of the various frequency components. Finally, we take the *inverse* Laplace transform to map the result back into the time domain. We thus see the desired output voltage (or current). As mentioned, the reason we get away with this "simplification" is that all the drudgery has already been done beforehand in the form of comprehensive lookup tables for Laplace and inverse Laplace transforms.

Understanding Delays

With the Laplace transform technique, we can show that a certain time delay is equivalent to a "phase lag" in the frequency domain, one which is proportional to the frequency of the component.

As previously mentioned, in switching power supplies too, data is not sampled and acted upon continuously. In other words, there are inherent delays on account of the stream of discrete pulses. That also implies there will be increasing phase lag as we approach frequencies close to the switching frequency. To avoid meeting our criterion of instability (T = -1) prematurely, we must start by fixing the crossover frequency to less than one-fifth the switching frequency. That makes sense because the term "phase" (angle) has no relevance unless we are talking in terms of a specific frequency and its associated time period (of repetition)—which we are doing in this case, through our frequency decomposition technique via the Laplace transform. The "compensator" (feedback network) of the control loop can also introduce additional delays, with corresponding frequency-dependent phase shifts, which add to the inherent delays present in the response of the plant to the disturbance. These compensator delays are easier to understand, because the feedback circuit typically involves several resistors and capacitors, with a bunch of interrelated RC time constants. In particular, the capacitors present may also need a finite time to charge or discharge to their new average values through all their accompanying resistors. And that incidentally leads to the compensator's "poles" or "zeros", depending on how the elements are arranged within the high-gain error amplifier of the compensator. We will talk of Type 2 and Type 3 analog compensators in more detail soon.

We have seen that mathematically, if T = -1, we get full-blown instability. Intuitively, we can visualize instability as a situation where the system is trying to respond to a completely *outdated/ delayed command* without realizing it (reading "up" instead of "down" for example), and thus continues to head in the opposite (wrong) direction every time. The delay is response now is exactly one half-cycle, the word cycle referring to the decomposed component's culprit frequency. This is full-blown instability, even though the system itself does not literally explode.

Large Signal Response in Other Topologies

As mentioned, the majority of "disturbances" or excitations encountered in switching power supplies are not "small-signal" as many probably still assume, but large-signal. Under such bulk stimuli, it can take a certain finite time for the current in an inductor to slew up or down to the new desired "average" value commensurate with the new steady load condition. See **Figure 4.4** once again. We also remember that this inability of the inductor and the switch to provide the required power during this "reinitialization" duration is *equivalent to a sudden drop in gain*—as the closed-loop system struggles to correct itself. This can lead to full-blown instability if "conditional stability" is already present.

The maximum slew rate of the inductor current is primarily dependent on the "timeless", and shall we say "stubborn", equation of an inductor: $V = L\,\Delta I/\Delta t$. Though there are ways to mitigate the resulting delay, such as hysteretic control, ultimately, this is the downside of using reactive (energy-storing) components (inductors and capacitors), compared to the fast-acting but wholly *inefficient* resistive elements employed in linear power supplies. We always need *time*, to either build up or deplete stored energy in a controlled manner. But we can dissipate almost immediately! That is equivalent to saying: resistors have no inherent delays.

And that is why, as mentioned earlier, if we suddenly go from 0A to 5A load in a switcher for example, the initial dip in the output voltage may have little to do with the control loop characteristics. Indeed, the control loop can always make things worse, but even if it is considered optimized here, the output response may eventually be determined only by the output bulk capacitance vis-à-vis the inductance, since the output capacitor needs to supply the entire additional energy demanded till the current in the inductor can ramp up and take over. And if the capacitance happens to be inadequate to start with, we need to refer back to **Figure 4.7**.

Similarly, as also explained in **Figure 4.7** if we suddenly go from 5A to 0A (unloading), we may find to our horror that though we saw the output voltage jump up and respond by even *halting switching action completely*, the output rail continued to rise, almost out of our control for a short while. The reason for that is the inductor stubbornly pushes out all the stored energy related to its initial current setting, into the output capacitors where it can be stored indefinitely as electrostatic energy as required, since the attached load isn't demanding energy anymore.

Perhaps that reminds us of the hazy outlines of our forgotten Physics 101 course: *energy can be converted* or dissipated as heat (in resistors)*, but never wished away*. That is why we need catch diodes (and output capacitors) in any switcher in the first place. To freewheel the current associated with the stored magnetic energy. Dispense with the diode, and we get a worthy spark ignition system instead of a switcher—lots of heat and light, but no useful power.

But there are still some remaining "subtleties" with regards to *non-buck* topologies, as shown in **Figure 4.4**. For example, during a line (input) disturbance, as opposed to a load transient, things become quite different for the boost and the buck-boost topologies, as compared to the simple scenario described in the figure for a buck topology on the left-hand side. The reason is, in a buck, the average inductor current equals the load current (in steady state), and is therefore constant during a line disturbance. There is no delay attributable to any inductor "reinitialization" problem, except of course for a load transient.

However, in a boost or buck-boost, the average inductor current *is* a function of the duty cycle, unlike a buck, and thus needs to move to a new average value if we vary the input voltage, even if the load current is held constant! So now, the inductor reinitialization issue returns to haunt us, along with the other inherent delays present in the control loop—even during a supposedly "pure line disturbance"! See the right-hand side of the figure.

This goes to show that not all "disturbances" are alike, *nor all topologies*, and we must be cognizant of quite a few such unexpected "subtleties" when we try to move basic control theory concepts over to switchers.

Pitfalls of Terminology

We remind ourselves once again that to prevent smaller errors of perception from snowballing into masses of confusion, we must be very clear about the exact meaning of all the terms in common use.

As mentioned, a prime example of that is the concept of "closed loop gain". And the other side of the very same silver coin: "(open) loop gain" or "T". Many power supply engineers continue to think that open loop gain is some sort of amplification factor that gets applied to a vague, unspecified "disturbance" when the feedback loop is *literally* "open": i.e. broken or non-existent. Then, as a corollary, they assume that in contrast, closed loop gain must be what we measure when the feedback loop is actually present! A sideshow of this confusion is that a hands-on power supply engineer may wonder why, when he or she runs Bode plots using a standard bench network analyzer, the machine claims that it is measuring the "open loop gain". "Why isn't it giving us the closed loop gain, considering the fact that the loop is in reality closed?" And so on. One wrong premise leads to many wrong conclusions. Some engineers wisely use the term "loop gain" instead of open loop gain, as we too eventually finally did on previous pages. Others prefer to call "T" the "round transfer function". It can get a bit confusing.

Stepping the Reference Voltage

Many switcher engineers/authors did realize early on that closed loop gain was V_{OUT}/V_{REF}, not V_{OUT}/V_{IN}. But then, instead of giving examples showing line/load disturbances, they inadvertently propagated the fallacy further by documenting the overshoots and ringing when the reference voltage is ramped up suddenly from 0V. For example, we will often see in related literature the case of a "1/s" "step disturbance" applied to the system, where $s = j\omega$ as usual. But in this case, the step is the *shape of the reference voltage*. You may wonder why it is relevant.

We need to remember that:
a) Every power converter starts up initially with its reference rising up to its set value, so that hardly qualifies as a "disturbance" of interest to us.
b) Besides, the reference voltage, and the output voltage, rarely come up "instantly", since in a practical case both are usually brought up gradually under the influence of a closed-loop soft-start circuit. The reference is typically slow to rise, as it comes via a 0.1 μF ceramic decoupling capacitor placed on the current-limited REF pin, which is charged up slowly.
c) Even if we assume V_{REF} did come up abruptly, the output itself would take a very long time after that, relatively speaking, to ramp up to its steady value, since it has to first charge the rather sizeable output bulk capacitance across it, through the intervening (slew-rate limiting) inductor. So this response scenario has really nothing to do with jerking/wiggling the reference voltage around, even if that is considered relevant.
Indeed, the way the output comes up *with no soft-start implemented*, may show some ringing which *qualitatively* mimics the ringing observed during line and load transients.

The K-Factor Method

We should also point out that a practical, often overlooked problem to measuring what some engineers still call "open loop gain", is that in a modern high-gain switcher, there may be no easy way to "break" up the loop, i.e. to literally open it, without causing disastrous effects on the output. Leave aside testing it successfully in that state.

Indeed, that can be done on occasion. As when trying to stabilize relatively *low-gain* mag-amp post-regulators, using the oft-mentioned K-factor method.

The venerable K-factor method from Mr. Venable, a subject of many popular articles on feedback control, implicitly requires knowledge of the gain with *no* feedback present, i.e. with the loop broken—as a means of optimizing the feedback loop when it is finally introduced. And so, even though some engineers insist that the K-factor is all that is ever required for stabilizing switchers, it is usually impractical in most modern cases.

Besides the practical aspects, the K-factor technique also unfortunately trades off gain for phase margin by reducing gain at low frequencies and increasing it at higher frequencies—quite the opposite of what we usually try to do, for reasons we will soon discuss. That is why the K-factor technique is perhaps only well-suited for post-regulators, where a steady almost ripple-free DC input rail is present to start with— such as mag-amps! But rarely otherwise. The K-factor method as applied to a Type 3 amplifier attempts to put two coincident zeros a factor of $1/\sqrt{K}$ below crossover and two coincident poles a factor of $\sqrt{K}$ above crossover. As we will see, that does nothing to the *peaking of the LC pole*, which is not only a potential cause of conditional stability, but affects the ringing at the output during line/load transients as we will soon uncover.

But the irony of it all is, perhaps all the effort that the K-factor based "optimization" effort represented, was ultimately for only a questionable improvement in something called the phase margin. Questionable, because no one seems to agree fully what is the "optimum phase margin", *and why*.

Practical Analog Compensation Strategy

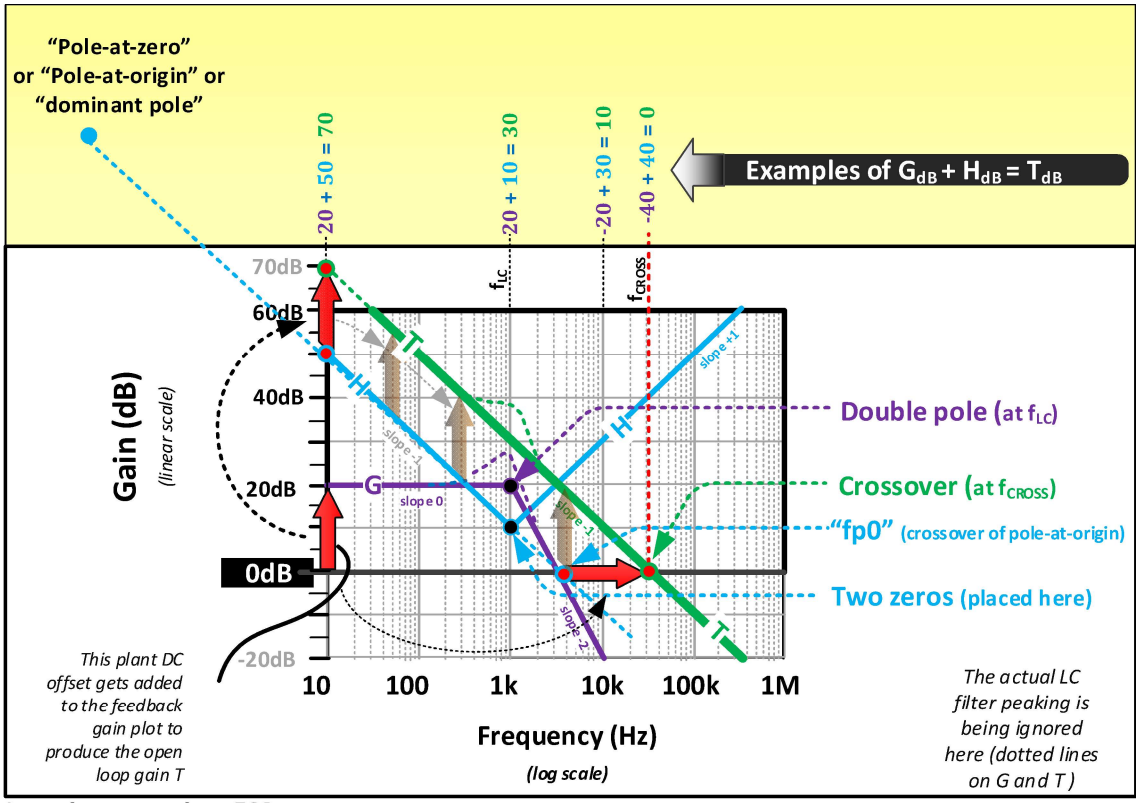
Let's recapitulate: stability ultimately depends on the following basic question: What happens if the disturbance undergoes various delays as it goes around the loop? For example, even in the simple case of a thermostatic room air-conditioner, a) it may take time for the sensor to feel the change if a window is thrown open. b) After that, the heater or air-conditioner will need some time to activate and respond. And so on. As also mentioned earlier, these delays can be modeled as frequency-dependent phase lags. And so, though we have usually been writing out the gain functions as simply T, G, H, etc., in reality they should be written as T(s), G(s), H(s), etc., indicating their *frequency dependence* and *inherent phase angle* too, besides their magnitude.

Similarly, for a switcher, we can now visualize a situation where an *additional 180° phase shift* can easily occur for some unspecified "harmonic" (frequency component) of the original disturbance. It will then reinforce itself after going around the loop. The result is the system could break up into oscillations.

For rejecting low- to mid-frequencies, it is now obvious that we need to try and maximize the DC gain. But we also need to deliberately *roll the gain off at higher frequencies* to avoid instability. Ultimately, we have to adhere to a simple criterion: *We have to ensure that at the specific frequency where an additional 180° of phase lag occurs, the gain falls below 1 to avoid oscillations (gain margin).* That will ensure any disturbance will get abated as it goes around the loop. Alternatively, we must ensure that when the signal comes around full-circle with a gain of 1, its phase lag is not enough to reach the ominous level of -180°.

Typical Analog Compensation exercise

In **Figure 4.22** we show a typical compensation exercise, but in terms of gain magnitudes only so far. It breaks up the "almost" straight line of T into its constituent G and H components. Note that the double LC "pole" of the plant, which is responsible for the "-2" slope of G after the breakpoint, has been (roughly) canceled out by placing *two "zeros" at the exact position of the LC, by the compensator*. So we are left with an (almost) straight line for T, essentially coming from the compensator's low-frequency gain profile. Though that is displaced vertically by the exact amount equal to the DC gain of the plant, as we will see shortly.
Note that the loop gain T is simply the product of the two cascaded gains, G and H, but on a log plane, things are easier. We can just sum the two curves of G_{dB} and H_{dB}, to give us T_{dB}. In other words, once we express the gains in decibels, we can just sum them up, rather than multiply them out, as shown in the figure.



Ignoring capacitor ESR-zero

Figure 4.22: Classic analog control loop design (simplified) and resulting loop gain

In the figure, we can also see we have set a very high DC gain (via the compensator) —as recommended for attenuating low- to mid-frequency disturbances. In fact, theoretically, the gain is infinite at 0 Hz, but in reality it gets limited by the inherent characteristics of the error amplifier, though that is not shown in the figure for the sake of simplicity, and also because it is practically impossible to display 0 Hz on a log scale anyway. But we have shown a dotted line extending to some very low frequency, and that is called the "pole-at-origin" or "pole-at-zero", among other monikers. But it has no location that we can really specify or draw out. What we do know however is, wherever it is located, it causes the gain to fall at "-1" slope thereafter. So its exact location, i.e. how low-frequency it really is, is reflected only by *the frequency at which it intersects the unity gain (0dB) axis*. That frequency is what we are calling "fp0" here. Indeed, we may place two zeros in the compensator well before the H_{dB} curve ever gets to cross the 0dB axis. But if we draw a dotted line to the 0dB axis, the intersection frequency is fp0 as shown in **Figure 4.22**.

"fp0", the crossover frequency of the pole-at-origin, needs to be set very carefully because it is the key parameter which ultimately determines the crossover frequency of interest to us: f_{CROSS} (the crossover of T). Both fp0 and f_{CROSS} are related through the DC gain of the plant, which in turn is completely responsible for the vertical arrow shifts shown in **Figure 4.22**. We will now see what that exact relationship is.

The plant, as we know by now, has three main constituents, and its gain is the product of its three stages.

$$G(s) = \frac{1}{V_{RAMP}} \times V_{IN} \times \frac{1/LC}{s^2 + s\left(1/RC\right) + 1/LC}$$

So its DC gain (the flat portion of G up to the break frequency in **Figure 4.22**) is simply V_{IN}/V_{RAMP} (or 20 log of that in decibels). As we can also see from this figure, this is the amount by which the compensator gain profile gets shifted upwards to create "T". Since we are dealing with a "-1" (inverse proportionality) curve for T, it is easy to see that the following relationship tells us how exactly we must position fp0 from the compensator, to achieve a certain desired f_{CROSS} for T.

$$fp0 = \frac{V_{RAMP}}{V_{IN}} \times f_{CROSS}$$

So, the loop gain curve in **Figure 4.22** "crosses over" at a frequency f_{CROSS}, which implies a gain of 1 at that frequency (on a log scale it is the zero of the y-axis since log 1 = 0). To stay well away from any phase lag effects causing total instability on account of the discrete/sampling issues related to the switching frequency of switchers, f_{SW}, it is customary to set f_{CROSS} to at least less than $f_{SW}/2$ ("Nyquist's sampling criterion"). But in fact, it is far better to set the crossover somewhere between $f_{SW}/10$ to $f_{SW}/5$. Not higher.
Note that the 180° inherent phase lag on account of negative feedback ("negative" though only at low frequencies as we now realize), is rarely plotted out. It is "understood". Only the *additional* phase shift introduced by the feedback network and the plant combined, is displayed on a typical Bode plot.

There are also other parasitics that come into the picture, which we have neglected so far. One is the ESR-zero, coming from the ESR of C_{OUT}. Using a Type 3 compensator, we try to kill this zero (from the plant), by placing a pole at the same exact position. But besides the pole-at-origin, a Type 3 compensator produces 2 zeros (both of which we have used up already, to cancel the LC double pole), but also 2 poles, one of which we can use to kill the ESR-zero. See **Figure 4.23**. That leaves us with one additional pole to play with, called "fp2" here. Some people say we should place it at $10\times f_{CROSS}$, others say that to attenuate the high-frequency ripple component, we need to place it at $f_{SW}/2$. Lloyd Dixon suggested placing it at exactly f_{CROSS}. This was called the "optimized solution" in A to Z, Second Edition.

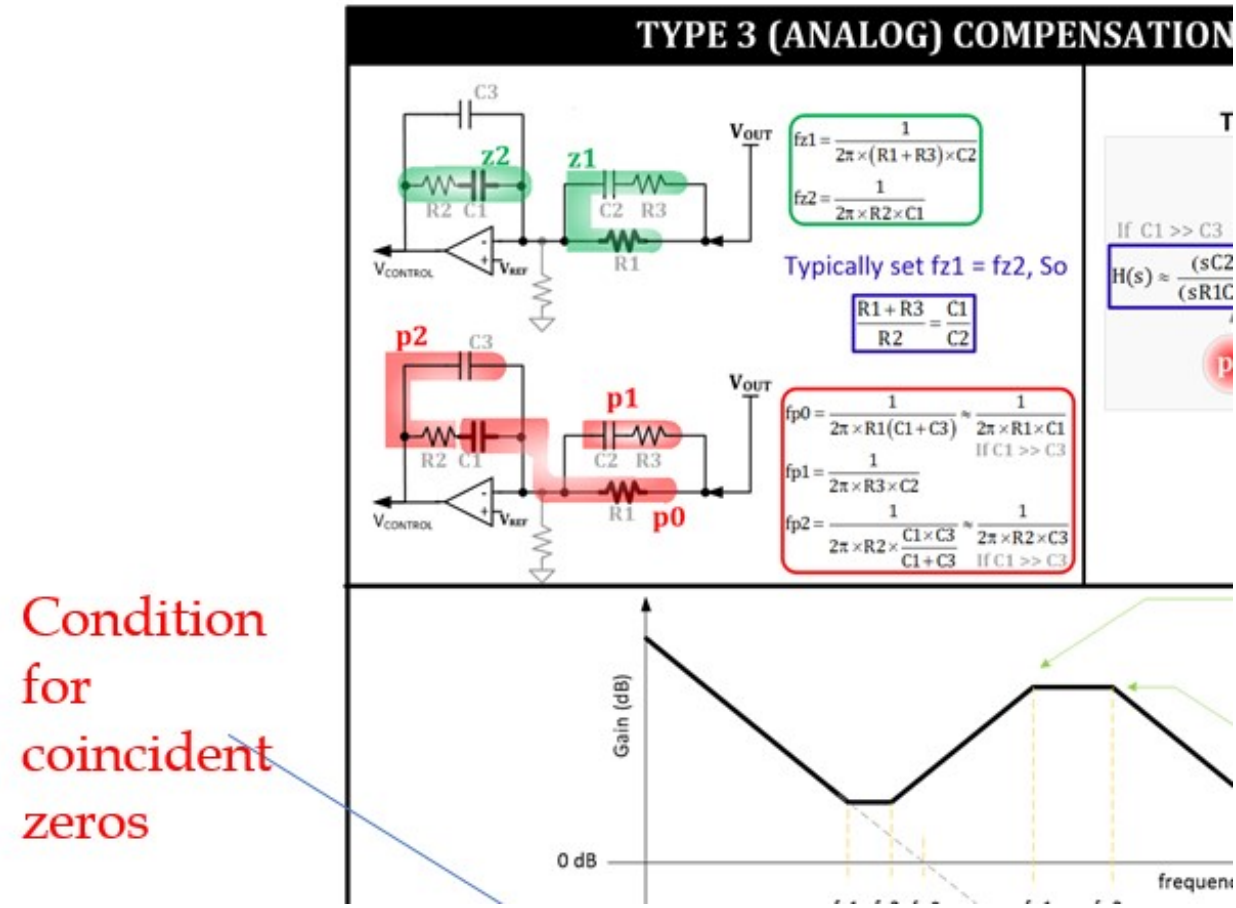


Figure 4.23: Type 3 compensator equations and strategy

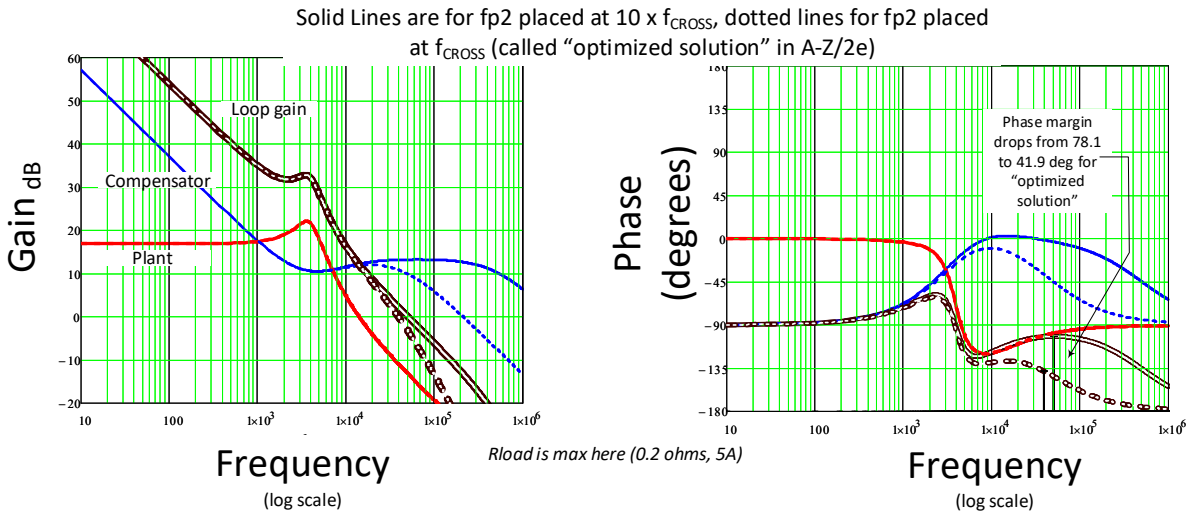


Figure 4.24: Results of an actual Type 3 compensation exercise at maximum load

In **Figure 4.24**, we have shown the poles and zeros from a typical analog control loop exercise, extracted from A to Z, Second Edition. Note that this is no longer the "asymptotic approximation" used in **Figure 4.22**. It has all the curved regions, plotted out using Mathcad.

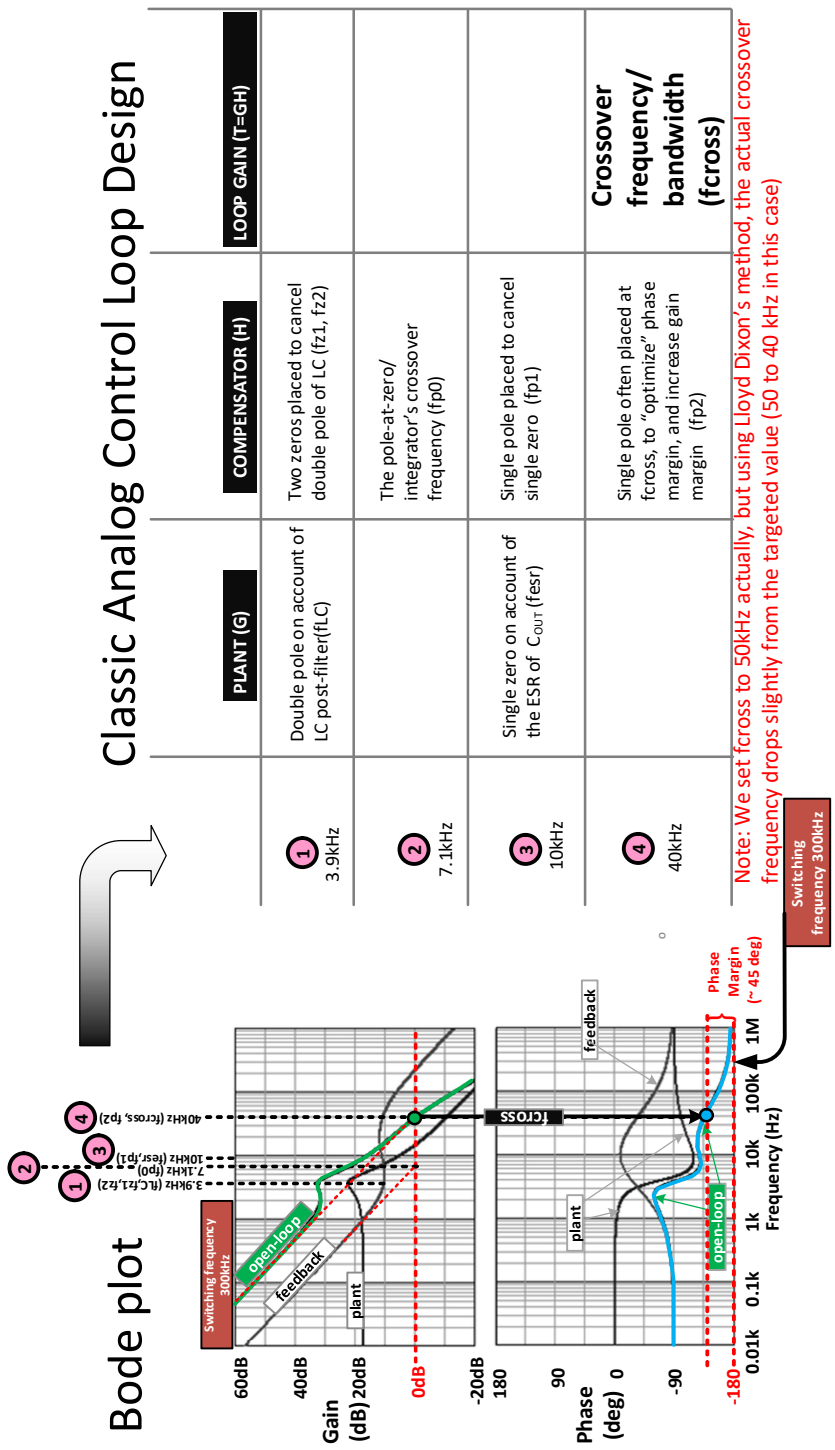
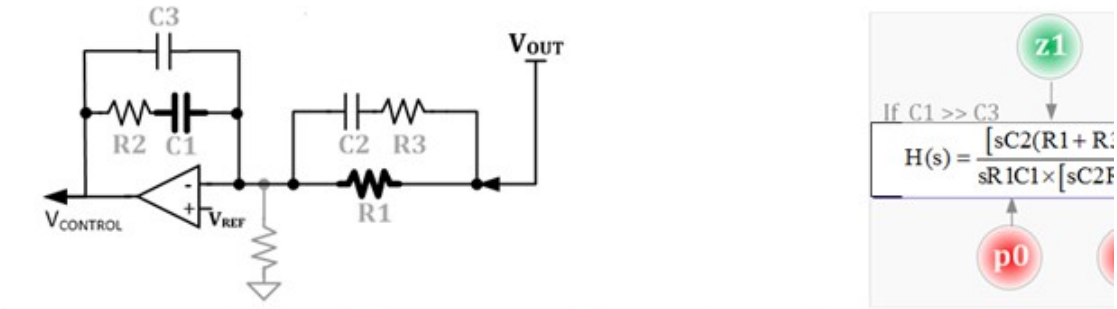


Figure 4.25: Summary of pole-zero placements for previous figure and example

In **Figure 4.25**, we summarize the compensation strategy we have been talking about, showing exactly what happens (in terms of the frequencies involved), for the plant G, compensator H, and of course "T".

In **Figure 4.26**, we present a table of the components of a Type 3 compensator, based on the transfer function equation in **Figure 4.23**, showing how all but one component, are involved in more than one pole/zero position. *Which is why it becomes so difficult to change anything in an analog loop.* Changing just one component can have a domino effect on the gain curves, with "unintended" consequences. We will discuss this in more detail shortly.



	p0	p1	p2	z1
R1	●			●
R2			●	
R3		●		●
C1	●			
C2		●		●
C3			●	

Figure 4.26: How each component of a Type 3 compensator plays a multi-role

Types 1, 2 and 3 Compensators Summarized

In **Figure 4.27**, we show other compensators too for completeness sake. Type 2 for example, offers only one pole and one zero (besides the pole-at-origin, which all three offer). It is therefore unsuitable for VMC since we need two zeros to cancel the LC double-pole. However, some older switchers try to use the ESR-zero for that purpose, along with one zero from the compensator, and therefore do not attempt to kill the ESR-zero. However that is not an optimum solution at all. Generally, Type 2 compensators can only be used with CMC, because its plant does not have the LC double-pole, instead featuring an RC-based (first order) "load pole" as shown in **Figure 4.28**.

The Type 1 compensator is of no practical use *on its own*, but is a key building block of Type 2 and Type 3 compensators. It is an integrator. It is the where the pole-at-origin, fp0, comes from. The key general function behind it is plotted out in **Figure 4.29**. It has the form

$$H(s) = \frac{A}{s} \equiv \frac{1}{s/A} \equiv \frac{1}{s/\omega_0}$$

This crosses over at $\omega_0 \equiv \omega p0 = A$. Or equivalently $fp0 = A/2\pi$. So, we can adjust it (translate it upwards or downwards), by changing A.

Note: *It is always preferable to write all pole and zero functions in the form* $(s/\omega_0)^x$*, to avoid needless confusion about the DC gain contribution from the functions.*

The math behind the op-amp embodiment of this function, the integrator, is shown in **Figure 4.30**. We get the equation for fp0 as:

$$fp0 = \frac{1}{2\pi RC}$$

Note: In a Type 3 compensator (**Figure 4.26**) this pole-at-origin (integrator function) is created by R1C1, not R1C3 as commonly and erroneously assumed. The reason for that is, we are working under the assumption C1>>C3; otherwise the mathematical solutions to the locations of the poles and zeros, are extremely intricate, and thus unusable.

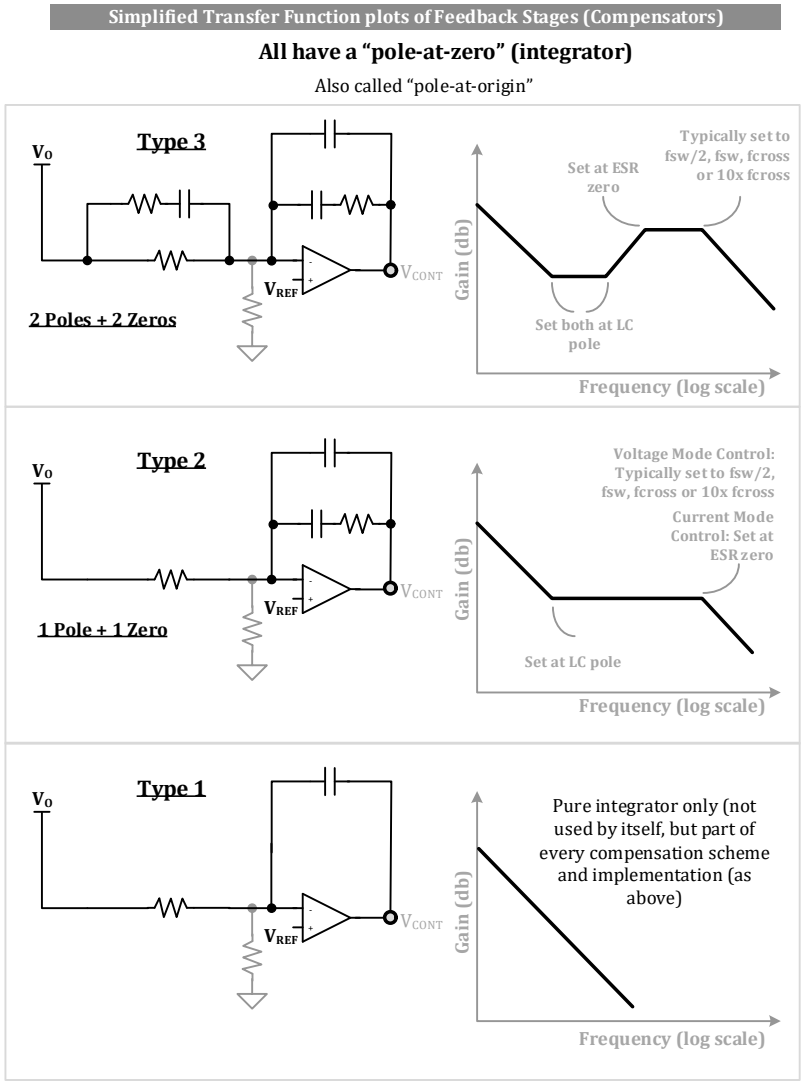


Figure 4.27: Type 1, 2 and 3 compensators

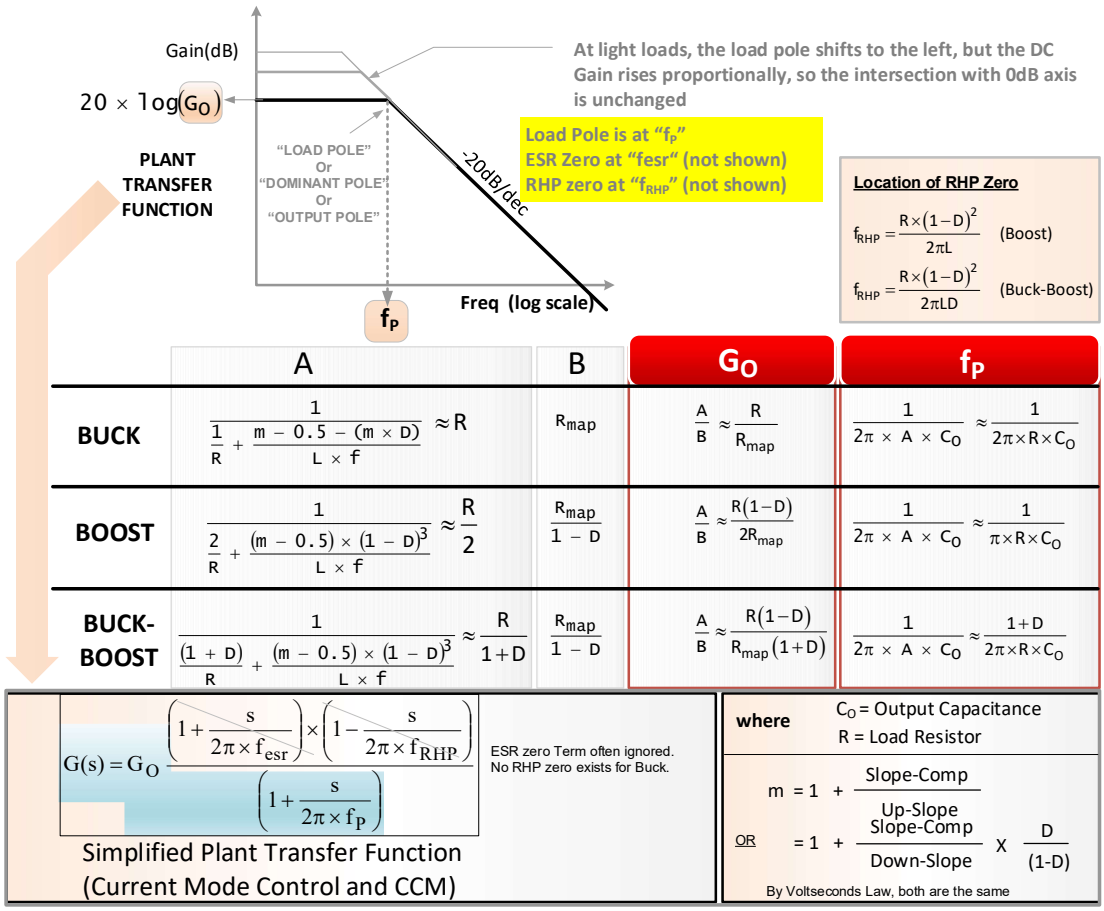


Figure 4.28: The plant in current mode control (suitable for Type 2 compensator)

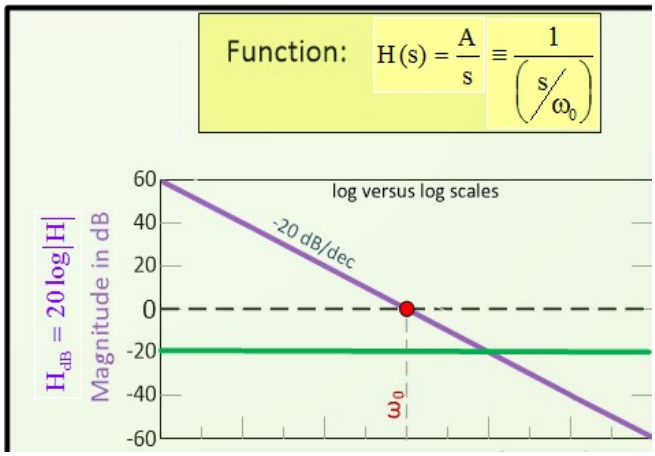
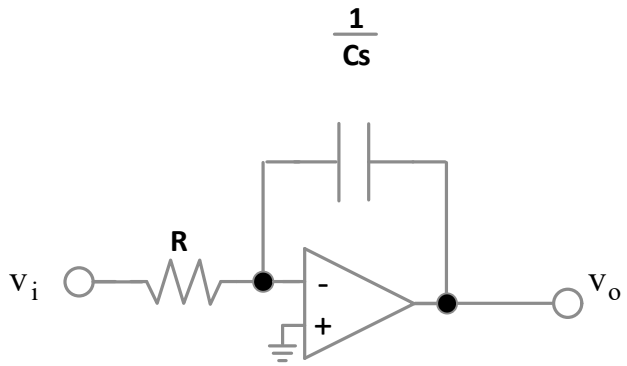


Figure 4.29: The integrator function (inversely proportional to frequency)



$$\text{Transfer function} = \frac{v_o}{v_i} = \frac{1/Cs}{R} = \frac{1}{RCs} \equiv \frac{1}{\left(\frac{s}{1/RC}\right)} \equiv \frac{1}{\left(\frac{s}{\omega p0}\right)}$$

$$\text{So}: \quad \boxed{fp0 = \frac{1}{2\pi RC}}$$

Figure 4.30: Analog implementation of an integrator

Compensating Other Topologies

As mentioned, we can luckily break up the switch and L-C combination of a boost and buck-boost into a switch followed by a distinct and equivalent cascaded post-filter stage, consisting of the same C (output capacitor), in series with an "equivalent inductor" of value:

$$L_{equivalent} \equiv \underline{L} = \frac{L}{(1-D)^2}$$

In effect, that makes the effective inductance *a function of the input voltage*. Hence the treatment can get rather complex, since the LC resonant frequency moves to higher and higher frequencies as we lower the input voltage (higher D). And that is, intuitively, what eventually contributes to the RHP zero instability mentioned earlier. The conventional solution to the RHP zero problem is to virtually accept that there is no solution! We just have to roll-off the loop gain at a much lower frequency than the typically targeted $f_{SW}/10$ to $f_{SW}/5$ for a buck. Maybe we need to go closer to $f_{SW}/20$, or even lower. Which is also why we can hardly expect excellent control loop response from, say, a typical power factor correction (boost) stage, or a "cheap and dirty" flyback (buck-boost).

Summary of Plant Transfer Functions

Finally, we present a summary of the plant functions for VMC and CMC, for easy reference. In **Figure 4.31, Figure 4.32, Figure 4.33 and Figure 4.34,** we have the buck, boost and buck-boost (all in VMC) respectively, followed by the buck in CMC.

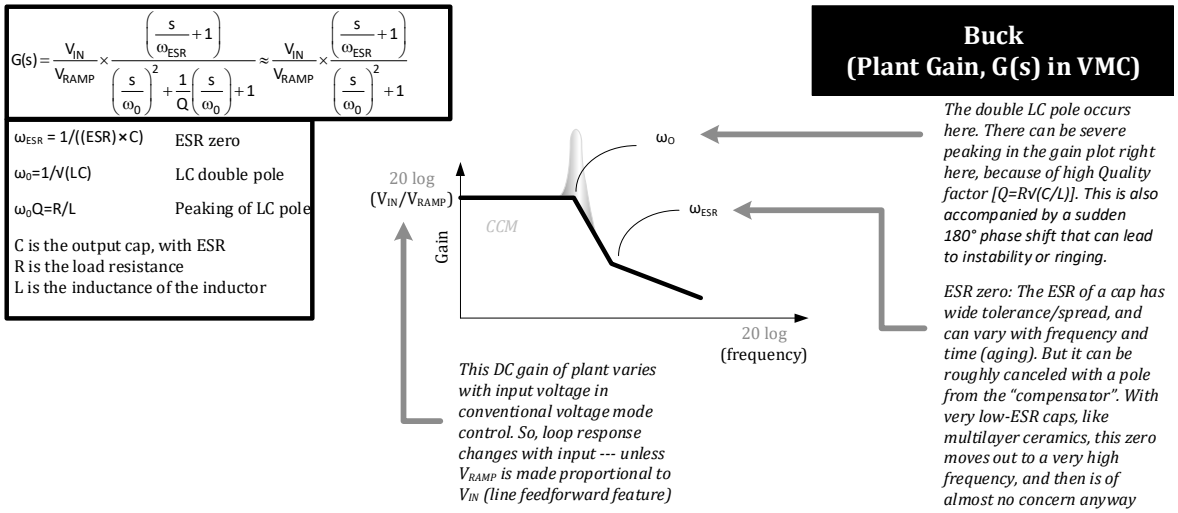


Figure 4.31: Buck in VMC

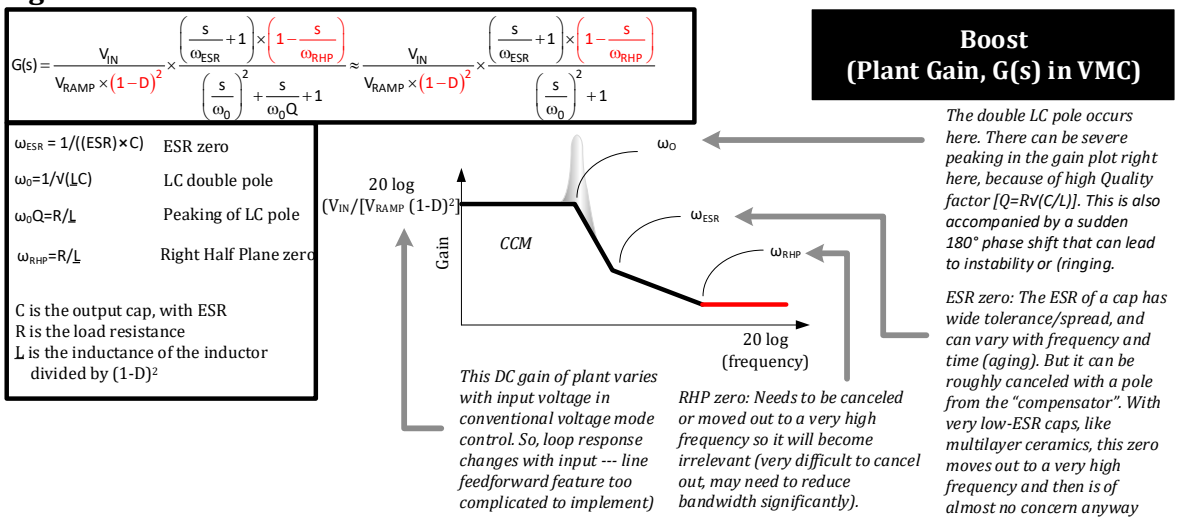


Figure 4.32: Boost in VMC

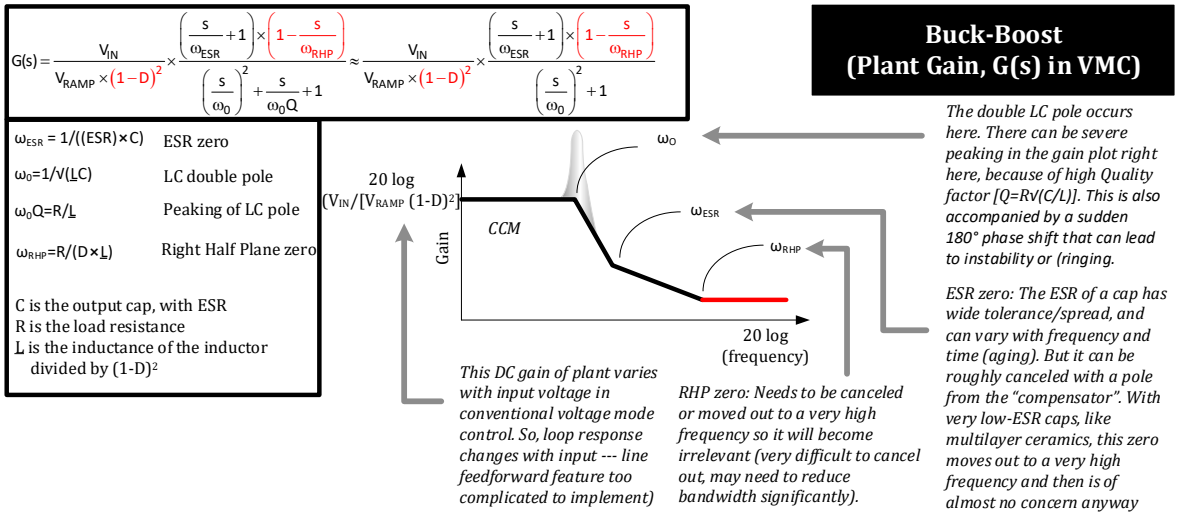


Figure 4.33: Buck-boost in VMC

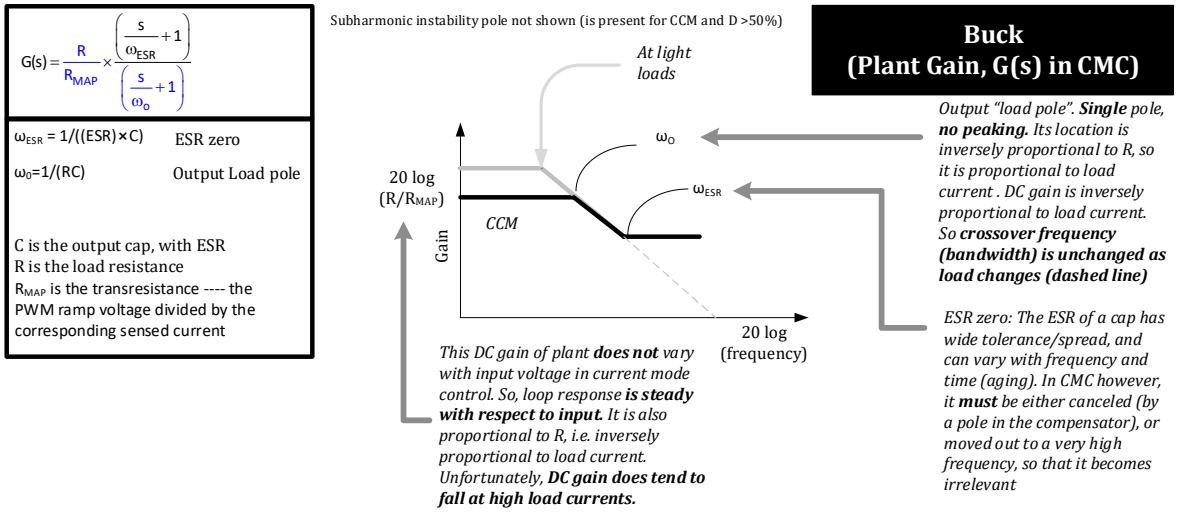
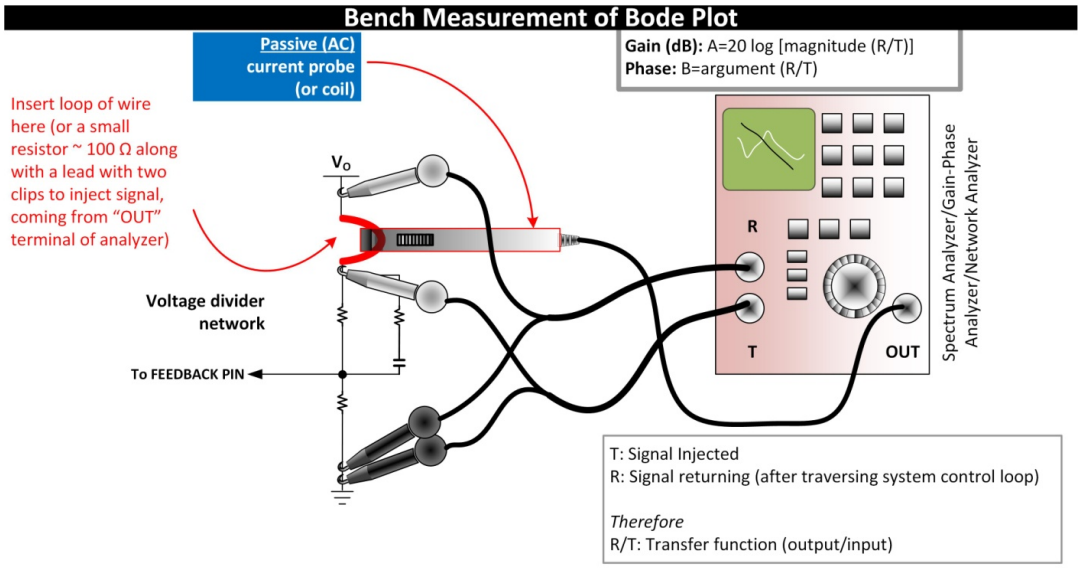


Figure 4.34: Buck in CMC

Measuring Loop Gain on the Bench

A simple way of collecting a Bode plot on a switcher is shown in **Figure 4.35**. A current loop and a *passive* current probe (snap-on coil) are the basic requirements. A standard HP/Agilent network analyzer such as the 4396B are required. No need typically, for more complicated setups as from Ridley or Venable.

And indeed, as the math in **Figure 4.36** shows, we do measure T ("open loop gain") on the closed loop system, provided we inject the signal at a suitable point, as in **Figure 4.35**.



Check SW (switching) node on scope while test is in progress: The "jitter" should not be more than ~10% of the time period of the switching cycle, and not less than ~2%. Otherwise adjust output amplitude/attenuation settings on spectrum analyzer.

Figure 4.35: A simple way of doing a loop gain-phase (Bode) plot on the bench

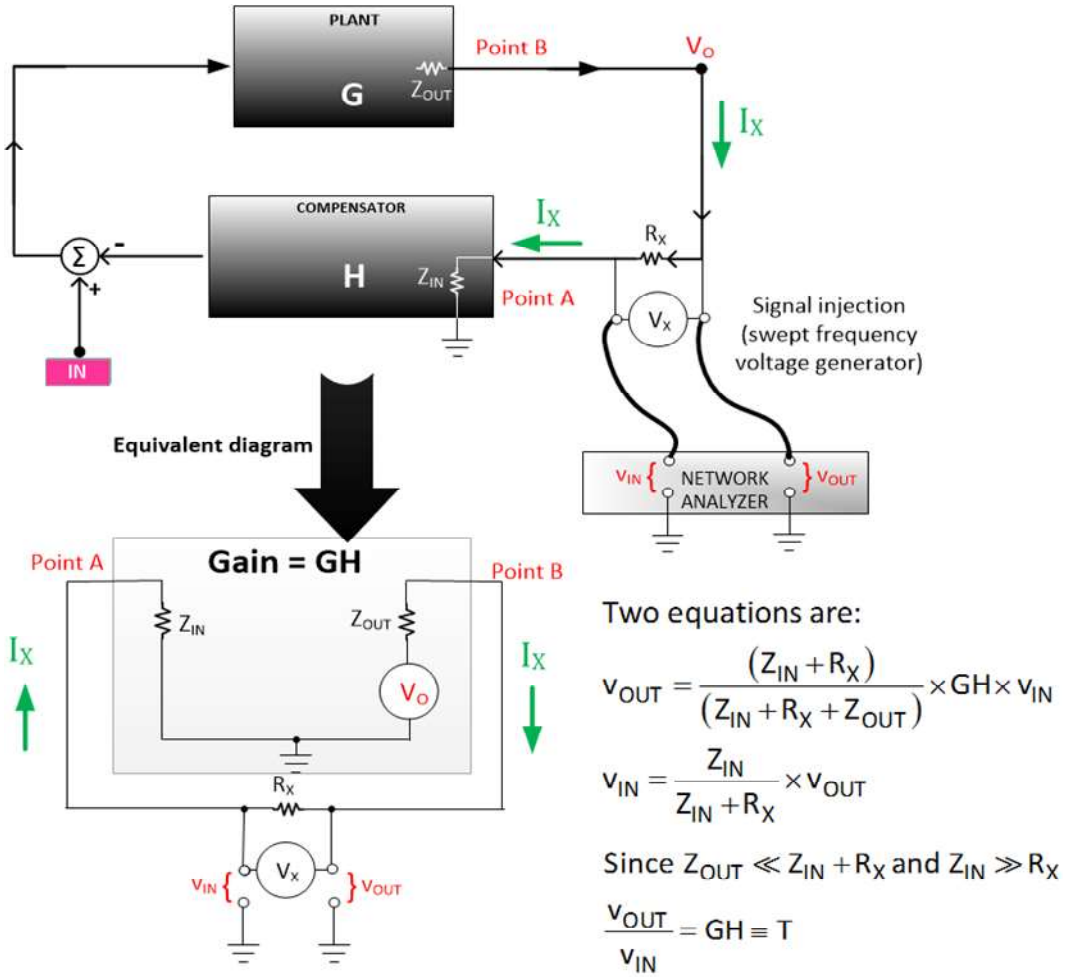


Figure 4.36: Indeed, we measure (open) loop gain on a closed loop setup

Tweaking a Type 3 Compensator

One of the big nuisances regarding typical analog compensators is the difficulty of tweaking any aspect of the gain profile.

We will illustrate this with an actual example shortly. But the baseline for that is the following solved example.

Example: *Using a 300 kHz synchronous buck controller we wish to step-down 15V to 1V. The load resistor is 0.2Ω (5A). The PWM ramp is 2.14V as per the datasheet of the part. The selected inductor is 5 μH, and the output capacitor is 330 μF, with an ESR of 48 mΩ*

We know that the plant gain at DC for a buck is V_{IN}/V_{RAMP} = 7.009. Therefore, (20 × log) of this gives us 16.9 dB. The LC double pole is at

$$f_{LC} = \frac{1}{2\pi \times \sqrt{LC}} = \frac{1}{2\pi \times \sqrt{5 \times 10^{-6} \times 330 \times 10^{-6}}} \Rightarrow 3.918\ \text{kHz}$$

Here we want to set the crossover frequency of the open-loop gain at $1/6^{th}$ the switching frequency, i.e. at 50 kHz. Therefore we can solve for the integrator's fp0 and thereby its "RC" using

$$fp0 = \frac{V_{RAMP}}{V_{IN}} \times fcross \equiv \frac{1}{2\pi \times RC}$$

So in our case, the integrator's RC is

$$R_1C_1 = \frac{V_{IN}}{2\pi \times V_{RAMP} \times fcross} = \frac{15}{2\pi \times 2.14 \times 50 \times 10^3} = 2.231 \times 10^{-5}\ \ s^{-1}$$

If we have selected R_1 (upper resistor of divider) as say 2kΩ, C_1 is then

$$C_1 = \frac{2.231 \times 10^{-5}}{2 \times 10^3} \Rightarrow 11.16\ \text{nF}$$

The crossover frequency of the integrator section of the op-amp is

$$fp0 = \frac{1}{2\pi \times R_1C_1} = \frac{10^5}{2\pi \times 2.231} \Rightarrow 7.133\ \text{kHz}$$

The ESR-zero is at

$$fesr = \frac{1}{2\pi \times 48 \times 10^{-3} \times 330 \times 10^{-6}} \Rightarrow 10.05\ \text{kHz}$$

The required placement of zeros and poles is:

$fz1 = fz2 = 3.918\,\text{kHz}$ (place both zeros at at the LC pole location)

$fp1 = fesr = 10.05\,\text{kHz}$ (place first pole to cancel ESR zero) s

$fp2 = 10 \times fcross = 500$ kHz (unoptimized/standard solution)

(We can set fp2 = f_{CROSS} for "better" phase margin)

The components required to make this happen are (solutions of several simultaneous equations)

$$C_2 = \frac{1}{2\pi \times R_1}\left(\frac{1}{fz1} - \frac{1}{fp1}\right) = \frac{1}{2\pi \times 2 \times 10^6}\left(\frac{1}{3.918} - \frac{1}{10.05}\right) \Rightarrow 12.4 \text{ nF}$$

$$R_2 = R_1 \frac{fp0}{fz2} = 2 \times 10^3 \times \frac{7.133}{3.918} \Rightarrow 3.641 \text{ k}\Omega$$

$$C_3 = \frac{1}{2\pi \times \left(R_2 fp2 - R_1 fp0\right)} = \frac{10^{-6}}{2\pi \times \left(3.641 \times 500 - 2 \times 7.133\right)} \Rightarrow 88.11 \text{ pF}$$

$$R_3 = \frac{R_1 \times fz1}{fp1 - fz1} = \frac{2 \times 10^3 \times 3.918}{10.05 - 3.918} \Rightarrow 1.278 \text{ k}\Omega$$

We already know C_1 is 11.16nF and R_1 was selected to be 2 kΩ. So here is a summary of all the components (with the voltage divider component highlighted, to indicate it is an input): **C_1=11.16 nF, C_2= 12.4 nF, C_3= 88.11 pF, R_1= 2 k, R_2 = 3.641 k, R_3 = 1.278 k.**

This baseline corresponds to the central (solid red) gain curve in **Figure 4.37.**

We first ask: how do we lower f_{CROSS}, *only*, without changing the basic location of the poles and zeros. In other words we simply want to translate the red solid curve vertically down. The first step is to double C_1, because R_1 and C_1 determine fp0 (the crossover of the pole-at-origin "p0" as per **Figure 4.38**), and R_1 is preferably fixed since it is part of the voltage divider. However, now looking at the interaction matrix in **Figure 4.38**, we see that C_1 is also part of the second zero "z2". And this doubling of C_1 will no doubt lower fz2. We can see this step #1, the red dashed gain curve in **Figure 4.37**. But that is not what we wanted. So looking again at **Figure 4.38**, we realize that to get fz2 back to where it was, we need to go through step #2: halve R_2.This is the blue dashed line in **Figure 4.37**. Unfortunately, since R_2 was also part of p2 as per **Figure 4.38**, halving R_2 has shifted fp2 to a higher frequency. We need to correct that too. This is done through step #3, where we double C_3. This gives us the solid blue line in **Figure 4.37**, and since C_3 is only part of p2, *the domino effect stops right here*, luckily. Similarly, if we want to raise f_{CROSS}, we can go through the three steps #A, #B and #C shown in **Figure 4.37**.

We can achieve our target of raising or lowering f_{CROSS} without changing the locations of the other poles/zeros, but with a total of three component changes!

It is not as simple as putting a decade box somewhere in the compensator, and blindly tweaking the Bode plot.

Now suppose we want to shift both coincident zeros to half their original frequency, perhaps because we changed the inductor or/and output capacitor to shift the LC pole to half the frequency. Looking at **Figure 4.39** we see that though shifting fz2 seems easy, we are unable to intuitively change fz1, since we get trapped in a strange circle.

Keep in mind that from the locations of the zeros, there is a constraining relationship which we may not have explicitly recognized so far.

$$fz1 = \frac{1}{2\pi(R_1 + R_3)C_2}$$

$$fz2 = \frac{1}{2\pi R_2 C_1}$$

So if fz1=fz2, we have

$$\frac{C_1}{C_2} = \frac{R_1 + R_3}{R_2}$$

This is the constraining relationship inherent in our strategy. Indeed we can confirm by plugging in the numerical values from the book, that this is true. *Further, it needs to be maintained wherever our LC pole is positioned, as per our compensation strategy.* If we do not, our simple strategy will break down, and all bets are off. We may be able to manually tweak crossover frequency and/or phase margin on the bench by using a decade box for one of the resistors involved, as is often done, but at best that would be a minor tweak. In reality as we see below, *many components have to be changed simultaneously*.

As mentioned, halving fz2 is relatively easy. All we need to do, is to double R_2. But since pole "p2" also depends on R_2, to keep it from moving we halve C_3. So that is over, because C_3 is involved only in "p2", not in any other pole or zero location.

Shifting fz1 is however very tricky, and cannot be done intuitively. First, it involves three components: R_1, R_3 and C_2. We don't want to be forced to change R_1, since that is part of the voltage divider. However, if we simply double C_2, this also affects pole "p1" and to keep that unchanged, we need to halve R_3. But R_3 is also involved in the zero "z1", and so the entire process seems convoluted. Luckily, since C_2 also gets multiplied by C_1 in the location of "z1", we do manage to move the location of z1, but by a certain weighted amount, based on the value of R_3.

$$fz1 = \frac{1}{2\pi(R_1 + R_3)C_2}$$

In other words, it is hard to predict what the values of the RC's are for changing fz1. We need to go back to the basic equations for calculating all the components from scratch (mathematics, not intuition). In our specific numerical example, we recalculate all the values if we change fLC (fz1 and fz2) from 3.918 kHz to 3.918/2 = 1.959 kHz. The before and after RC values are:

Before:
C_1=11.16 nF, C_2= 12.4 nF, C_3= 88.11 pF, R_1= 2 k, R_2 = 3.641 k, R_3 = 1.278 k.
After:
C_1=11.16 nF, C_2= 32.7 nF, C_3= 44 pF, R_1= 2 k, R_2 = 7.28 k, R_3 = 484.24.

Our conclusion is, just to shift the two zeros without changing the crossover frequency and the other poles and zeros of a Type 3 compensator, we need *four* component values to be changed every time, and it isn't straightforward either. We certainly can't do it "on the fly". But we can do that with digital techniques, as we will learn in the next part of this series.

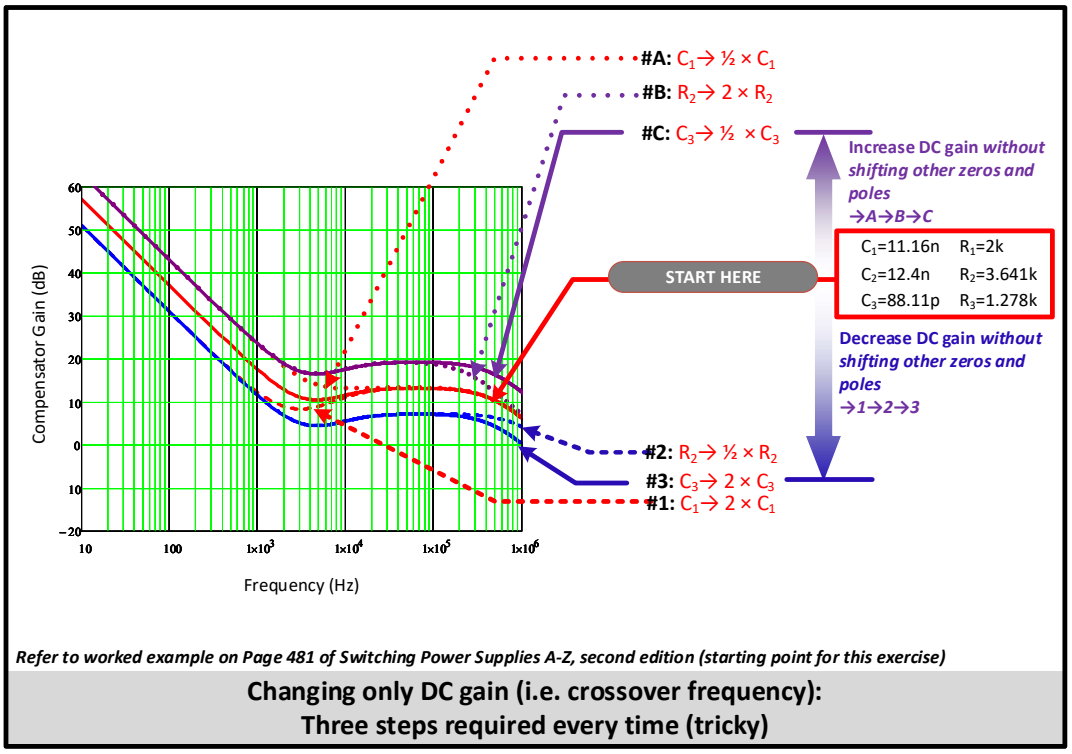


Figure 4.37: Changing crossover frequency (only) using a Type 3 compensator

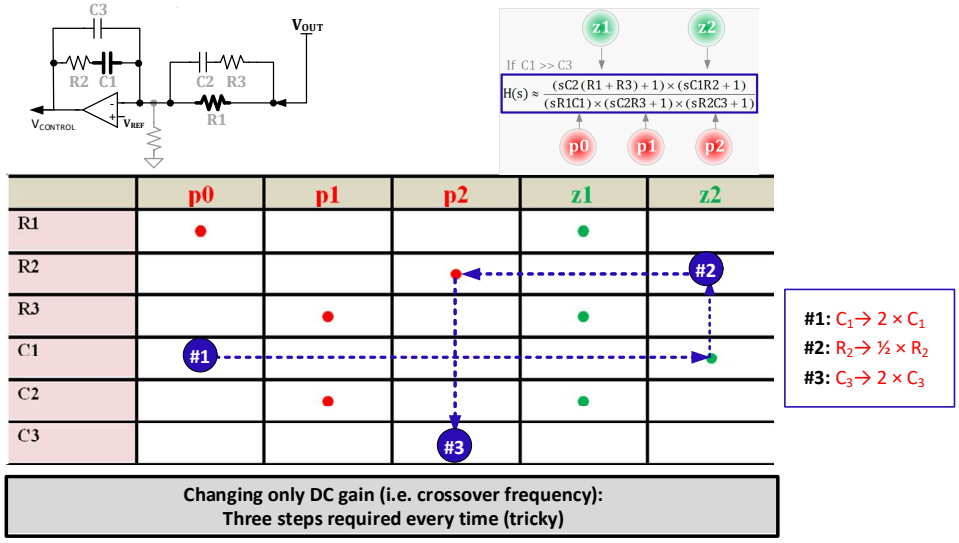


Figure 4.38: How we adjust component values for previous figure

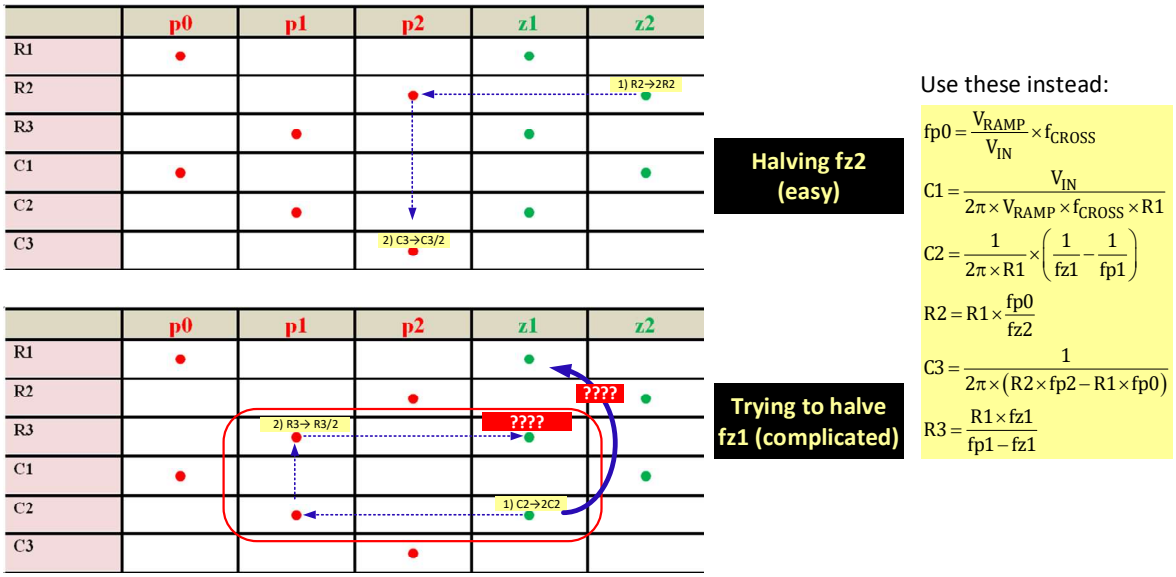


Figure 4.39: Trying to intuitively shift both coincident zeros (only)

Underlying Approximations of Type 3

Unfortunately, the trouble doesn't stop with our inability to tweak a Type 3 compensator easily or intuitively. For if we look at the capacitor values we have calculated for our application, they are not even close to standard values. Capacitors still come mainly in the E12 series: 10 12 15 18 22 27 33 39 47 56 68 82. The tolerance is +,- 10%. On top of that, unless we are using C0G capacitors, we have to include the effects of temperature, voltage, aging and so on. *Not to forget that our initial equations were based on an assumption: $C_1 >> C_3$.*

So the final placement of the poles and zeros, as also the bandwidth (f_{CROSS}) may be quite different from what we intended.

However just in case we want to derive more exact equations, without any approximations, here they are, with and without the $C_1 >> C_3$ approximation. But now we realize that for example, "p0", the pole at origin, is actually affected by *three* components: R_1, C_1 and C_3.

$$fp0 = \frac{1}{2\pi \times R_1 (C_1 + C_3)} \approx \frac{1}{2\pi \times R_1 C_1}$$

$$fp1 = \frac{1}{2\pi \times R_3 C_2}$$

$$fp2 = \frac{1}{2\pi \times R_2 \left(\frac{C_1 C_3}{C_1 + C_3} \right)} = \frac{1}{2\pi \times R_2} \left(\frac{1}{C_1} + \frac{1}{C_3} \right) \approx \frac{1}{2\pi \times R_2 C_3}$$

$$fz1 = \frac{1}{2\pi \times (R_1 + R_3) C_2}$$

$$fz2 = \frac{1}{2\pi \times R_2 C_1}$$

Matters get even more complicated and nothing is intuitive anymore. With that we head to digital control loops next.

Chapter 5
PID and Digital Control Loops

Introduction

Here it certainly gets a bit mathematical, and I (Sanjaya) apologize for that beforehand. But really, we have no natural feel for phase angles in particular. That's why! Can any of us claim to feel comfortable with something like $e^{j\omega t}$? Yet, in some strange way, $e^{j\omega t}$ expresses a phase relationship, which unfortunately, makes it not just necessary, but easier that its grislier alternatives such as solving differential equations. In other words, math is inevitable in feedback systems, and in fact anything in which phase is key.

However, the good news is, after gaining some mastery, by manipulating some common functions in the s-plane, a certain confidence starts to emerge. That has been my approach here— to drive the fear out. I have tried to have fun with math a bit, to get you to the level where you won't get shivers at the very sight of the Laplace transform. Because once you are past that, it becomes remarkably smooth and painless as you'll see.

Towards the end of this chapter, I have gone on to show that a historical "artifact" called conditional stability, is not as harmless as it looks. In fact, it may in some cases be the prime cause for the severe ringing on the output under large-signal events. I have thus introduced a way to lower that ringing by achieving much better matching of the power stage and the feedback section. The Q-matching technique published here is perhaps a stunning analogy… hitherto relatively unnoticed, but at par with my concept of current ripple ratio and also the now suddenly popular scaling lows, presented at an IEEE seminar by Sanjaya, and then a webinar by Nicola.

In the previous chapter, we learned that the basic intent of creating a closed loop (i.e. corrective) system is that by introducing negative feedback we can reduce the effect of any disturbance on the output, be it a line variation, a load step, or a *wiggle in the reference* (however relevant)—by the factor 1/(1+T) for *all* of them, as compared to their effect on the output without a closed loop corrective system in place. Here "T" is in effect the *cumulative* gain of all the cascaded stages of the loop, consisting of the plant (gain = "G") and feedback (gain = "H"). This term "T", may be referred to in literature as the "open-loop gain", or "loop gain", or "round gain" and so on. But whatever it is called, we need to keep in mind that T = G×H. So, the constituent gains are simply multiplied together—or added up on a log scale, since log T = log G + log H. In terms of our shorthand: $T_{dB} = G_{dB}+H_{dB}$. Note that this presupposes *cascaded* gain stages. We also showed that the voltage divider cannot be extricated as a separate gain stage, if we are using conventional error amplifiers in the compensator. Also, nor can the LC post-filter stage be separated out, except for a buck. But by introducing something called the "equivalent inductance", we do manage to do that, even for a boost and a buck-boost.

In principle, we always try to set DC gain as high as possible. Because at least for DC, or very low frequencies, there is no associated phase shift to consider, and for that reason the "loop gain" (overall transfer function, T) has a magnitude, with no imaginary component (i.e. it is a real number). In that simple case the following approximation makes intuitive sense: $1/(1+T) \approx 1/T$. In other words, a high DC gain is very helpful in rejecting DC or low-frequency (e.g. line AC frequency of 50/60 Hz) disturbances in particular, by the factor 1/T.
But we soon learn we can't afford to set a high gain for all the frequencies components, because higher frequencies are almost inevitably accompanied by frequency-dependent phase shifts. We always need to ask: what is the cumulative impact of all these phase shifts? *Does it cause instability*? That condition is defined by the condition T = -1, i.e. a magnitude of unity and a phase of -180°. We thus defined the phase margin as the difference between the phase lag and the ominous threshold of -180°.

Note that phase angles of cascaded gain stages always add up arithmetically, as do gains provided the latter are expressed in logs (db). In other words $\varphi_T = \varphi_G + \varphi_H$. In particular, we need to prevent any frequency component of any disturbance from ever reaching 180° of phase lag by the time it propagates full circle through the plant and the compensator stages. Because if the net phase lag ever reaches that value, then combined with the 180° baseline offset always present on account of *negative* feedback, we will end up with $180+180 = 360 \equiv 0°$ phase lag. Intuitively this means that after going around the loop, the disturbance has returned "in-phase" for the frequency component under discussion. It can wreak havoc, *if the corresponding magnitude is also the same as the starting value (gain =1 or 0 dB)*. Note that this gain equality condition is hard to visualize intuitively, but that is what is implied by T = -1. It is not the same as acoustic feedback for example, where the signal can come around full circle in phase, and with increased magnitude, and cause a huge squeal.

In other words, just a phase lag of 180° does not guarantee instability. It might cause severe ringing on the output, but it cannot sustain itself, and will eventually decay. So at best, phase angle reinforcement is only bad news. One more condition needs to be met to cause full blown instability: *the magnitude of the frequency component on returning full circle needs to also be equal to its starting value* ("0dB"). If these two conditions for phase and gain, are met *simultaneously*, only then the disturbance will become self-sustaining.

To avoid this "doomsday scenario", we need to incorporate a certain robust, safety or stability margin (figuratively the "distance from disaster"). We can talk about this margin, either in terms of "phase margin", i.e. the phase angle short of the 180° phase lag level when the gain falls just below unity (at crossover frequency), or in terms of "gain margin", i.e. the gain below the unity (0dB) level at the frequency where the phase reaches 180° phase lag (if at all, of course).

However, not forgetting that we are dealing with switchers, not continuous control as in older analog systems such as room air-conditioners, we also need to keep in mind that we have to stay well clear of Nyquist's sampling limit of $f_{SW}/2$, where f_{SW} is the switching frequency. This causes an additional frequency-dependent phase lag which can also contribute to system instability. We also need to recognize that the effect of switching can be felt well below $f_{SW}/2$. So we need to stay clear of that by quite a margin! Typically crossing over (0dB) at less than $f_{SW}/5$.

Summarizing: In a practical case, we usually attempt to set a very high DC gain to reduce the effect of disturbances in general. But realizing that frequency-dependent phase shifts will always occur, we strive to ensure that T "crosses over" (i.e. falls below unity, or 0dB axis on a log scale) typically between $f_{SW}/10$ to $f_{SW}/5$, and the phase lag at that point must be well short of -180°.

Occasionally, engineers still ask: don't we try to set current-mode controlled (CMC) systems to achieve even higher bandwidths? Say closer to $f_{SW}/3$? Yes indeed. We may try. But what can happen is this: there is a subharmonic peaking in the gain plot exactly at $f_{SW}/2$, and it can have severe consequences. Let us briefly revisit CMC, to get that over with.

Problems with CMC

Sometimes, in current mode control (CMC), we try to go as high as $f_{SW}/3$ for the crossover frequency. But keep in mind that a sharp peaking in the gain curve appears at exactly $f_{SW}/2$, on account of subharmonic instability *if D > 0.5*. We can see this peak on a Bode plot if we look closely and have high-enough resolution. See **Figure 5.1.**

We discover that if this "parasitic peak" rises upwards and ever intersects the 0dB line, it will cause the system to go into unrecoverable instability, with no understandable "Bode plot" thereafter. We will see "alternate pulsing" on the switch mode (one big pulse, followed by a narrow pulse, repeating indefinitely). The transient response will be as bad as it gets, even though under steady-state conditions, we may not notice any difference.

Solutions to that are lowering the "Q" of this subharmonic peak—by steps such as adding more slope compensation, increasing the inductance, or simply lowering the crossover frequency to even lower values than the supposed max $f_{SW}/5$ limit of voltage-mode control (VMC)! This will then increase the "safety margin" requirement of CMC as shown in **Figure 5.1.**

To lower the Q to a reasonable value *of less than 2 (maybe even to 1 or 0.5)*, we need to ensure a certain minimum inductance as shown in **Figure 5.2**. But increasing the inductance causes a new problem, as shown in **Figure 5.3**. This is the issue related to the *leading edge spike*. It can cause jitter and in severe cases, an inability to deliver full power.

Suppose we increase the inductance to evade this, we could cause premature termination of the switching pulses. Because in doing so, we are inadvertently raising the pedestal on which the spike is riding on. And since, less than the required energy will be delivered for that (prematurely terminated) cycle, in the next cycle the converter will try to compensate by a larger duty cycle. In this process it gets some unexpected help because after the early termination of the previous pulse, the inductor current had a longer time to slew down, and thus the pedestal on which the leading-edge spike is riding comes down, usually enough to help it evade early pulse-limiting in the next cycle.

What we see on the oscilloscope are alternate wide and narrow pulses, *which mimic what we get under subharmonic instability*, as mentioned above.

We are surprised, because we had thought that a high inductance should be helping us avoid subharmonic instability, but here it seems to be aggravating it!

We could of course set a large 'blanking time' for current mode control, and/or we can add some delay to the current limit detect circuit. But we also then run the danger of not being able to react fast enough to an actual abnormal load condition, especially if the inductor starts to saturate. The transient response will also worsen, since as mentioned, any delay in sampling is equivalent to a frequency-dependent phase lag. If we add slope compensation indiscriminately, we are in effect converting CMC to VMC, and the LC peaking will once again start showing up in the Bode plot!

And so on. We go around in circles. No wonder, VMC with input feedforward is becoming the preferred choice nowadays. We therefore ignore CMC from this point on.

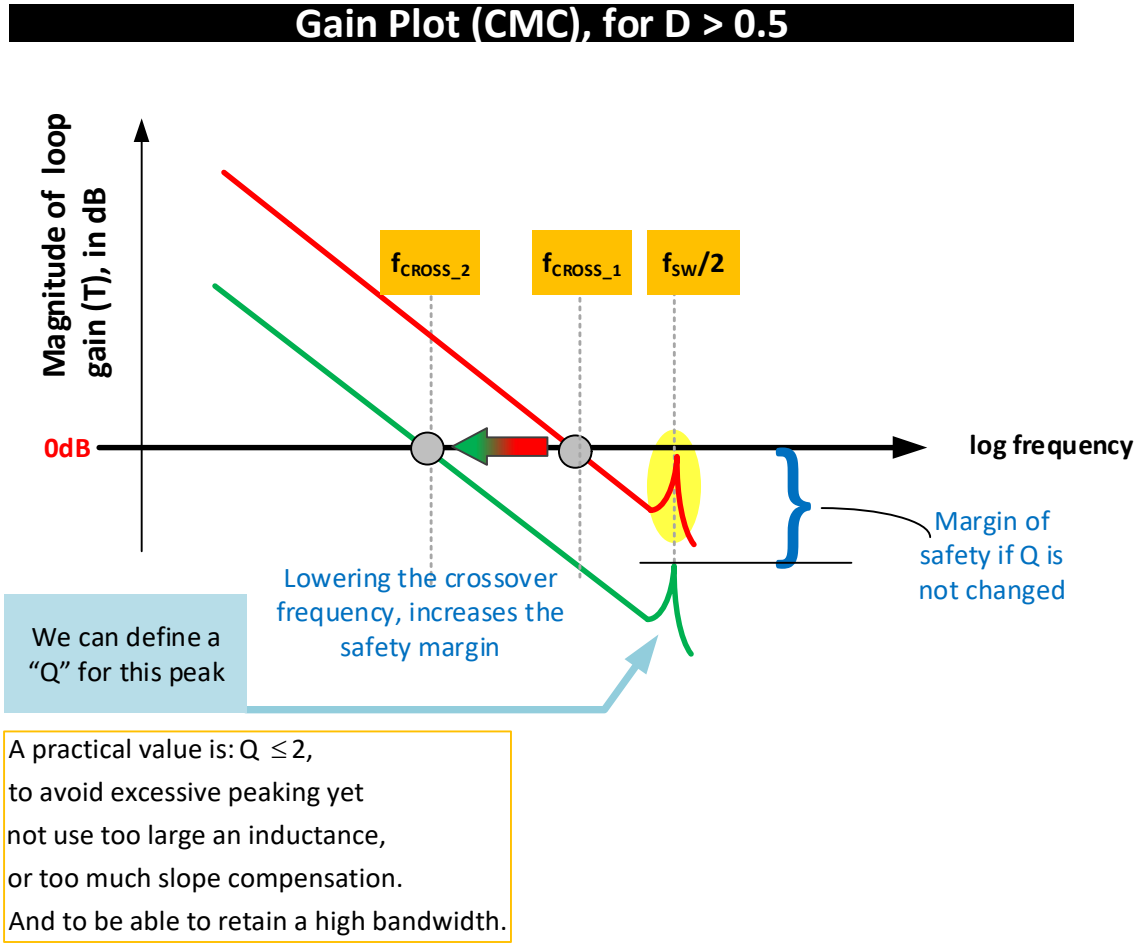
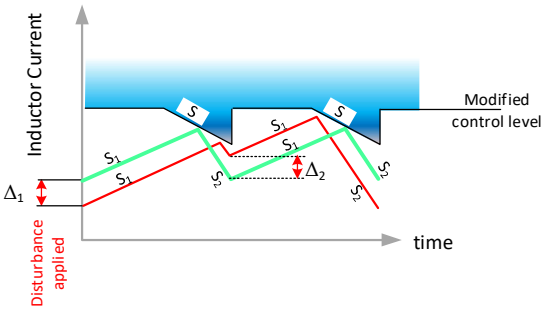


Figure 5.1: Reducing crossover increases the safety margin in CMC

How slope compensation (S) causes any disturbance in the inductor current waveform, to converge ($\Delta_2 < \Delta_1$)



Note that in related literature:

a) the slope compensation we call "S" is often called "S_e"

b) the up-slope we call "S_1" is often called "S_n"

c) the down-slope we call "S_2" is often called "S_f".

If L is in µH, S is in A/µs & recommended $Q \leq 2$, we get

$$L_{\mu H} \geq V_{IN} \times \frac{(D-0.34)}{SlopeComp_{A/\mu s}} \quad \text{(Buck)}$$

$$L_{\mu H} \geq V_O \times \frac{(D-0.34)}{SlopeComp_{A/\mu s}} \quad \text{(Boost)}$$

$$L_{\mu H} \geq (V_{IN}+V_O) \times \frac{(D-0.34)}{SlopeComp_{A/\mu s}} \quad \text{(Buck-Boost)}$$

Q (quality factor of the pole at half switching frequency) is defined for all topologies as:

$$Q = \frac{1}{\pi[m_c D' - 0.5]}$$

where $m_c = 1 + \frac{S_e}{S_n}$; D'=1-D

We can simplify this to

$$S \times L = \frac{V_{ON}}{1-D} \times \left(\frac{1}{\pi Q} + D - 0.5\right)$$

Note: V_{ON} is voltage across L during T_{ON}.

In steady state:

$V_{ON} / (1-D) = V_{OFF} / D$.

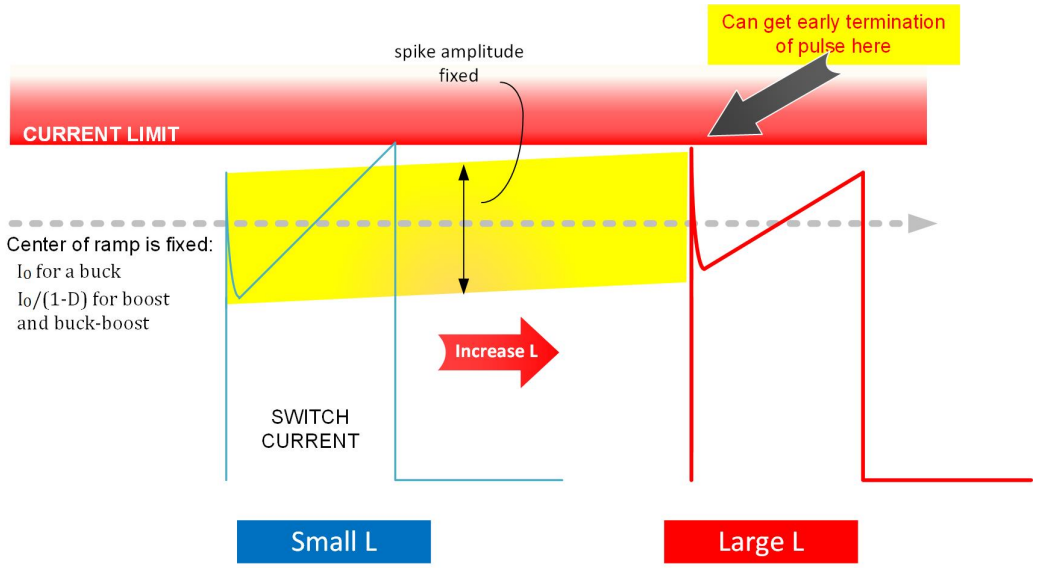
We can use the equations for V_{ON} and D for each topology to get:

$$S \times L = V_{IN} \times \left(\frac{1}{\pi Q} + D - 0.5\right) \quad \text{(Buck)}$$

$$S \times L = V_O \times \left(\frac{1}{\pi Q} + D - 0.5\right) \quad \text{(Boost)}$$

$$S \times L = (V_{IN}+V_O) \times \left(\frac{1}{\pi Q} + D - 0.5\right) \quad \text{(Buck-Boost)}$$

Figure 5.2: Minimum inductance related to a given slope compensation, to achieve $Q \leq 2$



One major concern is the leading edge spike. This can cause jitter and in severe cases, a consequent inability to deliver full power. If we increase the inductance we could cause premature termination of the switching pulses in current mode controllers, because in doing so, we are inadvertently raising the pedestal on which the spike is riding on. But since less than the required energy will be delivered for that (prematurely terminated) cycle, in the next cycle the converter will try to compensate by a larger duty cycle. In this process it gets some unexpected help, because after the early termination of the previous pulse, the inductor current had a longer time to slew down, and thus the pedestal on which the leading edge spike is riding comes down, probably enough to help it evade early pulse limiting in the next cycle. Finally, what we may see on the oscilloscope are *alternate wide and narrow pulses*, which mimic what we get under genuine subharmonic instability. We may be surprised, because we always thought that a high inductance should be helping us avoid subharmonic instability, but here it seems to be aggravating it. The leading edge spike also causes erratic responses from the current limit protection circuit for both current mode and voltage mode control. We can't set an effective current limit based on a spike, especially since we will discover that this spike varies from unit to unit, being based on uncontrolled and/or uncharacterized parasitics. We could of course set a large 'blanking time' for current mode control, and/or we can add some delay to the current limit detect circuit. But we also then run the danger of not being able to react fast enough to an actual abnormal load condition, especially if the inductor starts to saturate. So we may end up oversizing our inductor purely because we started with a burning desire to somehow decrease current ripple ratio 'r'.

Figure 5.3: How increasing the inductance can cause alternate pulsing too

Contribution of Plant and Feedback Blocks

As we saw in **Chapter 4**, eventually in a closed loop of cascaded blocks, it really does not matter whether a block is considered within the plant or within the feedback stage. And that is essentially what is also implied by writing T = G×H, which could instead be written as T = H × G, or in general as T = $\mathbf{\Pi}$ G_i. Any disturbance is attenuated by the correction factor 1/(1+T), where T = $\mathbf{\Pi}$ G_i compared to its effect on the output if closed-loop feedback were not present! The only critical element of any practical control loop is negative feedback—for creating corrective action. The other stages can be viewed as "democratic" constituents of the closed loop system from the viewpoint of the disturbance. It doesn't matter where the gain or phase contribution is coming from. We are only concerned with the net loop gain, i.e. "T" (with all dB's of cascaded stages added logarithmically and phases added arithmetically). That is why in the previous chapter we had suggested not even bothering to allocate different symbols for the plant and feedback, such as G and H, or H and G.

However, in this part we are still retaining the distinction between G and H, just for the purpose of descriptive clarity. So, here G is still the plant and H the feedback block. As mentioned, in literature, some reverse even that. Beware!

Note that phase shifts occur in both the blocks, G and H. Unfortunately, the plant gain/phase profile, (G), is largely out of our control. On the other hand, the compensator (feedback block, H) is almost completely in our hands. That is the only major difference between the two from our viewpoint, in a closed loop system. We can adjust H to literally "compensate" for, or "tune out", any undesirable aspects of the plant gain/phase H. Hence this topic of study is often called "loop *compensation*".

The first "weakness" of the plant profile G, in terms of creating a desirable T profile, is that it has a low DC gain—almost flat initially till a certain cutoff frequency determined by the resonance of the inductance and output capacitor (LC). We can ask: what exactly is the low-frequency/DC (flat) value of G? There are three simple equations to answer that, for each of the three basic topologies. But we will focus only the buck regulator here. We can approximate that equation further if the DC resistance of the inductor (DCR), the equivalent series resistance (ESR) of the output capacitor are both almost zero, as follows:

$$G(s) = \frac{V_{IN}}{V_{RAMP}} \times \frac{1}{\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q}\left(\frac{s}{\omega_0}\right) + 1}$$

The gain contribution from the PWM comparator stage (part of G) is $1/V_{RAMP}$. Similarly, the power (switching buck) stage (also part of G) provides the V_{IN} term. The LC post filter (also part of G) provides the rest of the plant gain above, i.e. the *frequency-dependent* part (involving s).

There are many ways the above equation is written out in literature actually.

$$G_{LC}(s) = \frac{1/LC}{s^2 + s\left(1/RC\right) + 1/LC}$$

$$G_{LC}(s) = K\frac{1}{s^2 + \frac{\omega_0}{Q}s + {\omega_0}^2}$$

$$G_{LC}(s) = \frac{1}{\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q}\left(\frac{s}{\omega_0}\right) + 1}$$

$$G_{LC}(s) = \frac{1}{1 + 2\zeta\left(\frac{s}{\omega_0}\right) + \left(\frac{s}{\omega_0}\right)^2}$$

where $\omega_0 = 1/(LC)^{1/2}$. Note that above, $K = \omega_0^2$. Q is the quality factor. We have several forms: $Q = (R/L)/\omega_0$ or $Q = R \times (C/L)^{1/2}$, where R is the load resistor in the case of a switcher.

Alternatively, we can express this transfer function in terms of ζ, the damping factor, where $\zeta=1/2Q$.

Note: *The LC post filter has no DC gain (0 dB). Any DC gain in the plant comes from the V_{IN}/V_{RAMP} term, coming from the two other constituent stages. So to avoid confusion, it is best to use only the last two forms above, which try to express the gain in terms of (s/ω_0). It then becomes clear that the term is not causing any DC gain. Or if it is!*

In other words, stick to this form:

$$G_{LC}(s)=\frac{1}{\left(\frac{s}{\omega_0}\right)^2+\frac{1}{Q}\left(\frac{s}{\omega_0}\right)+1}$$

Or for the plant (combining the post filter with the comparator and switch), this is the best way to write it out:

$$G(s)=\frac{V_{IN}}{V_{RAMP}}\times\frac{1}{\left(\frac{s}{\omega_0}\right)^2+\frac{1}{Q}\left(\frac{s}{\omega_0}\right)+1}$$

So for example, in our case, we can now clearly see that the only contribution to the DC gain value of G above comes from the term V_{IN}/V_{RAMP}. If the PWM ramp is 1V, and the input voltage is 12V, then the DC gain contribution from the plant, *irrespective of the output voltage* is 12/1 = 12. In terms of decibels, this is 21.5 dB. This amount of DC gain is usually insufficient to reject low-frequency ripple or transients, as described in Part 1, and we therefore we have to use the feedback network to somehow increase the product G×H = T fairly dramatically. The circuit to do that is the integrator, as also discussed in **Chapter 4**.

At this point, we describe the behavior of the LC post filter, and provide more detailed forms of the LC filter's transfer function, now involving parasitics.

LC Post Filter Analysis

We are again going to look at this through our numerical example in Chapter 4 (also in A to Z, Second Edition)

Example: *Using a 300 kHz synchronous buck controller we wish to step-down 15V to 1V. The load resistor is 0.2Ω (5A). The PWM ramp is 2.14V as per the datasheet of the part. The selected inductor is 5μH, and the output capacitor is 330μF, with an ESR of 48mΩ.*

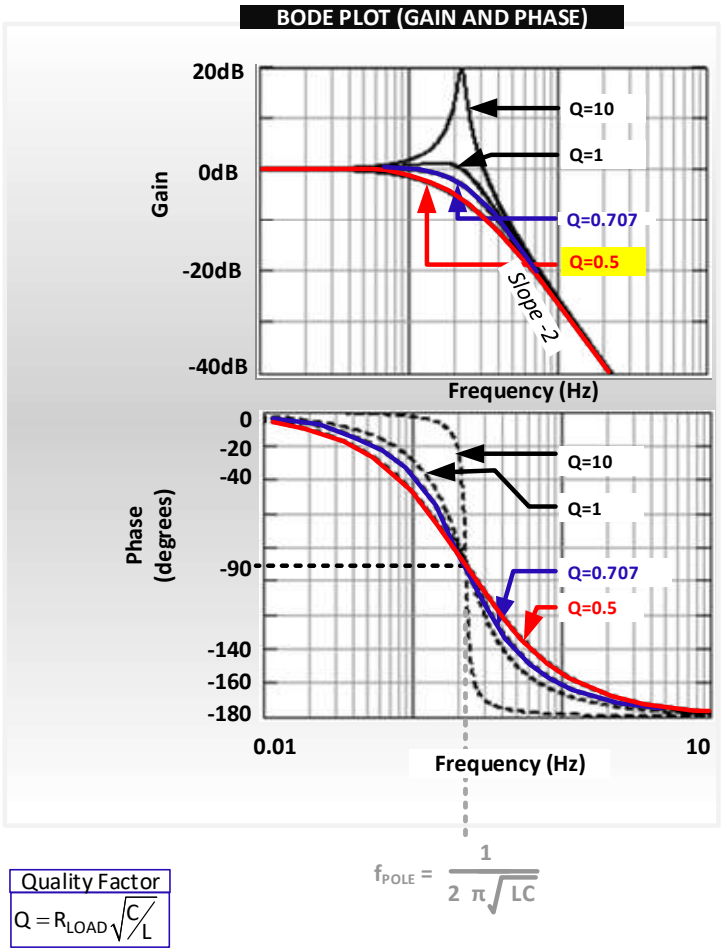


Figure 5.4: LC post-filter response as load is varied, assuming zero DCR and zero ESR

To start with we ignore ESR too. As we vary the load we get the filter response shown in **Figure 5.4**. Note the extremely abrupt phase shift of 180° at the break (resonant) frequency f_{LC} and the peaking in gain, especially as we *increase the load resistor* (increase Q). The load resistor is the only dissipative term so far which is damping out the LC resonance. So if it were not present, we would tend towards the following approximation.

$$G(s) = \frac{V_{IN}}{V_{RAMP}} \times \frac{1}{\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q}\left(\frac{s}{\omega_0}\right) + 1} \approx \frac{V_{IN}}{V_{RAMP}} \times \frac{1}{\left(\frac{s}{\omega_0}\right)^2 + 1}$$

We reach a conclusion here: *The resonant frequency is determined by the coefficient of the term involving "s^2". The damping is determined by the coefficient of the term involving "s".*
In general, if we arrive at a second order equation of the type:

$$A(s)^2 + B(s) + 1$$

We can conclude that the resonant frequency and Q are as follows

$$\omega_0 = \frac{1}{\sqrt{A}}; Q = \frac{\sqrt{A}}{B}$$

If A is unity, as it usually is, then remember this: *Q is the reciprocal of the coefficient of the s-term.*

Now let us add some ESR to this, *but still at max load*. The resultant gain function is shown in **Figure 5.5**. It introduces an ESR-zero, and a good approximation of that is

$$G(s) = \frac{V_{IN}}{V_{RAMP}} \times \frac{\left(\frac{s}{\omega_{ESR}} + 1\right)}{\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q}\left(\frac{s}{\omega_0}\right) + 1}$$

Now let us set the ESR back to zero, and instead add DCR (of inductor) to the transfer function. This is plotted out in **Figure 5.6** for max load.

The detailed transfer function with no approximations (ESR and DCR included) is:

$$G_{LC}(s) = \frac{\left(\frac{s}{\omega_{ESR}} + 1\right)}{\left(\frac{s}{\omega_0}\right)^2 \left(1 + \frac{ESR}{R}\right) + s\left[\frac{L}{R} + \left\{DCR \times C \times \left(1 + \frac{ESR}{R}\right)\right\} + (ESR \times C)\right] + \left(1 + \frac{DCR}{R}\right)}$$

where $\omega_0 = 1/(LC)^{1/2}$ as before, R is the load resistor and C is the output capacitor with some ESR. L is the inductor with a certain DCR.

It is difficult to write a simple form for Q above.

We make the following conclusions:

a) There is an ESR-zero at the angular frequency $\omega_{ESR} = 1/(ESR \times C)$. So $f_{ESR} = 1/2\pi(ESR \times C)$.
b) The ESR starts dramatically changing the -2 slope of the LC double pole, closer to a -1 slope, as it approaches the LC break-point frequency from the right.
c) The ESR affects the break-point frequency, since it appears in the term involving s^2. The DCR does so too, but only slightly through the term (1+DCR/R) at the end.
d) Both the ESR and the DCR affect the Q as expected.

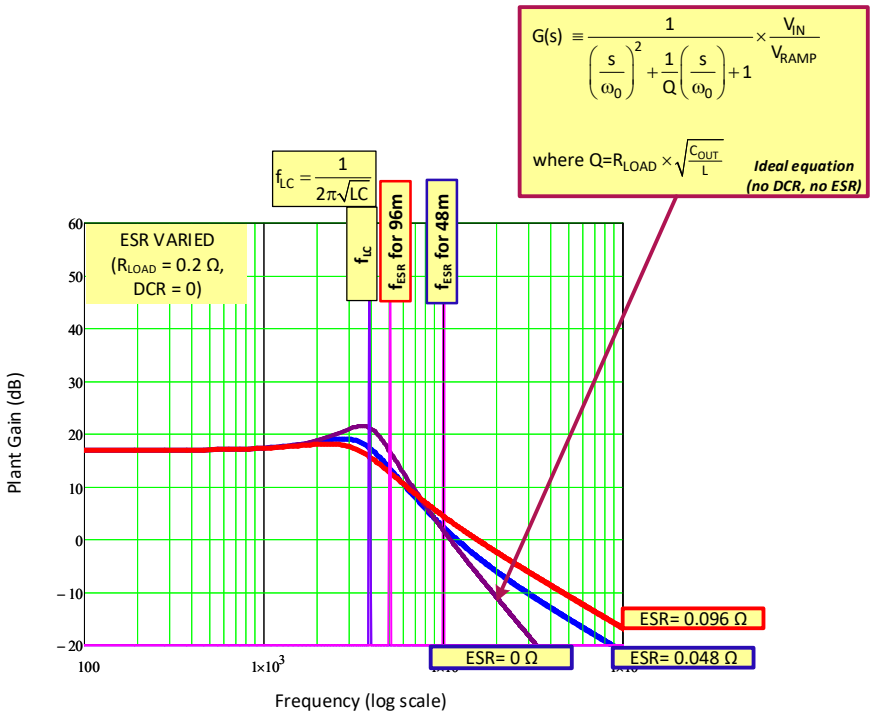


Figure 5.5: LC post-filter response as ESR of cap is varied, assuming zero DCR and MAX LOAD

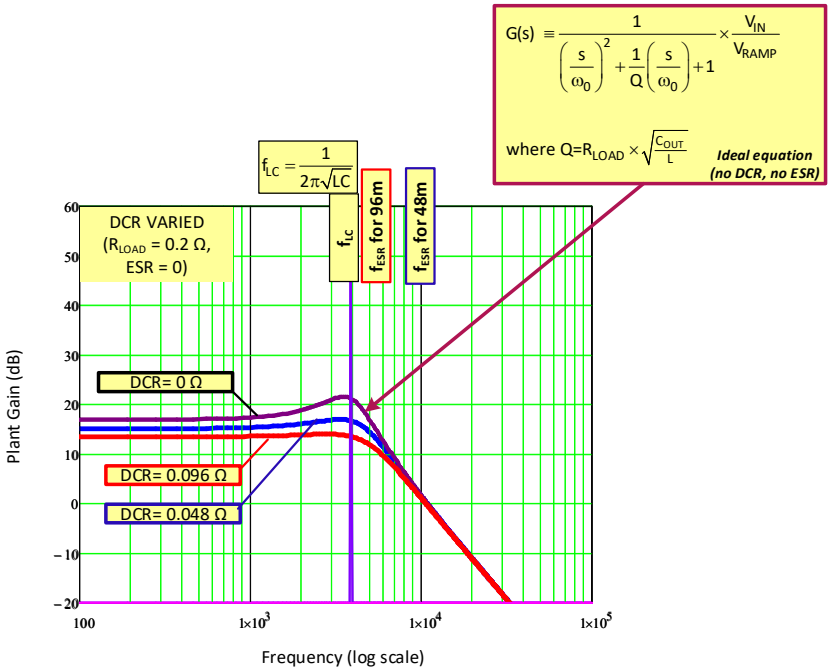


Figure 5.6: LC post-filter response as DCR is varied, assuming zero ESR and MAX LOAD

Developing Intuition

Before we go any further with our analysis, we pause. The reason for the pause is we still need to develop a certain "intuition" to succeed. We need to understand how the frequency domain (s-plane) came about in the first place, and how the Laplace transform can help us. *And how it sometimes can't!*

Intuition in control loop theory is like the "taste of coffee". You need to acquire it, since it is not necessarily natural. Feedback in general, is definitely an area where prior intuition can actually be misleading. Because for one, we are not intuitively cognizant of phase angles. That is perhaps why, even the "basic" phenomenon of interference fringes from light beams, as demonstrated by Thomas Young, came as late as 1801, and was still not accepted for years thereafter by his peers. But it is that experiment which underlined the importance of phase angle, as in feedback theory subsequently.

Our approach here is in a sense, to put the cart before the horse, and first understand the behavior of some of the key underlying transfer (gain) functions in the "s-plane". The s-plane with all its real and imaginary, positive and negative frequencies is also, initially at least, unintuitive. The good news is, after we feel comfortable navigating the s-plane, we will hopefully develop a certain form of *acquired* intuition for control loop theory, which will see us through the complexities. On to digital loops!

Logarithms are Natural

Note that while plotting the gain-phase curves, we may have discovered it was completely equivalent whether we took the variable being plotted (gain or frequency), and used a log scale, or took the log of the variable beforehand, and then plotted it on a linear scale. We remind ourselves here that a 10× variation is the same as 20 dB. Similarly, a (1/10)× variation is the same as -20dB. Similarly, a 10 dB variation is almost the same as a 3× variation, since 20 × log(3) ≈ 10 dB. Note that no one really talks of frequency in terms of decibels, so we have also refrained from that here. Instead, we are using a log scale for frequency in our figures. However keep in mind that <u>if</u> we had expressed frequency in decibels in the same manner as for gain (i.e. as 20×log f), we would have realized that for *both* gain and for frequency, a "20 dB" shift is equivalent to a 10× variation, which is more commonly referred to as a "decade" shift when we talk specifically about frequency instead of gain. In other words, intuitively, "20 dB/decade" or slope of "1" is actually the same as 20 dB/20dB. Which in terms of tan(θ) is literally a slope of 1 (i.e. tan 45°, if the x and y-axis are equally proportioned). That is why +20dB/decade is often called a "+1" slope. Similarly, -20dB/decade is a -1 slope. And -40dB/decade is a "-2" slope.

In other words, a typical slope of "-1", i.e. -20 dB/decade, simply means that if the gain changes by a factor of 10, so does the frequency, by the very *same* factor!

But isn't that *inverse proportionality*?

In general, a variation in gain by any arbitrary factor "Z", corresponds to a variation in the frequency by the same factor "Z"— for a line with slope "-1". For a line with slope "-2", as for the plant gain above the LC double pole corner frequency f_{LC}, a variation of -40dB/decade, or "-2" slope, simply means that if the frequency increases by a factor "Z" (say 2×), the gain falls by a factor Z^2 (4×). This is $\propto 1/f^2$. And so on.

Another way of expressing slope is to talk in terms of *octaves* of shift in frequency, instead of decades. An octave is simply a doubling of frequency. But since 20 × log (2) = 6 dB, some engineers prefer to say that a slope of "-1" is the same as -6 dB/octave instead of -20 dB/dec. Which is also just -6 dB/6dB, or -1. Same as -20 dB/20dB, and so on.

We realize that what all this implies is:

a) If the gain is inversely proportional to frequency, i.e. gain $\propto 1/f$, we get a slope of -1 (-20 dB/dec) on a log vs log scale. This usually comes from one reactive element (such as an RC combination).
b) If the gain is inversely proportional to frequency^2, i.e. gain $\propto 1/f^2$, we get a slope of -2 (-40 dB/dec) on a log vs log scale. This usually comes from two reactive elements (such as the LC double pole in VMC).
c) If the gain is proportional to frequency, i.e. gain $\propto f$, we get a slope of +1 (+20 dB/dec) on a log vs log scale.
d) If the gain is proportional to frequency^2, i.e. gain $\propto f^2$, we get a slope of +2 (+40 dB/dec) on a log vs log scale.

Logarithms are simply the most natural form of progressions in our world. Because they reflect constant ratios! They are essentially "geometrical progressions", the same as exponential functions.

Example: *Consider 10000 power supplies in the field with a failure rate of 10% every year. That means in 2010 if we had 10000 working units, in 2011 we would have 10000 $\times$ 0.9= 9000 units. In 2012 we would have 9000 $\times$ 0.9= 8100 units left. In 2013 we would have 7290 units left, in 2014, 6561 units, and so on. If we plot these points --- 10000, 9000, 8100, 7290, 6561 and so on, versus time, we will get the well-known decaying exponential function. See* ***Figure 5.7*** *on the left side. We have plotted the same curve twice: the curve on the right has a log scale on the vertical axis. Note how it now looks like a straight line. It cannot however ever go down to zero!*

The simplest and most obvious initial assumption of a *constant* failure rate has led to an *exponential* curve. That is because the exponential curve is simply a succession of evenly spaced data points (very close to each other), which are in *geometric progression*—i.e. the ratio of any point to its preceding point is a constant number. So if x is the horizontal axis, we get the function on the vertical axis as $y(x) = a \times a \times a \ldots$(multiplied x times), i.e. $y(x) = a^x$. Most natural processes behave similarly. Such as radioactive decay (half-life), population etc. Note that all curves based on geometrical progressions may see "exponential" to us, but to be accurate, the truly exponential curve is the one where the ratio of the geometric progression is such that the slope of the function a^x at any point equals the function itself. In other words

$$\frac{d(a^x)}{dx} = a^x \text{ only if } a = 2.72 \quad (\text{i.e. "e"})$$

This property of "e" provides a huge simplification when trying to solve differential equations, which is why "e" is so ubiquitous.

We also recall that a logarithmic scale is simply a different scaled version of the exponential scale, so they possess the same qualities. That is why if the log of any number is multiplied by 2.303, we get its natural log ("ln"). Conversely, if we divide the natural log by 2.303 we get its log. This follows from

$$\ln(10) = 2.303 \text{ and } \frac{1}{\log(e)} = 2.303$$

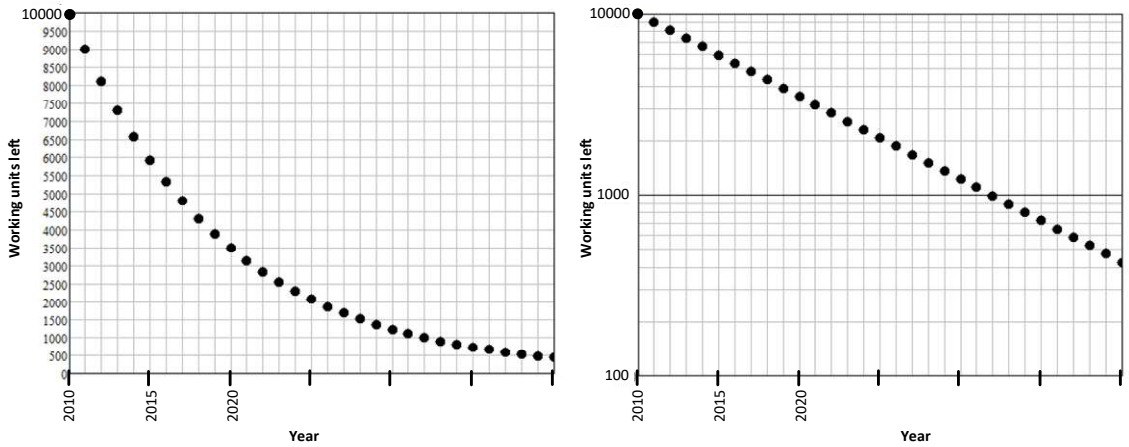


Figure 5.7: How geometrical progressions appear as straight lines on logarithmic scales

Relating this to our relationship with nature, we perceive loudness and brightness close to logarithmic. We tend to perceive decibels, not factors. That way our senses can handle a very wide range of sight and sound amplitudes, by "squeezing them together"—almost logarithmically.

See the conversion tables for logs/decibels and factors provided in **Figure 5.8**.

The other great thing about using logs as mentioned previously is that if T = GH, we get log |T|= log |G| + log |H|. We have simply written this out in shorthand as $T_{dB} = G_{dB} + H_{dB}$, implying that we can arithmetically sum the decibels, since it is now all logarithmic.
Next, we show the development of the Laplace transform.

From Switching Power Supplies Design and Optimization, 2e

Factor	20 × log (factor)
× 1	0 dB
× 1.5	3.5 dB
× 2	6 dB
× 3 (2 × 1.5)	9.5 dB (6 + 3.5)
× 4 (2 × 2)	12 dB (6 + 6)
× 5 (10/2)	14 dB (20 – 6)
× 6 (3 × 2)	15.5 dB (9.5 + 6)
× 7	17 dB
× 8 (4 × 2)	18 dB (12 + 6)
× 9 (3 × 3)	19 dB (9.5 + 9.5)
× 10 (2 × 5)	20 dB (6 + 14)

dB	Ratio	Easier to remember (Ratio)
1	1.122	
2 (= 12 – 10)	1.265 (= $4/\sqrt{10}$)	
3	1.414 (= $\sqrt{2}$)	$\sqrt{2}$
4 [= (20 – 12)/2]	1.581 [= $(10/4)^{1/2}$]	
5 (= 10/2)	1.778 [(= $\sqrt{10})^{1/2}$]	
6	2	2
7	2.24 (= $\sqrt{5}$)	
8 (= 20 – 12)	2.5 (= 10/4)	2.5
9 (= 6 + 3)	2.828 (= $\sqrt{8}$) (= 2 × $\sqrt{2}$)	$\sqrt{8}$
10 (= 20/2)	3.17 (= $\sqrt{10}$)	$\sqrt{10}$
11 (= 8 + 3)	3.536 (= 2.5 × $\sqrt{2}$)	
12 (= 6 × 2)	4 (= 2^2)	4
13 (= 10 + 3)	4.472 (= $\sqrt{10} \times \sqrt{2} = \sqrt{20}$)	
14 (= 7 × 2)	5 (= $\sqrt{5^2}$)	5
15 (= 12 + 3)	5.657 (= 4 × $\sqrt{2}$)	
16 (= 8 × 2)	6.25 (= 2.5^2)	
…	…	…
20 (= 10 + 10)	10	10

Figure 5.8: Converting from factors to decibels and vice versa

Fourier Series to Laplace Transform

Let us start with what we learned in high school. We broke up a repetitive waveform into discrete harmonic frequency components, analyzed the effect of each harmonic separately, then summed them all up to reconstruct the result. The decomposition was in terms of the fundamental frequency ω_0. In addition there was a DC offset expressed by $a_0/2$ (sometimes called just a_0 in literature).

The tricky thing was to apply that to switchers. We have to go from angles expressed in (dimensionless) radians to time and frequency. Because sine and cosines can't be applied directly to time. Time is not dimensionless. However, the conversions to use are indicated in **Figure 5.9**. The key is the following transformation

$$\frac{\theta}{2\pi} \leftrightarrow \frac{t}{T} \text{ , so } \theta \leftrightarrow \frac{2\pi t}{T}$$

We can also write this as

$$\theta \leftrightarrow 2\pi f \times t \equiv \omega t$$

With this, we get for a time-repetitive waveform, as used in switchers:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n \times \omega_0 t) + \sum_{n=1}^{\infty} b_n \sin(n \times \omega_0 t)$$

The amplitude of each harmonic is

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{-\pi} f(t)dt$$

$$a_n = \frac{\omega_0}{\pi} \int_{0}^{T} f(t)\cos(n \times \omega_0 t)dt$$

$$b_n = \frac{\omega_0}{\pi} \int_{0}^{T} f(t)\sin(n \times \omega_0 t)dt$$

In effect, we are going between time into a "frequency domain", except that *so far this frequency domain is simply a discrete array of frequencies*, separated by something we call the fundamental frequency.

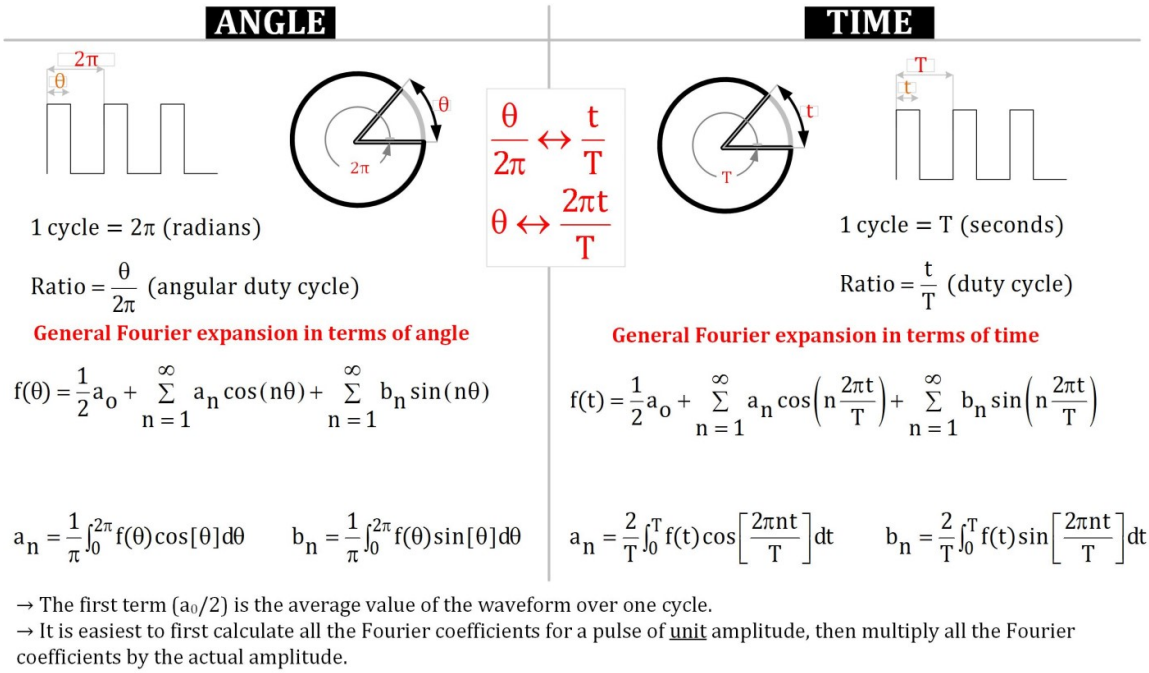


Figure 5.9: Going between angles and time in Fourier series for example

This technique was further developed into the *complex Fourier series*, by invoking the exponential function. But it was still just a mathematical construct to simplify computations. And it was based on the following well-known relationships

$$e^{j\theta} = \cos(\theta) + j\sin(\theta)$$
$$e^{-j\theta} = \cos(\theta) - j\sin(\theta)$$

$$\sin(\theta) = \frac{e^{j\theta} - e^{-j\theta}}{2j}$$
$$\cos(\theta) = \frac{e^{j\theta} + e^{-j\theta}}{2}$$

Note that in standard electrical analysis, we set F9 = F12t as mentioned.

As an example, using the above equations, we can see derive the magnitude and phase of the exponential function $f(\theta) = e^{j\theta}$ as follows

$$\text{Magnitude}(e^{j\theta}) = \sqrt{\cos(\theta)^2 + \sin(\theta)^2} = 1$$
$$\text{Argument}(e^{j\theta}) = \tan^{-1}\left(\frac{\sin(\theta)}{\cos(\theta)}\right) = \tan^{-1}\tan(\theta) = \theta$$

So if for example we started with the following Fourier series

$$5\cos\theta + 12\sin\theta$$

Using the exponential function we get

$$5\cos\theta + 12\sin\theta = (2.5 + 6j)e^{-j\theta} + (2.5 - 6j)e^{j\theta}$$

So now we are operating with real and imaginary harmonic amplitudes. It is just a mathematical construct though.This is called the "complex Fourier series".

In general:

$$f(\theta) = \sum_{-\infty}^{\infty} c_n e^{jn\theta}$$

All the frequencies are a discrete interval apart, and can be both negative and positive.

In our case, we replace θ with $\omega_0 t$. We can then solve for c_n as follows

$$c_n = \frac{\omega_0}{2\pi}\int_0^T f(t)e^{-jn\omega_0 t}dt$$, or equivalently $$c_n = \frac{1}{T}\int_0^T f(t)e^{-j\frac{2\pi nt}{T}}dt$$

This is how we can go back and forth between the domain involving time (t) and frequency as represented by the discrete array $n\omega_0$.

There is no smooth one-dimensional, or two-dimensional spread of frequencies yet! That comes next.

Historically, the next step was to try the same decomposition technique with non-repetitive waveforms. This led to the Fourier transform method. Here the summation over discrete frequencies changed to a smooth integration. The decomposition was as follows

$$f(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}F(\omega)e^{j\omega t}d\omega$$

So, instead of $n\omega_0$, we now have a "smooth" (continuous) variable ω.

The "amplitude of each harmonic" in a sense, is now $F(\omega)$, taking the place of c_n in the Fourier series. It is called the Fourier transform. With it we can study the effect of each frequency component ω. We also define $s = j\omega$, and "s" is then the frequency plane or domain. *But it is actually not a plane yet, just a one-dimensional continuous spread. In terms of the modern two-dimensional s-plane with orthogonal real and imaginary axes, we can think of the Fourier transform as operating only along the imaginary axis of this "plane".*

The Fourier transform ("amplitude of harmonic") as a function of $s = j\omega$, is

$$F(s)=\int_{-\infty}^{\infty}e^{-st}f(t)dt$$

Note that alternatively, we could have calculated this using

$$F(\omega)=\int_{-\infty}^{\infty}e^{-j\omega t}f(t)dt\equiv\int_{-\infty}^{\infty}(\cos\omega t)f(t)dt-j\int_{-\infty}^{\infty}(\sin\omega t)f(t)dt$$

In other words, the real part of the Fourier transform is found by multiplying the time domain signal by the cosine wave ("harmonic") and then integrating (from -∞ to +∞). The complex part is found by using the sine wave instead above.

However there are many functions (applied disturbances), such as a ***step function***, which do not give us a finite number for F(s) if we integrate it directly from -∞ to +∞ as in a Fourier transform As a result, a new decomposition technique was introduced, called the Laplace transform.

It attempted to add an exponentially decaying part to the integral, to force convergence. In effect, we were doing this (where σ is a real number):

$$F(s)=\int_{0}^{\infty}e^{-st}f(t)dt=\int_{0}^{\infty}e^{-\sigma t}\times e^{j\omega t}\times f(t)dt\Rightarrow \text{Integral}\left[(\text{exponential enevelope})\times(\text{oscillatory})\times(\text{function})\right]$$

$$F(\omega)=\int_{0}^{\infty}e^{-\sigma t}e^{-j\omega t}dt$$

$$F(\omega)=\int_{0}^{\infty}f(t)\,e^{-\sigma t}e^{-j\omega t}dt=\int_{0}^{\infty}f(t)\,e^{-(\sigma+j\omega)t}dt$$

$$F(s)=\int_{0}^{\infty}f(t)\,e^{-st}dt\quad\text{where } s=(\sigma+j\omega)$$

This is actually still the Fourier transform if σ = 0. The limits of integration have also changed a bit, because in Laplace transform analysis, we typically always assume that there was no signal/impulse/disturbance prior to t = 0.

The Laplace transform is given a new symbol **L.**

$$L\{f(t)\} = F(s) = \int_0^\infty f(t)\, e^{-st} dt$$

Once again we can visualize this as just a "harmonic amplitude", akin to c_n in the complex Fourier series. Of course, we no longer have harmonics, but a continuous spread of frequency components.

We can also go backwards into the time domain by using

$$F(s) = \int e^{-st} f(t)dt = \int e^{-(\sigma+j\omega)t} f(t)dt = \int e^{-\sigma t} f(t)dt + \int e^{-j\omega t} f(t)dt$$

$$F(s) = \int e^{-\sigma t} f(t)dt + \int (\cos\omega t - j\sin\omega t) f(t)dt$$

$$F(s) = \int e^{-\sigma t} f(t)dt + \int (\cos\omega t) f(t)dt - j\int (\sin\omega t) f(t)dt$$

In other words, the real part of the Laplace transform is found by multiplying the time domain signal by the cosine wave ("harmonic") and then integrating (say from -∞ to +∞) and adding the first term above to it. The complex part is found by using the sine wave instead above.

Note that σ is just the real part of the s-plane (horizontal axis). So as we stretch out to the right or left side of the s-plane, we are in effect, including *progressively higher envelope tailoring terms to try and force convergence for successful integrations.* We can "correct" for that when we go back into the time domain by using the inverse Laplace transform.

Note that this technique still may not force convergence unconditionally, and so in a g eneral Laplace transform with a general stimulus/signal f(t), we may need to watch out for "regions of convergence" (ROCs) too. It is out of scope here.

Note that now the s-plane can have both real and imaginary values, besides being positive or negative. It is still best viewed as a mathematical construct. *But it is now a two-dimensional continuous spread of decomposed frequencies, and the relative phases of the decomposed frequency components is accounted for by the use of two orthogonal axis (one real called σ, one imaginary called jω).*

However it can also be shown that when the time domain is entirely real, the upper half of the s-plane is a mirror image of the lower half. In other words, any pole or zero will have a "reflected version" below "sea level" (0dB axis). We will come to that soon.

Enough of math! Let us try to visualize what we really did. In **Figure 5.10** , we show that after applying an exponential term to "precondition the harmonic amplitude", with the amount of preconditioning dependent on the horizontal distance away from the center, we land up with vertical slivers of imaginary frequencies, which are basically weighted Fourier transforms as indicated. These constitute the "harmonics", which in this case are really part of a smooth spectrum.

Finally, in **Figure 5.11** we present a table of common Laplace transforms, akin to the logarithmic tables we discussed in Part 1. The relative ease with which Laplace transforms can be manipulated, depends on the fact that the drudgery was already done long ago while creating these lookup tables. Note that the Laplace transform of the unit step function, as shown in **Figure 5.12** exists, and is equal to 1/s. With a Fourier transform, the integration would have "exploded."

To reiterate: the basic value of the Laplace transform in our case seems to be:

a) If we have an accurate model of the plant and compensator (in the frequency domain) and have calculated the transfer function of each in the s-plane, then multiplying that with an arbitrary impulse (expressed as a Laplace transform), such as a "stepped reference", we can figure out the response of the system.
b) Then going back into the time domain, we should be able to see the ringing on the output, as a result of the applied disturbance.

But as Lloyd Dixon mentioned (see Chapter 4), our models for the plant and compensator may at best be valid only for "small-signal" events. Do they apply if we slam the converter with a load transient from 0 to max load? Or even to half max load?

If we do that, we actually discover there is another major problem, which none of our analysis probably revealed. We may have encountered it in the past, but simply moved on by noting that "the bench results were 'mysteriously' different from our theoretical calculations". We blamed it on "parasitics" and came out clean!

The truth is: to understand the output ringing under large-signal events, and then suppress it, requires a rather different angle to our ongoing study, as we proceed to reveal now. It goes beyond what Mr. Dixon had hinted upon when he talked about conditional stability.

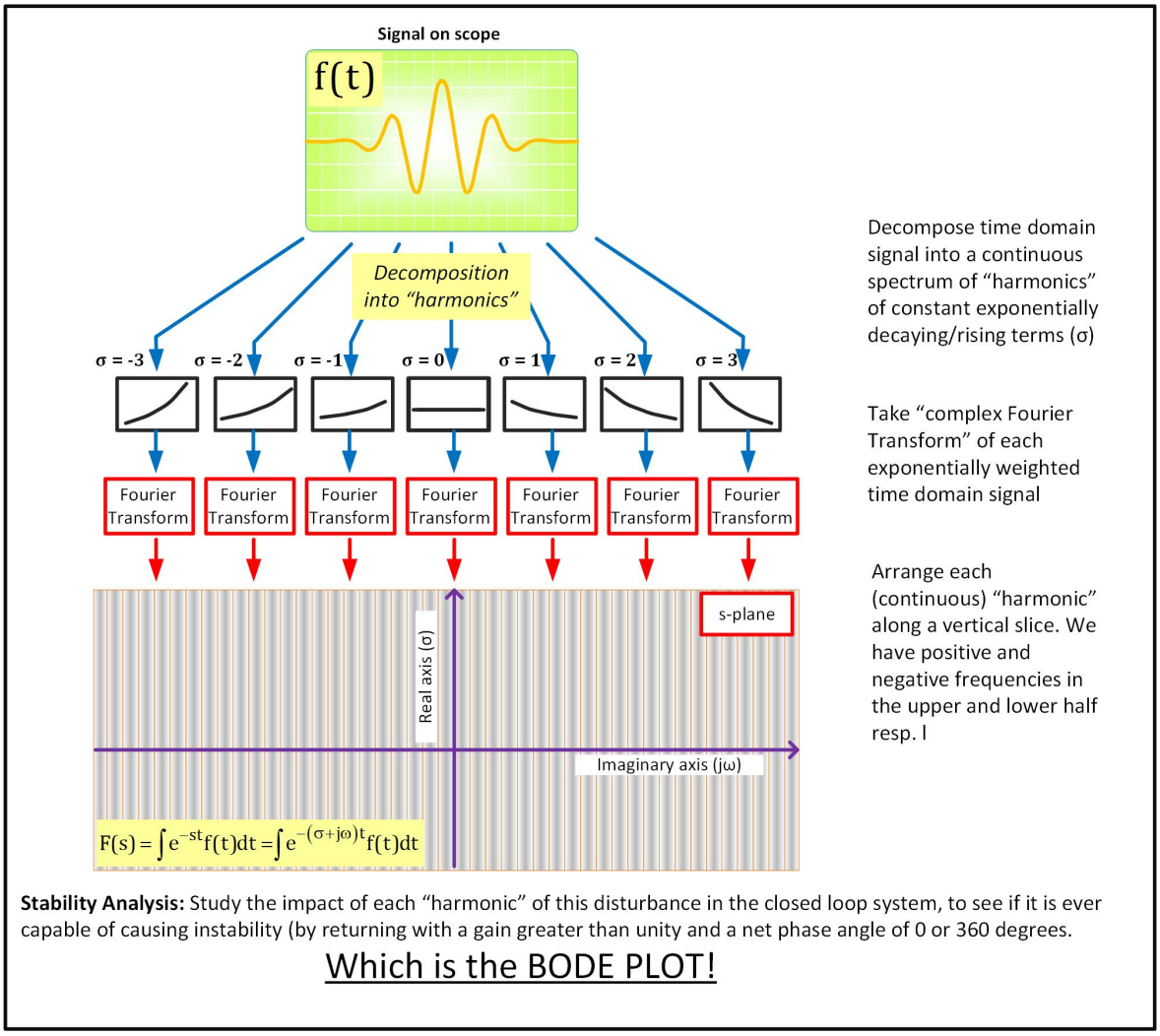


Figure 5.10: The Laplace transform decomposition process explained

Transforms of Special Functions

	f(t)	F(s)
Unit impulse :	$\delta(t)$	1
Unit step :	H(t)	$\frac{1}{s}$
Ramp:	tH(t)	$\frac{1}{s^2}$
Delayed Unit Impulse:	$\delta(t-T)$	e^{-sT}
Delayed Unit Step:	H(t-T)	$\frac{e^{-sT}}{s}$
Rectangular Pulse:	H(t)-H(t-T)	$\frac{1-e^{-sT}}{s}$

TRANSFORMS OF STANDARD FUNCTIONS

f(t)	F(s)
1	$\frac{1}{s}$
$e^{-\alpha t}$	$\frac{1}{s+\alpha}$
$\frac{1}{T}e^{-\frac{t}{T}}$	$\frac{1}{1+sT}$
$1-e^{-\alpha t}$	$\frac{\alpha}{s(s+\alpha)}$
$te^{-\alpha t}$	$\frac{1}{(s+\alpha)^2}$

f(t)	F(s)
$\sinh \beta t$	$\frac{\beta}{s^2-\beta^2}$
$\cosh \beta t$	$\frac{s}{s^2-\beta^2}$

Figure 5.11: *Some Laplace transforms*

$f(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } \geq 0 \end{cases}$

$F(s) = \frac{1}{s}$

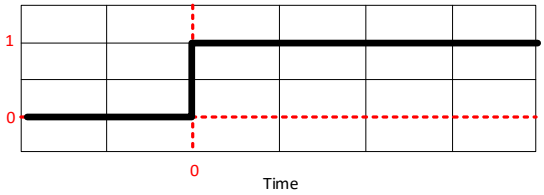


Figure 5.12: Unit step function and its Laplace transform

Building Blocks: Poles and Zeros

In our ongoing attempt to acquire mathematical intuition, we now try to gain some mastery over the gain-phase profiles of common functions in the s-plane. Because they often form the building blocks of the transfer functions of the plant and compensator.

Historically, the existence of poles and zeros was explained in the following manner. A transfer function can be eventually written out in the following general form:

$$T(s) = \frac{V(s)}{U(s)} = k\frac{a_0 + a_1 s + a_2 s^2 + a_3 s^3 + ...}{a_0 + a_1 s + a_2 s^2 + a_3 s^3 + ...}$$

$$= K\frac{\frac{s}{Z_0}\left(\frac{s}{z_1}-1\right)\left(\frac{s}{z_2}-1\right)\left(\frac{s}{z_3}-1\right)\times...\times\left(\frac{s}{Z_1}+1\right)\left(\frac{s}{Z_2}+1\right)\left(\frac{s}{Z_3}+1\right)...}{\frac{s}{P_0}\left(\frac{s}{p_1}-1\right)\left(\frac{s}{p_2}-1\right)\left(\frac{s}{p_3}-1\right)\times...\times\left(\frac{s}{P_1}+1\right)\left(\frac{s}{P_2}+1\right)\left(\frac{s}{P_3}+1\right)...}$$

The terms with "-" signs were troublesome as their "solutions" or locations were of the form s = z_n for zeros, or s = p_n for poles. These were in the right half plane (of the s-plane) and it could be shown that when we go over to the time domain, such terms tend to produce waveforms with exponentially increasing amplitudes, which we want to avoid. In contrast, the terms with a "+" sign seemed kosher, because their "solutions" or locations were of the form s = -Z_n for zeros, or s = -P_n for poles. These were in the left half plane (of the s-plane) and it could be shown that when we go over to the time domain, such terms tend to produce waveforms with exponentially decreasing amplitudes. So the effect of a disturbance eventually subsides, as we desire.

However, the above way of writing the poles and zeros, though not wrong, can be misleading, by virtue of *what it hides*. After all, the P_n could be negative! That would take it to the RHP! It could be imaginary too. So how does it really behave? And so on.

Also consider the fact that the LC double pole has the form

$$A(s)^2 + B(s) + 1$$

If we solve, we get

$$s = \frac{-B \pm \sqrt{B^2 - 4A}}{2A}$$

We thus have two "conjoined" solutions, and we could write this as

$$\left(s - \frac{-B + \sqrt{B^2 - 4A}}{2A}\right) \times \left(s - \frac{-B - \sqrt{B^2 - 4A}}{2A}\right)$$

In the worst-case, 4A can exceed B^2, the locations of the poles can be imaginary. And in that case *the solutions are complex conjugates of the form a+jb and a-jb*. That means every pole has a "reflection" below "sea level" ("sea level" being the 0dB axis!).
So, it is hard to see how, or even *why*, we would want to break the above LC pole transfer function up into something like (s-x) × (s-y), as in the generalized form above.

Plotting some Transfer Functions that we may encounter (or may not!)

Realizing that there is more complexity than implied by the "generalized" pole-zero equation above, we decide return to the basics and take a look at some common functions we are likely to encounter. Using Mathcad, we will plot them out and see what they look like.

Six conceivable functions which produce *poles* are shown in **Figure 5.13**. Six similar functions which produce *zeros* are shown in **Figure 5.14**. All the twelve plots are numbered sequentially as indicated in the figures, and we will refer to those in the discussion below. Finally, their locations in the s-plane are shown in **Figure 5.15**.

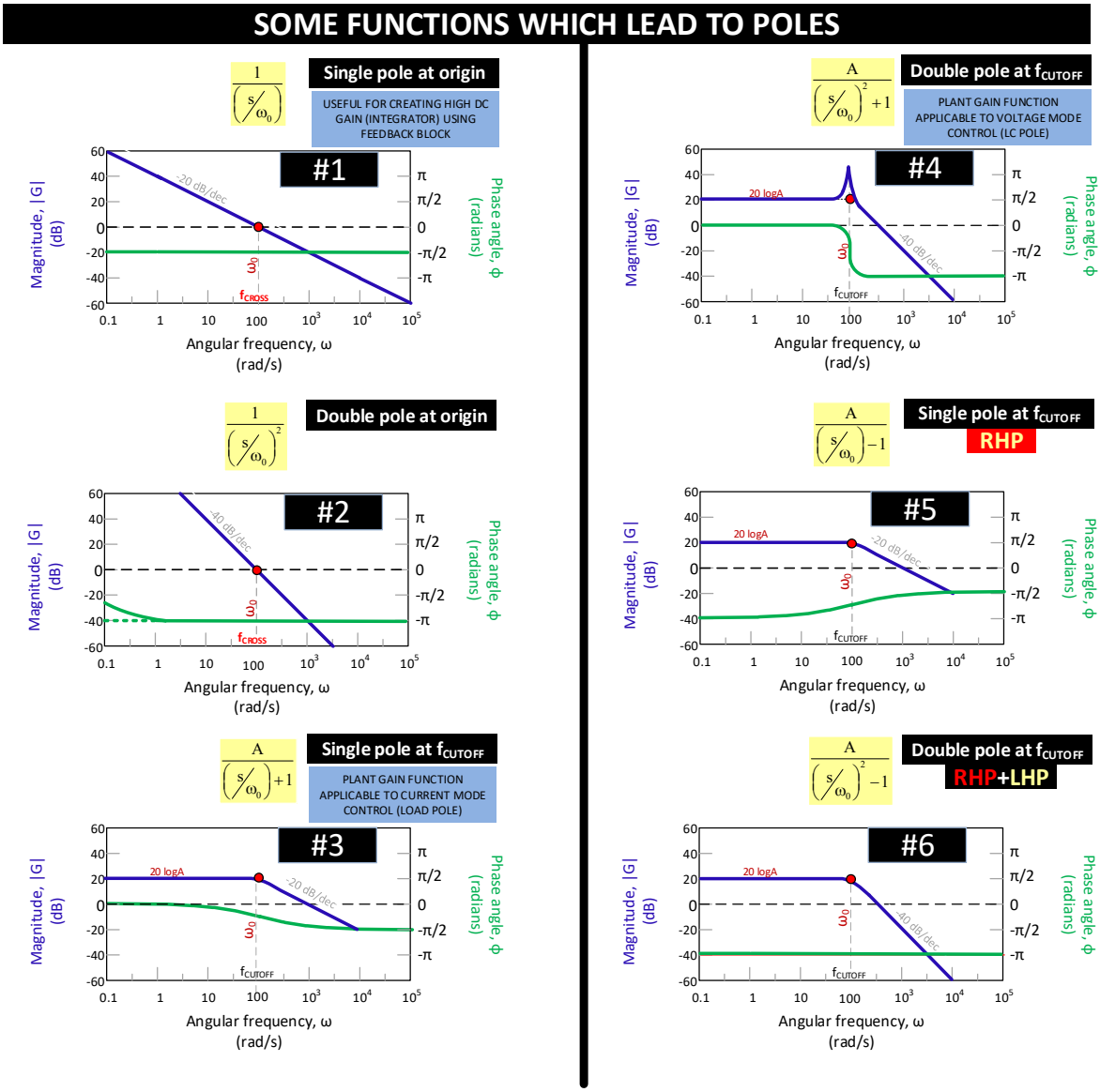


Figure 5.13: Functions which lead to poles

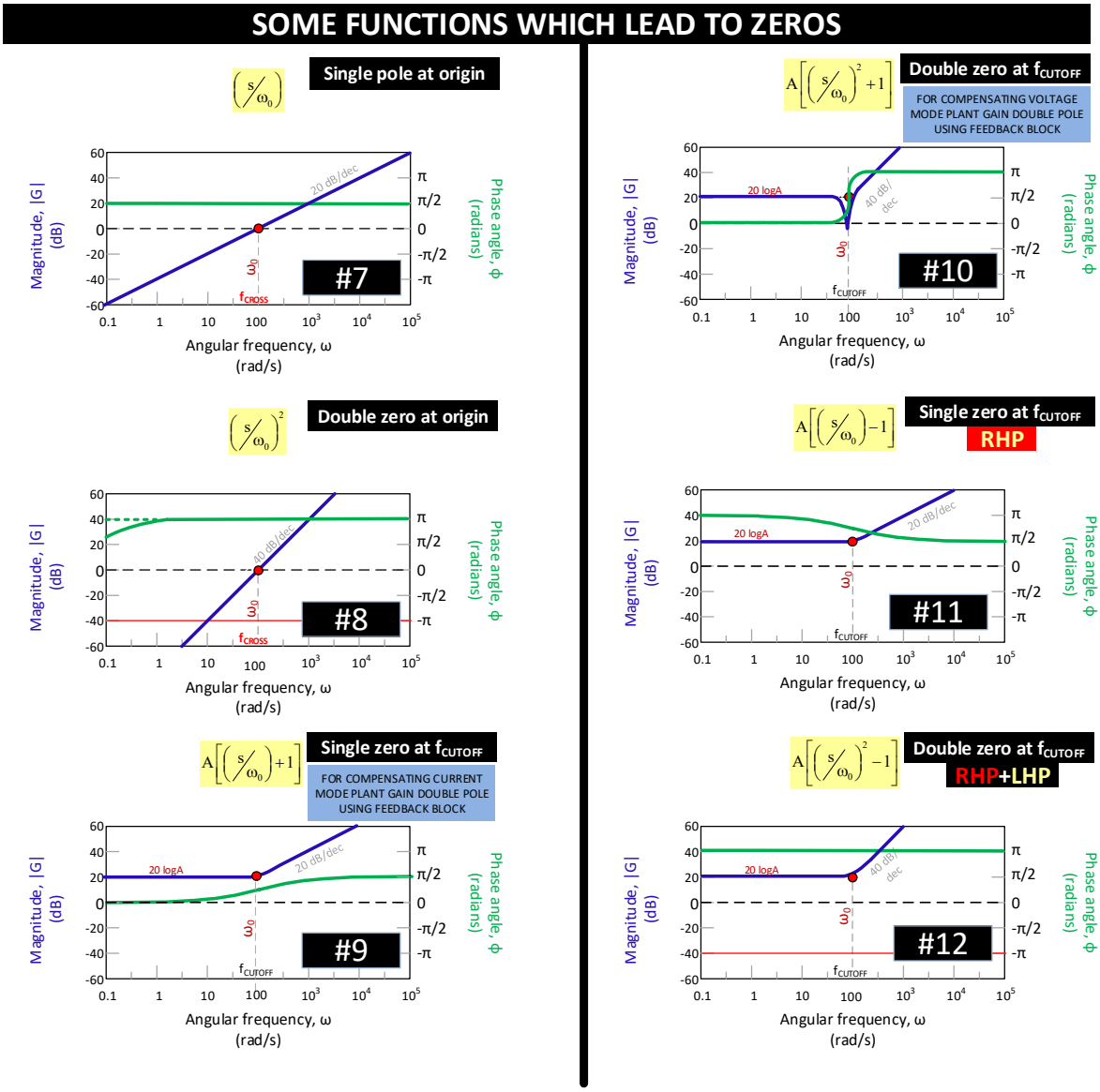


Figure 5.14: Functions which lead to zeros

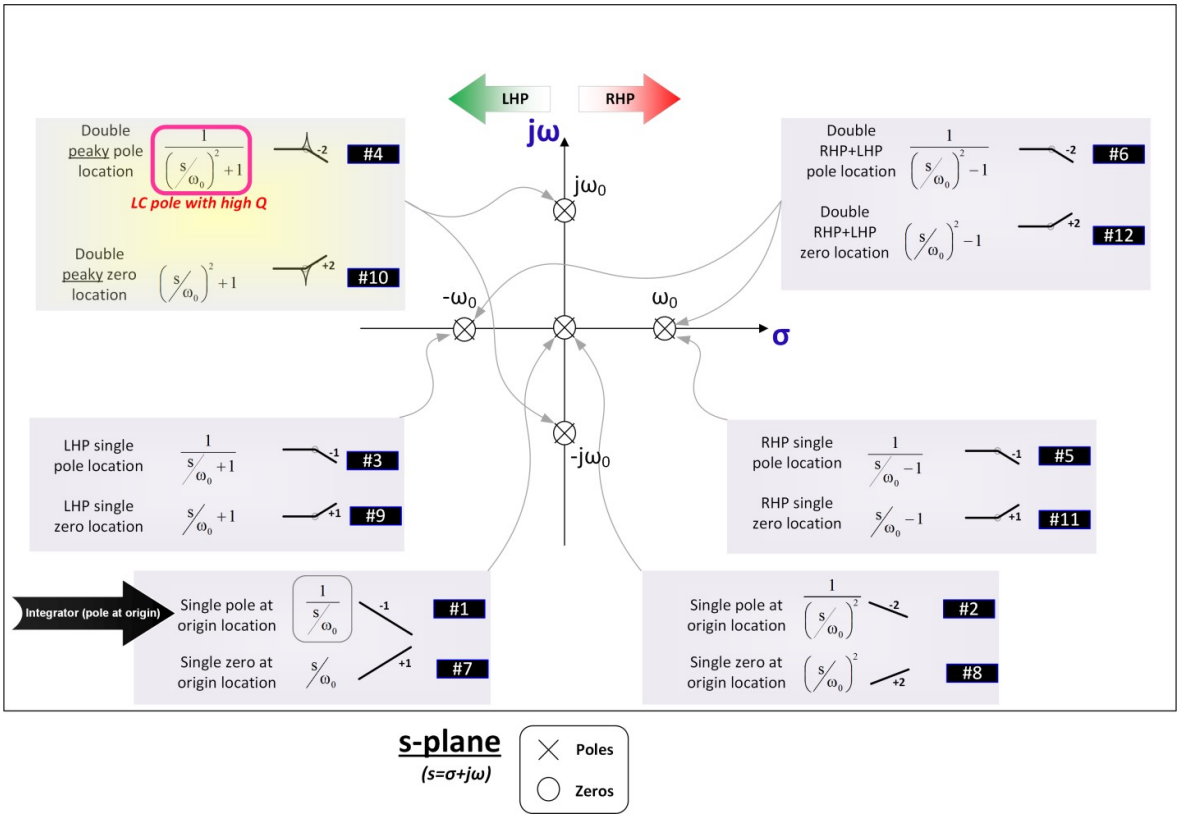


Figure 5.15: Locations of the zeros ("O") and poles ("X") of the two previous figures

Let us discuss the curves in **Figure 5.13** and **Figure 5.14**.

1) Plot #1 is a key function we will encounter. It has the form

$$H(s) = \frac{1}{s/\omega_0}$$

This has infinite DC gain, and falls off at the rate of -20dB/dec ("-1"). It crosses over at an (angular) frequency ω_0, or equivalently, a frequency $f_0 = \omega_0/2\pi$. It is a simple (first-order) "pole-at-origin" or "pole-at-zero. It happens to be the integrator function we discussed in **Figure 4.30**. Also see its implementation in **Figure 5.16**. It is an inevitable part of any analog compensator.

This function has an associated phase angle of -90°, so combined with the -180° from the negative feedback (inverting), we get -270°, which gives us 360-270 = 90° phase margin, as mentioned previously. That is why this is the preferred shape for the final loop gain T too.

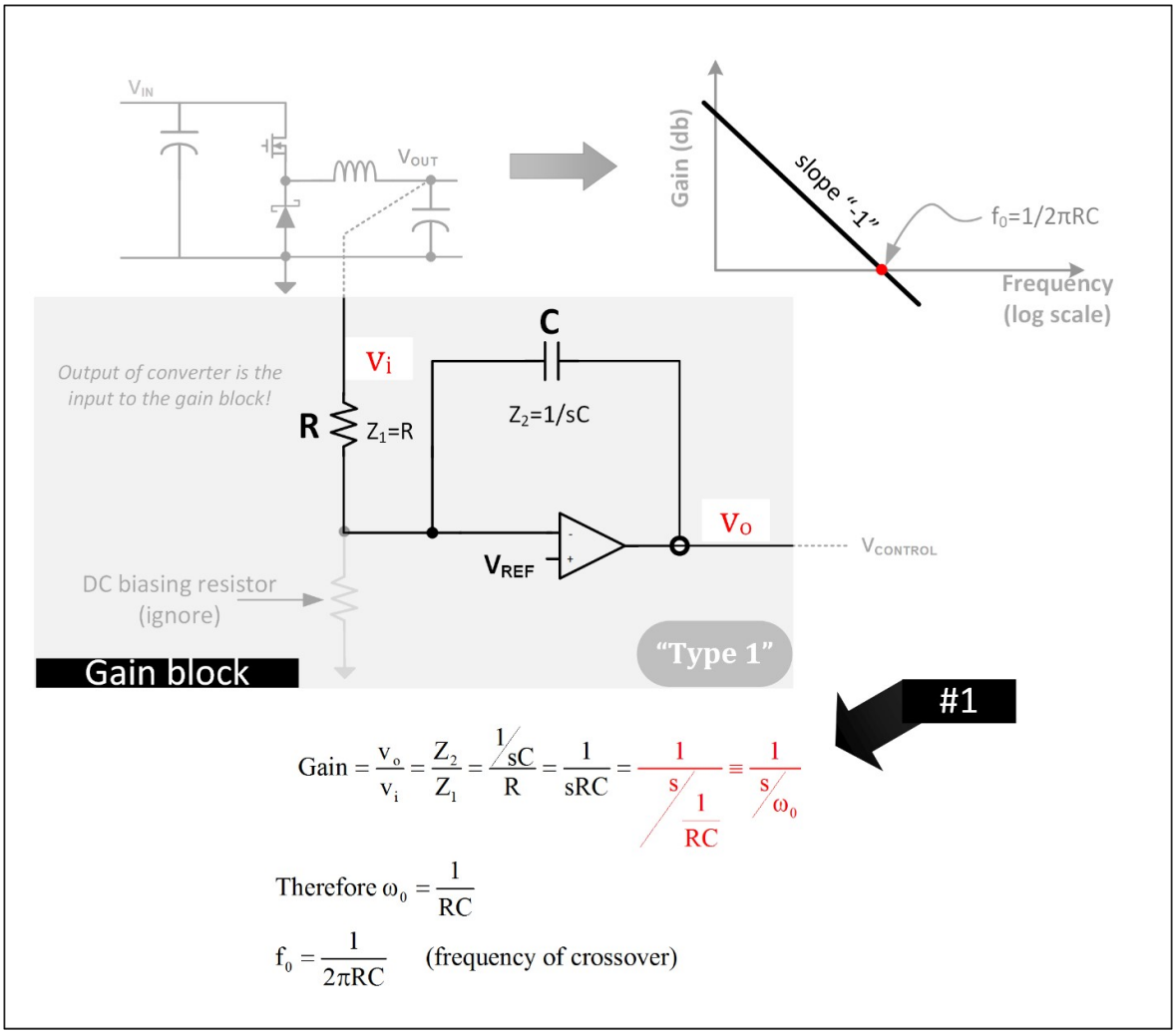


Figure 5.16: How the integrator is created in a Type 1 compensator

2) Plot #2 is of the form

$$H(s) = \frac{1}{\left(\frac{s}{\omega_0}\right)^2}$$

We can simply consider this as two coincident poles-at-origin.

$$H(s) = \frac{1}{\left(\frac{s}{\omega_0}\right)} \times \frac{1}{\left(\frac{s}{\omega_0}\right)} \equiv H_1(s) \times H_2(s) \Leftrightarrow H_{dB}(s) = H_{1\_dB}(s) + H_{2\_dB}(s)$$

The reason this is not useful to us is that the net phase lag from it is 180°, so it would give no phase margin at all. We ignore it.

3) Plot #3 is of the form

$$G(s) = \frac{1}{\left(s/\omega_0\right) + 1}$$

The location of the pole is the frequency at which the gain function intuitively "explodes" when the denominator is zero. This occurs at the following frequency.

$$\left(s/\omega_0\right) = -1. \quad \text{So} \quad s = -\omega_0$$

We know by now that in the most general representation, $s = \sigma + j\omega$, where σ is the real part of the frequency and ω its imaginary component. The s-plane is typically drawn with a vertical axis of $j\omega$ and a horizontal axis of σ.

So this function produces a simple (first-order) pole at $-\omega_0$, which is along the real axis in the left half plane (LHP). See its location in **Figure 5.15**.

Note that since LHP poles and zeros are the ones we commonly run into, the default is considered to be LHP, unless otherwise stated, i.e. as "RHP" (right half plane).
Note also that at this location, the gain function is said to "explode" because the denominator is zero. But that is only *intuitively* correct. In reality, because we are dealing with imaginary numbers, the magnitude of the gain is *not* infinite at this frequency. If we plot it out, we will discover it actually rolls off—exactly as shown in **Figure 5.13**.

4) Plot #4 is of the form

$$G(s) = \frac{1}{\left(s/\omega_0\right)^2 + 1}$$

We find its solution below:

$$\left(s/\omega_0\right)^2 = -1. \quad \text{So} \quad (s)^2 = -\omega_0{}^2. \quad \text{Solving: } s = \pm j\omega_0$$

We thus get complex conjugates: one above "sea level"/0dB axis, one below. These are shown in **Figure 5.15**.

Now, though we have apparently plotted #4 out in **Figure 5.13**, we had to add "something" to it to be able to do so. Because *imaginary frequency locations* such as this lead to (very) peaky gain responses. In fact when trying to plot out #4 using Mathcad, we soon realized that the peak at resonance was *infinite* for this function. In addition, its 180° associated phase shift was extremely abrupt at its break (resonant) frequency. In fact we couldn't even plot either the gain or phase out correctly unless we modified plot #4 a bit.

Here is what we did: *we introduced a small damping term.* We know by now that damping comes from a small "s" term. So we introduced "v", as shown:

$$G(s) = \frac{1}{\left(s/\omega_0\right)^2 + v\left(s/\omega_0\right) + 1}$$

Basically we "allowed" very small, non-zero values for "v". Then we plotted it out. Alternatively, using Q, we can write the above equation as

$$G(s) = \frac{1}{\left(s/\omega_0\right)^2 + \frac{1}{Q}\left(s/\omega_0\right) + 1}$$

where Q is the quality factor (1/v here).

The complete function, *with Q included*, is a very common function we encounter. It accrues from the plant, not the feedback circuit. (Hence we used "G", not "H" for it here). Specifically, it comes from the LC post-filter of the power converter. See **Figure 5.1.**

An infinite Q (small "v) leads to infinite peaking (no "s" term remains), and no math tool can plot it out properly, because the gain is infinity. But it really can't exist in nature either!—it's not just a limitation of math tools. Q can be very large, but is always a manageable and finite number.

We also realize that the moment the poles/zeros move out of the real axis, acquiring imaginary values, their corresponding functions exhibit higher and higher peaking. In that sense, the frequency locations are anything but imaginary! They are as real as it gets.

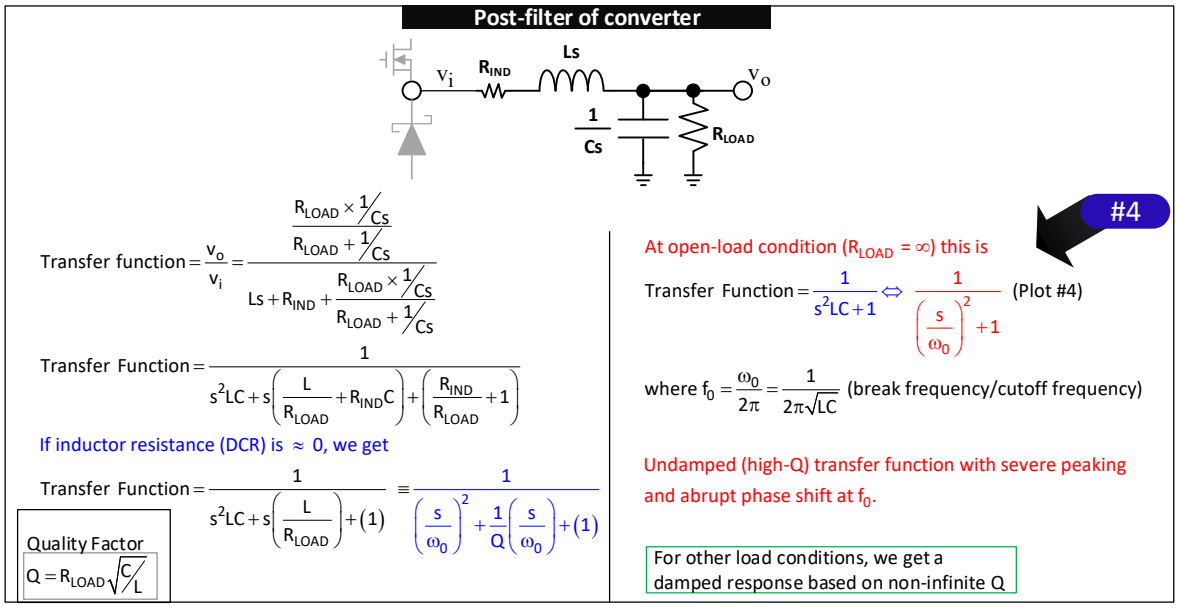


Figure 5.17: LC post filter with very high Q (output unloaded)

5) Plot # 5 is of the form

$$G(s) = \frac{1}{\left(s/\omega_0\right) - 1}$$

It has a solution (location) at s = ω_0. It is a RHP pole, so unlike regular (LHP) poles whose phase falls past the break frequency, for this it rises. It is virtually incompatible with the LHP poles and LHP zeros seen in most switchers. We therefore ignore this pole.

6) Plot #6 is of the form

$$G(s) = \frac{1}{\left(s/\omega_0\right)^2 - 1}$$

This does produce two poles, but in fact we can consider them as two coincident poles, one RHP and one LHP, by factoring it out the denominator as follows

$$\left[\left(s/\omega_0\right) - 1\right] \times \left[\left(s/\omega_0\right) + 1\right]$$

The phase contributions from both are opposite so there is no net phase shift as we see from **Figure 5.13**. There is a net gain, but why resort to unnecessary complication to produce DC gain?

7) Plot #7: this is the "opposite" of #1. It is a zero-at-origin compared to pole-at-origin. It lowers the DC gain, and is of no practical use to us. We ignore it.

8) Plot #8: this is just two of #7, i.e. two coincident zeros-at-origin. We ignore it too, for the same reason as #7.

9) Plot #9: This is a first-order (simple) LHP zero of the form

$$H(s) = \left(s/\omega_0\right) + 1$$

We could place *two* of these single zeros at the LC pole frequency to try and cancel it out. That is exactly what we do in standard Type 3 compensation as detailed in Part 1.

10) Plot #10 is the complement of the double pole peaky LC response is the double zero gain dip. Once again, we needed to add a term involving Q to be able to plot it out.

$$H(s) = \left(s/\omega_0\right)^2 + \frac{1}{Q}\left(s/\omega_0\right) + 1$$

This function can theoretically be used to provide the two zeros in the feedback block at the exact position of the double LC pole coming from the plant. Hence we used the symbol "H" not "G" for it here. This tells us something that is rarely discussed:
Just as we talk about the Q of the plant, why can't we talk about the Q of the compensator too? That could also be a way to enact our standard compensation strategy, instead of just using two zeros of the type plot #9, as discussed above.

Far more on the use of this function (plot #10) will follow soon.

11) Plot #11 is of the form

$$G(s) = \left(s/\omega_0\right) - 1$$

Plotting its magnitude and phase, we see that it is a zero, because the gain starts sloping upwards eventually. But this increase in gain is accompanied by a *decrease* in phase, not an increase as for a "regular" LHP zero, shown in plot #9. We try to avoid LHP poles and zeros as far as we can, because they tend to produce *escalating* responses when we try to go back to the time domain. More on this later.

Unfortunately, this RHP zero does appear in the plant of boost and buck-boost topologies, but not for the buck. Intuitively, that zero is often explained by saying that in a boost and buck-boost, energy is delivered to the output only during the switch OFF time. But if for example there is a sudden load increase, and the output dips, the control loop responds by increasing the switch ON-time in an effort to build up more energy in it as required. Unfortunately that reduces the time for the energy to be delivered to the output, so that dips even further momentarily.

LHP zeros are very hard to stabilize or avoid, and usually the only solution is to roll off the loop gain at a much lower frequency than we usually do for a buck.

12) Plot #12 is of the form

$$G(s) = \left({}^{s}\!/_{\omega_0} \right)^2 - 1$$

This does produce two zeros, but in fact we can consider them as two coincident zeros, one RHP and one LHP, by factoring it out as follows

$$G(s) = \left[\left({}^{s}\!/_{\omega_0} \right) - 1 \right] \times \left[\left({}^{s}\!/_{\omega_0} \right) + 1 \right]$$

As we see in **Figure 5.15**, the solutions are along the real axis, at $+\omega_0$ and $-\omega_0$. So there are two zeros of opposite types, and as a result, there is *no associated phase shift* at the resonant frequency as we can confirm from the plot in **Figure 5.14**. They "cancel" each other out in terms of phase, but do produce gain. Not typically useful, as indicated for its pole version (plot #6).

Summary of Common Functions

We have understood the behavior of common transfer functions in the frequency domain. We know the locations of their poles and zeros. We can recognize some of them from our previous experience dealing with analog control loops.

We are also now keenly aware of the fact that not just the phase *shift* (if any), say at the resonant frequency, is of importance, but the actual (net) phase angle too—so as to prevent 180° shift from ever occurring with an accompanying gain of unity (0dB) or greater. (That is another thing glossed over in related literature sometimes). For example, the function "$1/(s/\omega_0)$" has a constant phase lag of 90°, and its "breakpoint" is at a very low un-definable/un-plottable, literally invisibly low frequency. But because of its seemingly frequency-independent, flat phase lag of 90°, it leaves us with only 90° to play around with in terms of phase margin. Which is why we *must* cancel out the pole coming from the plant, whether it is VMC (180° phase lag from L and C_{OUT}) or current-mode control (90° phase lag from R_{LOAD} and C_{OUT}).
We also learn to avoid RHP poles/zeros as far as possible. They tend to give escalating responses in the time domain, whereas the LHP tends to give responses that decay, as we desire. To show this more clearly, let us start with the definition of the Laplace transform

$$F(s) = \int_{-0}^{\infty} e^{-st} f(t) dt$$

We calculate the Laplace transform of the exponentially decaying function e^{-at} (starting from zero at t = 0, as we always assume in normal Laplace transform analysis). It turns out to be $1/(s+a)$. See also the tables in **Figure 5.11**. Equivalently stated, the *inverse* Laplace transform of the step function $1/(s+a)$ is e^{-at}, which is a nice exponentially decaying function—for any "a" that is real and positive.

But what exactly is the function $1/(s+a)$? It is by definition, a pole at frequency "-a"—which is in the *left* half of the s-plane, assuming "a" is real and positive. We therefore realize that the unit step function ($1/s$ or $1/(s+a)$) corresponds to a "well-behaved" function in the time domain, one which *decays* exponentially with time.

In power supplies, the poles and zeros we deal with are usually on the left half plane. We simply try to avoid the right half plane (RHP) altogether. We thus lower the bandwidth (crossover frequency) significantly, in boost and buck-boost topologies, where the notorious RHP zero emerges. There is almost no other solution to this particular RHP problem.

Analog to Digital

With this new understanding, the basic idea behind ***classic analog control loop compensation*** in a voltage-mode-controlled buck is therefore:

A) Create a pole at the origin (p0) in the feedback block (using plot # 1)
B) Note the location of the double pole at the LC resonant frequency (plot #4)
C) Place two (LHP) zeros (z1 and z2) at the LC pole frequency (two of plot #9)
D) Note the location of the ESR-zero (plot #9)
E) Place a simple (LHP) pole (p1) at the location of the ESR-zero (plot #3)
F) Place a simple high-frequency (LHP) pole (p2) at f_{CROSS}, $10\times f_{CROSS}$ or $f_{SW}/2$ (plot #3)

Refer to the standard analog compensation strategy shown in **Figure 5.18**.

In our version of digital compensation, we replace step C above with this new step (to be discussed in detail soon):

C) <u>Place one second-order zero at the LC pole frequency (plot #10)</u>

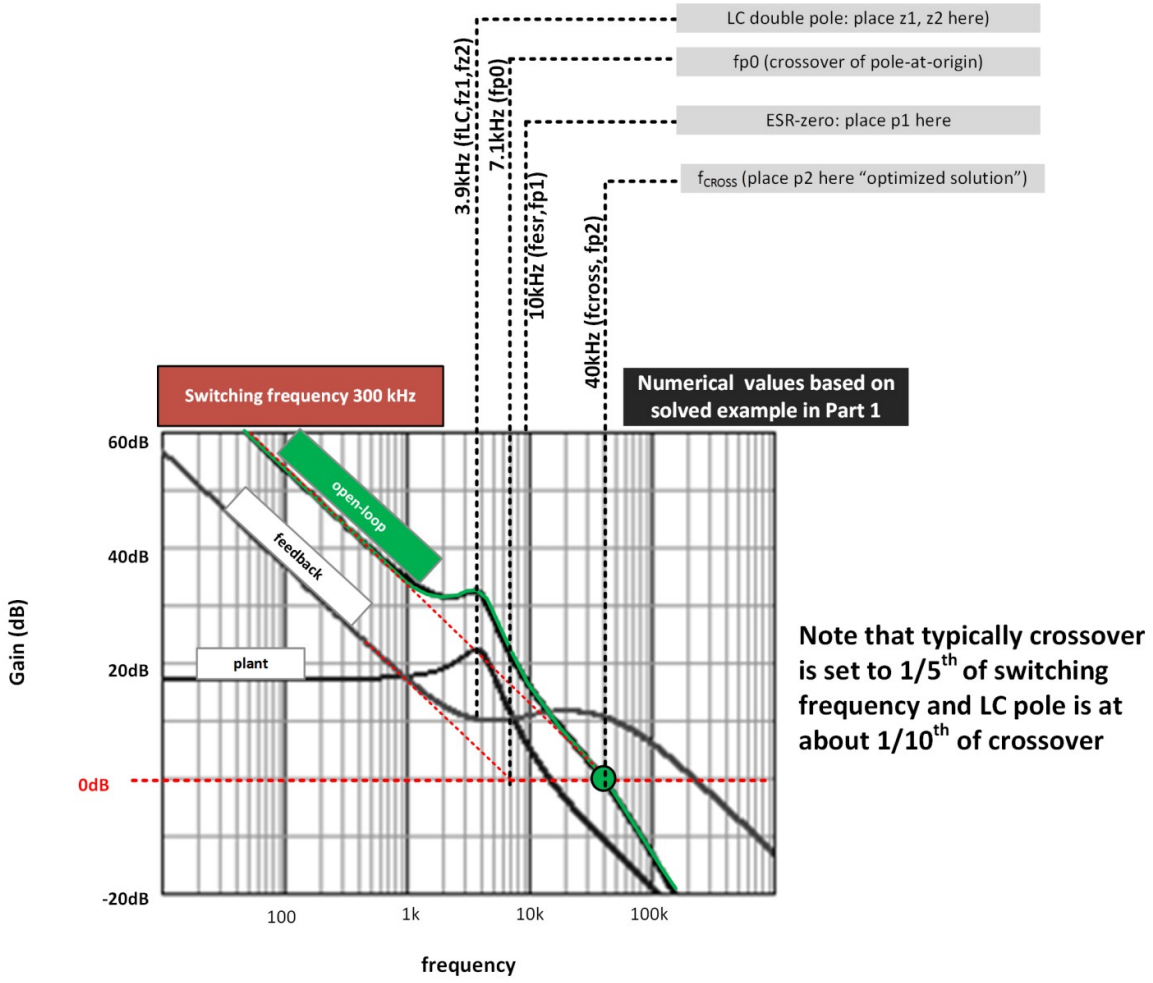


Figure 5.4: Analog compensation strategy

"Peakiness" of Transfer Functions

Certain concepts we may have realized from the previous pages after plotting out various functions in **Figure 5.13** and **Figure 5.14**, are captured more clearly in **Figure 5.19**. For example, if poles and zeros lie along the imaginary axis, we get a "peaky" response (in theory, an infinite peak exactly along the vertical axis). Also, if the poles and zeros lie along the real axis, we get a "damped" response.

Looking at **Figure 5.19**, we try to understand which term in the transfer function leads to the "peaking". As mentioned, we had actually added a small (yet undeclared) term in "s" to be able to Plot #4 in the first place. Otherwise the peaking would have been infinite and the associated phase shift extremely abrupt at the resonant frequency ω_0.

The transfer function of the LC post-filter stage of the plant (in VMC), with this "s" term highlighted is as follows. It is essentially plot #4 with a certain Q included.

$$G(s) = \frac{1}{\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q}\left(\frac{s}{\omega_0}\right) + 1}$$

From the definition of ω_0 ($=1/(LC)^{1/2}$), and Q ($= R_{LOAD} \times (C/L)^{1/2}$), we conclude that the magnitude of the term involving "s" does <u>not</u> (noticeably) affect the resonant frequency, which is determined (almost) solely by the term involving "s^2", but it does determine the "peakiness" of the response at that resonant frequency.

In general, if the coefficient of s is small, the "peakiness" is much higher.

We have a similar but inverted plot (for zeros) in Plot #10. And so we realize we may need to start thinking in terms of the Q of the compensator too, not just the plant.

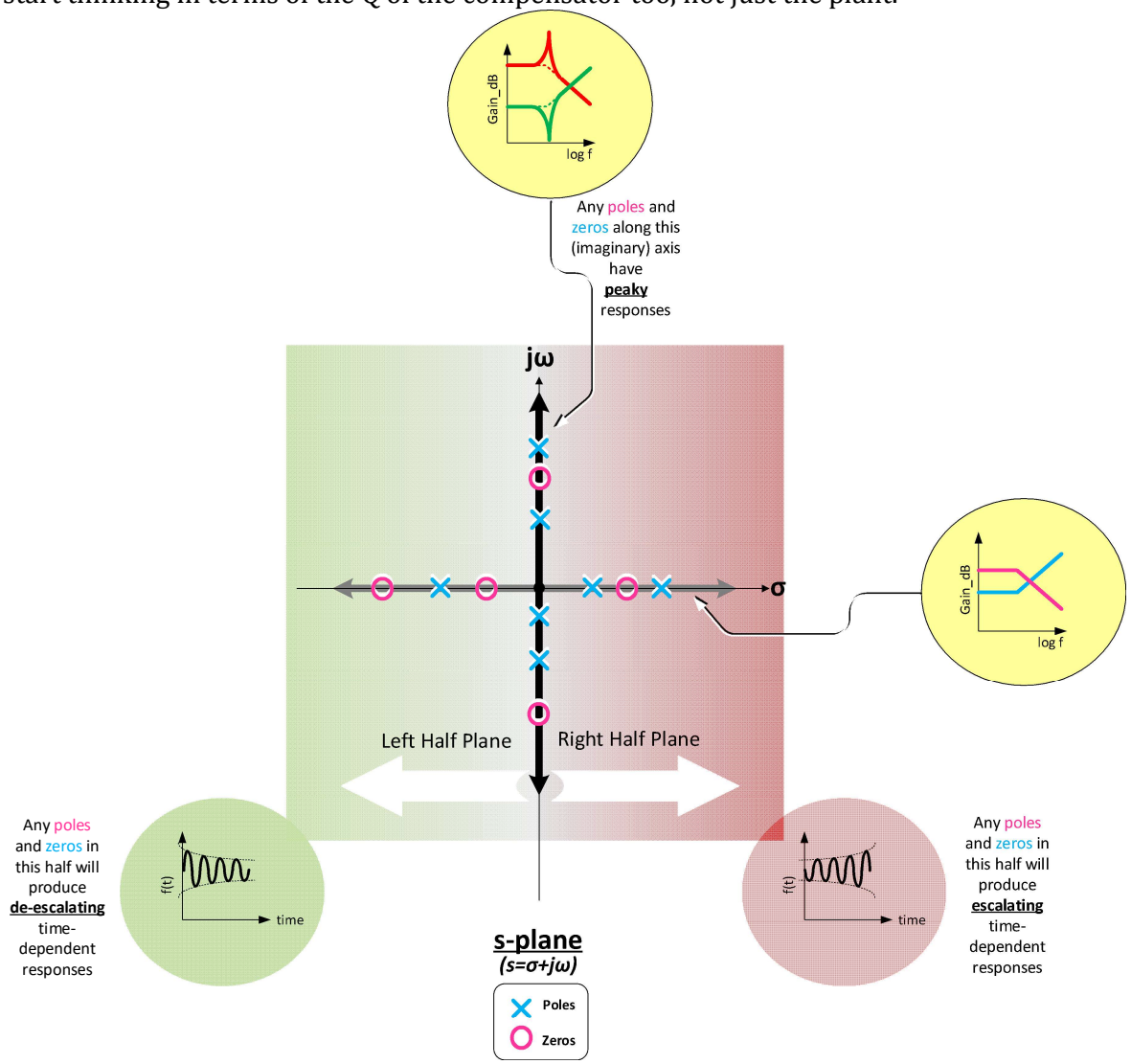


Figure 5.19: Summary of behavior of poles and zeros of common transfer functions

Limits of Analog Compensation

Despite our presumed familiarity with analog control, and even accepting the fact that we can't change the compensation "on the fly", analog loops are difficult to tweak as detailed in Chapter 4. We can do that far more easily in digital implementations. Also, analog compensation is inherently limited *in a certain way* which we will soon describe, though that is rarely clarified in related literature.

In **Figure 5.20**, we once again present the equations and typical strategy for Type 3 compensation as used in voltage mode controlled regulators. As discussed earlier, the key is introducing two coincident zeros in the compensator, at the exact position of the double LC-pole of the plant. The purpose of that is to "cancel" the double pole. We also require a pole at the origin, to get the DC gain high, for rejecting low-frequency disturbances in particular. But we also roll-off the loop gain to avoid oscillations.

What we *really* and unavoidably require from the compensator is *one pole-at-origin and two zeros*. However the way a Type 3 compensation is set up, using three capacitors and two resistors, there is a lot of interaction between all the R's and C's, so we end up *with more than that*.

The following is the well-known gain function of a Type 3 compensator:

$$H(s) \approx \frac{(sC_2(R_1+R_3)+1)\times(sC_1R_2+1)}{(sR_1C_1)\times(sC_2R_3+1)\times(sR_2C_3+1)}$$

Equivalently

$$H(s) \approx \frac{s^2[C_1C_2(R_1+R_3)R_2]+s[C_2(R_1+R_3)+C_1R_2]+1}{sR_1C_1(sC_2R_3+1)(sR_2C_3+1)}$$

The term in "s" is, as usual, defining a certain "Q" of the compensator, and we need to start being cognizant about that.

We now compare this with our previous generalized equation:

$$T(s)=\frac{V(s)}{U(s)}=K\frac{\frac{s}{Z_0}\left(\frac{s}{z_1}-1\right)\left(\frac{s}{z_2}-1\right)\left(\frac{s}{z_3}-1\right)\times\ldots\times\left(\frac{s}{Z_1}+1\right)\left(\frac{s}{Z_2}+1\right)\left(\frac{s}{Z_3}+1\right)\ldots}{\frac{s}{P_0}\left(\frac{s}{p_1}-1\right)\left(\frac{s}{p_2}-1\right)\left(\frac{s}{p_3}-1\right)\times\ldots\times\left(\frac{s}{P_1}+1\right)\left(\frac{s}{P_2}+1\right)\left(\frac{s}{P_3}+1\right)\ldots}$$

We thus realize that Type 3 compensation gives us two normal (LHP) zeros as desired, but also two normal (LHP) poles, besides the much-desired pole-at-origin. Note that all the poles and zeros are based on RC values and are therefore "single-order". They will provide an upward or downward slope of 20dB/decade past the point defined by the applicable break frequency ($\omega_X = 1/RC$). *There will be no "peakiness" however, because these frequency locations are along the real axis (see **Figure 5.19**).*

The thing to note here is *we need to do something with the two extra poles*, even if we don't think we need them. One obvious use of one of the "extra poles" is to cancel out the ESR-zero coming from the plant. This may however be at a very high frequency nowadays, because of the low ESR of modern ceramic capacitors. So it may be irrelevant to cancel.

The remaining "extra pole" is a subject of some debate as indicated in the figure. Some suggest putting it at fsw/2, others at fsw/10. At the Unitrode Seminar in Germany in 1996, Lloyd Dixon suggested putting it at the crossover frequency, f_{CROSS}.
Another problem is, as we know from Part 1 is that we need to simplify the actual transfer function of a Type 3 compensator to make it usable—by introducing the approximation C1>>C3. Which implies, we already have an inherent error in all our calculations. Besides that, we could have an additional +, - 10% tolerance, just based on the nominal value of available caps. And so on. *There is barely any precision in analog compensation*.

Besides that, how difficult is it to tweak the poles and zeros? We discussed that in Part 1 too. The overall answer is *very difficult*.

But the biggest problem with standard Type 3 analog compensation is yet to come! We now discuss that.

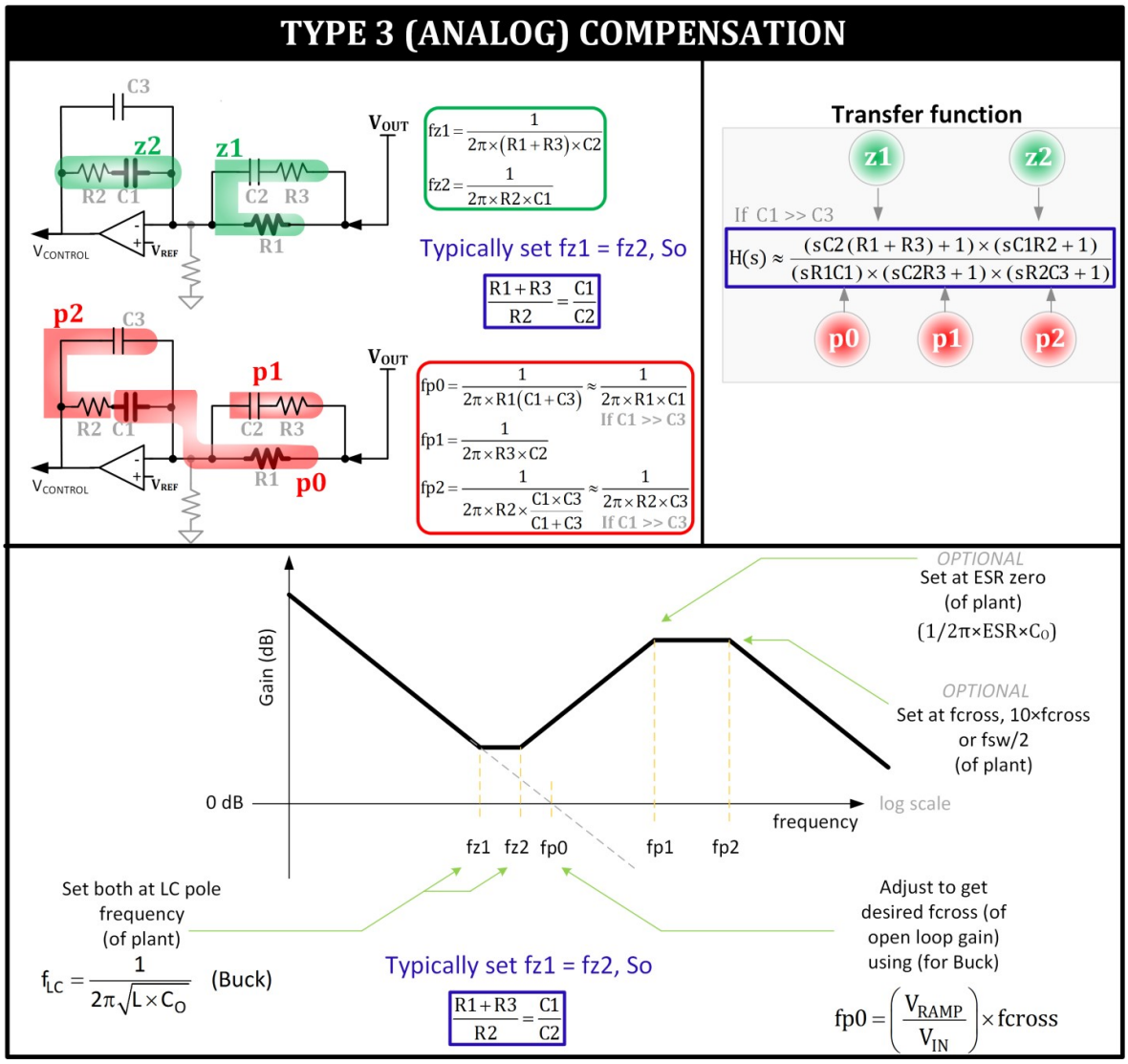


Figure 5.20: Summary of Type 3 compensation

Unrecognized Effects of "Q-Mismatch" in Type 3 Analog Compensators

One of the most common and *erroneous* assumptions of loop compensation is the statement:

"We can place two single-order zeros at the same frequency location, and that is equivalent to a double zero".

This is what we always implicitly assume when we place two zeros in the compensator to "kill" the LC double pole of the plant. ***But that cancellation is actually only partial as we will discover***.

Suppose we place two single zeros at the exact same frequency location. That gives us the following transfer function

$$H(s)=\left(\frac{s}{\omega_Z}+1\right)^2 \Rightarrow \frac{s^2}{\omega_Z{}^2}+2\left(\frac{s}{\omega_Z}\right)+1$$

Whereas, we know that the plant gives us the following transfer function

$$G(s)=\frac{1}{\left(s/\omega_0\right)^2+\frac{1}{Q}\left(s/\omega_0\right)+1}$$

For the two zeros to cancel the pole, we need to have the two cancel each other out

$$\frac{\left(\frac{s^2}{\omega_Z{}^2}+2\left(\frac{s}{\omega_Z}\right)+1\right)}{\left[\left(s/\omega_0\right)^2+\frac{1}{Q}\left(s/\omega_0\right)+1\right]}=1$$

Clearly that happens only for one condition: Q of plant equal to 0.5. Otherwise not.

Because as load changes, the Q of the plant can vary from small values (at max load) to very large values (light loads).

In other words, to cancel each other out, not only the LC pole frequency (ω_0 or ω_{LC} in this case) must equal the frequency of the two coincident zeros (ω_Z), but the crucial terms involving "s", which determines the amount of damping as discussed, must be equal. And under light load conditions, Q of the plant can become very high, so the "s" term becomes very small, and so the LC peaking is very severe. Whereas in the analog compensator with coincident zeros, we have the factor "2" in the term involving s, which is far from negligible at the resonant frequency ω_Z. As a result of that, *the response of the compensator is extremely well-damped. But not the plant's response. How can they hope to cancel each other out?*
Note that in the following article:
http://powerelectronics.com/power_systems/simulation_modeling/Transient-response-phase-margin-PET.pdf , the author, Mr. Basso, tweaked the plant gain (via simulations as usual), to get a Q of 0.5, and noted that it gave the best transient response. But instead of recognizing that it was simply a case of Q-matching, because the analog compensator had the same Q=0.5, he seems to have *attributed it to the phase margin* (76°) at that point, and thus declared that 76° was the optimum phase margin (despite the rest of the industry saying it should be 45°). Actually, this was almost clear evidence that *by simply matching the Q of the plant and the compensator*, we get the best transient response because the pole-zero cancelation is near-perfect. Of course it is not realistic to set Q of the plant to 0.5, because that occurs for only one specific load. A better approach would be to let Q of the plant be what it is, and try to adjust the Q of the compensator to match the Q of the plant, for a wide range of loads. That is what we can accomplish very easily using digital control.

Digging deeper into "Q-Mismatch"

We have just stumbled upon the topic of ***"Q-mismatch"***. Because the Q of the plant can vary anywhere from decimal values at max load to very big numbers at light loads. Whereas the Q of the compensator (and we almost forgot to even ask *what* it was), is *fixed*!

At what value?

Looking at the s-terms:

$$\frac{1}{Q}\left(\frac{s}{\omega_0}\right) \text{ compared with } 2\left(\frac{s}{\omega_z}\right)$$

we realize *the Q of a Type 3 analog compensator (with coincident zeros) is always **0.5**, irrespective of the R and C's used in the feedback network.*

This is the major issue plaguing standard analog compensators—its inability to cancel the LC pole properly, leading to *conditional stability*, which as it turns out, is not as harmless as we may have thought.

Before we discuss conditional stability again, note that there is another way to state the Q-mismatch issue. In terms of the functions we plotted in **Figure 5.13** and **Figure 5.14,** *we realize we are essentially guilty of trying to cancel the peaky LC double pole and its complex conjugate poles (spread vertically parallel to the imaginary axis, above and below "sea level", i.e. the 0 dB axis), with two simple zeros, both completely constrained on the real axis. How can they ever cancel each other out totally, if they are not located on top of each other in the s-plane?*

Conditional Stability Re-examined

The pole-zero "cancelation" is barely perfect in standard analog techniques, and as a result *the loop gain shows severe peaking and "wobbling" of phase around the LC pole frequency.* See **Figure 5.21.**

But why is that a problem? As mentioned in Chapter 4, this is called "conditional stability". Lloyd Dixon warned us a bit about that in reference to gain collapsing; Dr Ray Ridley thought it was "rugged" anyway, and so on. No one really seems to have realized it was a problem related to the *output ringing* under *large-signal* transients.

But now let us look at the evidence. We took a buck converter with the following parameters:
L= 330 nH, C = 546 µF, ESR = 520 µΩ, DCR = 8.53 mΩ ,V_{IN} = 12V , V_{OUT} = 1.2V, I_{OMAX} = 5A

Its calculated LC pole frequency is 11.86 kHz. Next, we applied a 0-2A load transient. The waveform in **Figure 5.22** emerged. Eyeballing the output, the ringing is at ~12.56 kHz, very close to the theoretically calculated LC pole frequency. And these are not simulations!

Question: Why do we see ringing at frequencies *close to the LC resonant frequency*, not at the supposed "likeliest" frequency at which instability occurs, i.e. f_{CROSS}?

We could argue along the lines of Lloyd Dixon here: that under large-signal events such as this, the inductor is unable to provide power temporarily for several cycles as the current ramps up to its new value (the "inductor reinitialization" problem mentioned in Chapter 4). As a result the gain collapses and f_{CROSS} temporarily falls close to the LC pole frequency, triggering temporary oscillations at that frequency. The steep wobble in the phase at the LC pole frequency, especially at light loads, as shown in **Figure 5.21**, is clearly not helping us.

We conclude that conditional stability is not as harmless as made out to be commonly.

And perhaps none of this will ever emerge if we continue relying on simulations using small-signal models. That was also the opinion voiced by Lloyd Dixon too, in 1996: *"...there has been a lack of balance and a tendency to try to force behavior that is uniquely related to switching phenomena into linear equivalent models (with sometimes uncertain results). Many of the major significant problems with switching power supplies do not show up in the frequency domain, or in the time domain using averaged models, unless these problems are anticipated in advance and provided for in the models.* ***Simulation in the time domain using switched models, although slower, reveals these problems that would have been hidden****."*

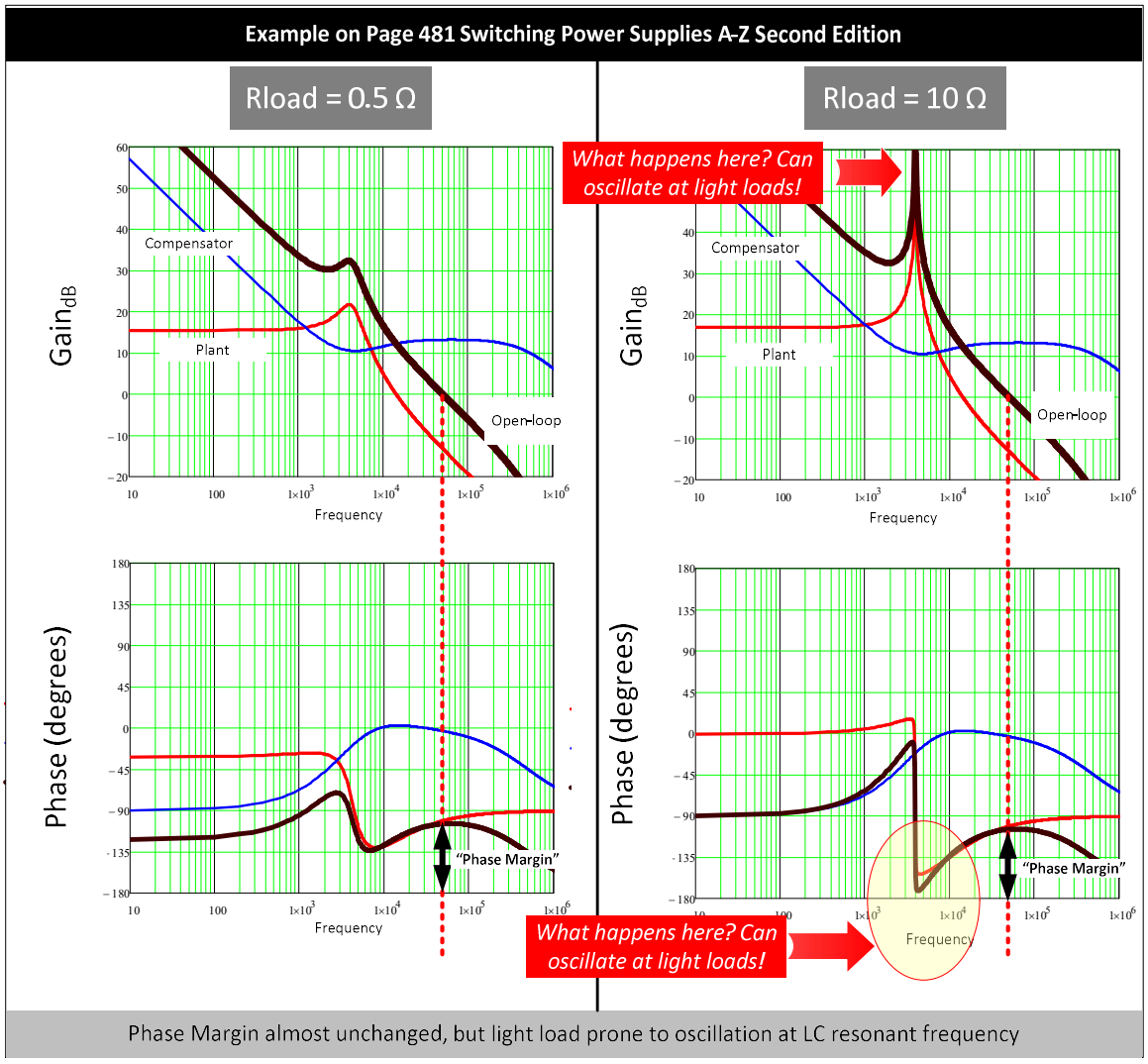


Figure 5.21: Conditional stability at light loads (high Q)

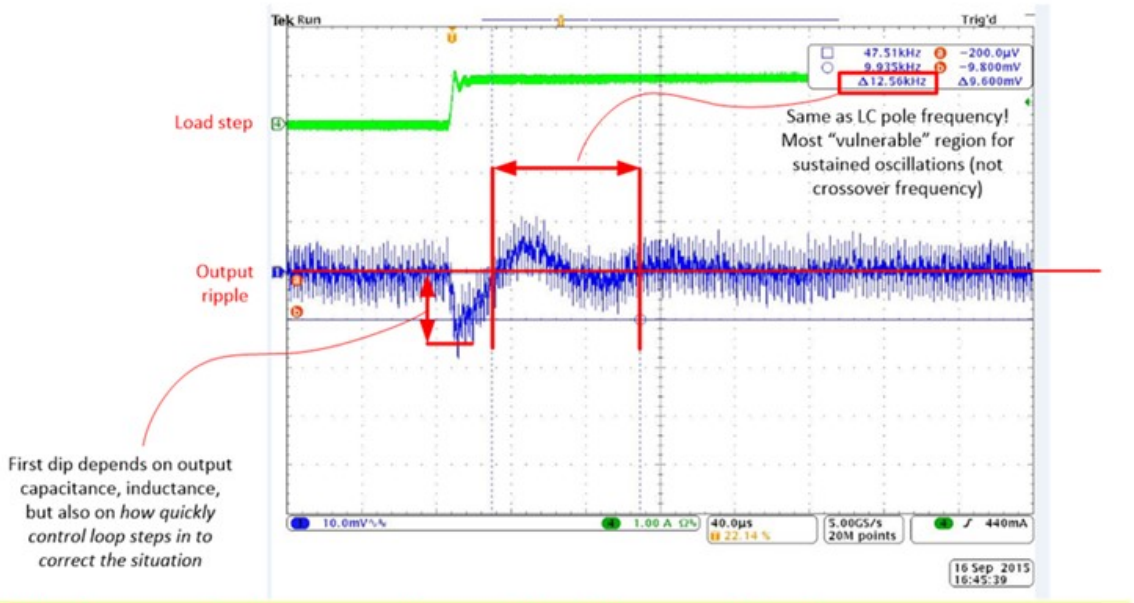


Figure 5.22: Output ringing analysis

"Forcing" an Analog Compensators to be More Effective

Let us use Mathcad to do some mathematical explorations on a Type 3 compensator. Its transfer function (for $C_1 >> C_3$) is

$$H(s) \approx \frac{(sC_2(R_1+R_3)+1)\times(sC_1R_2+1)}{(sR_1C_1)\times(sC_2R_3+1)\times(sR_2C_3+1)}$$

Equivalently

$$H(s) \approx \frac{s^2[C_1C_2(R_1+R_3)R_2]+s[C_2(R_1+R_3)+C_1R_2]+1}{sR_1C_1(sC_2R_3+1)(sR_2C_3+1)}$$

Now, knowing that the term involving "s" causes the compensator response to be excessively damped, compared to the relatively undamped (peaky) LC pole response of the plant, we try to mathematically induce additional "peakiness" ("dippiness" or "antiresonance" in this case!) in the compensator response, by introducing an arbitrary factor "v" as shown below.

$$H(s) \approx \frac{s^2[C_1C_2(R_1+R_3)R_2]+s[C_2(R_1+R_3)+C_1R_2]\times v+1}{sR_1C_1(sC_2R_3+1)(sR_2C_3+1)}$$

So by *reducing* the value of v (the "s" term) we can create a severe dip in the compensator gain similar to plot #10 in **Figure 5.14**. If we carefully control this antiresonance, we can almost completely cancel out the LC resonance for any load under consideration.

To prove this, in **Figure 5.23**, we compare the light-load response of a standard analog compensator to the response that is obtained by forcibly tweaking it by setting v = 0.01. *We see an almost complete cancelation of the LC pole occurring now, in terms of both the loop gain magnitude and phase.* Most importantly, conditional stability is significantly reduced if not wiped out. We expect a much better transient response too now, as we will soon see.

Indeed, this was just mathematical. There is no easy way to implement it using analog control. However, we will now proceed to show that digital control provides us the capability to implement such a unique transfer function for the compensator.

In brief: Instead of putting two coincident single-order zeros based on plot #9 at the LC pole, we can use digital techniques to produce a second-order zero with a more appropriate response curve based on plot #10.

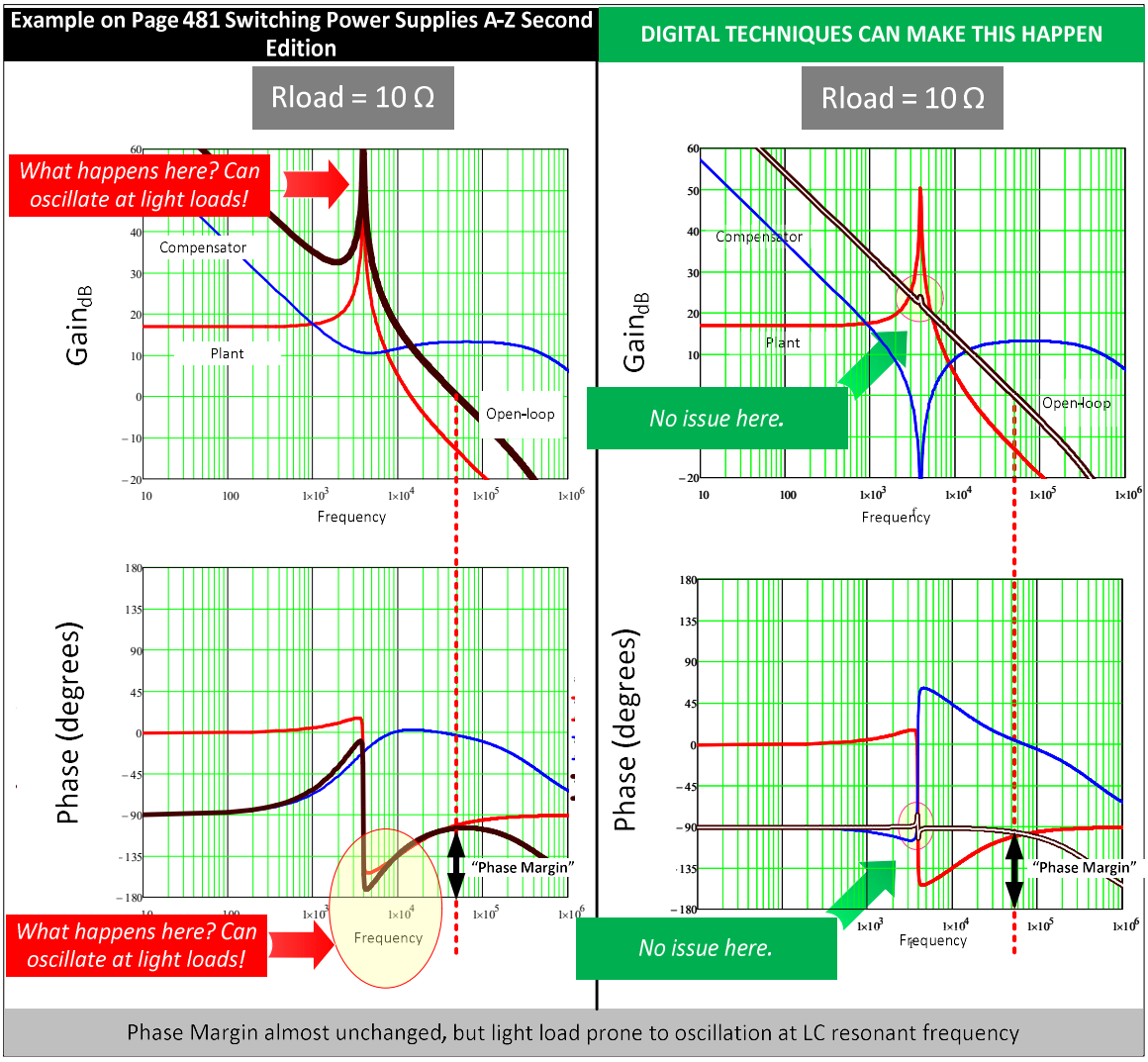


Figure 5.23: What digital techniques can accomplish

What is Damping?

Let us now define a general transfer function and see when we get an imaginary component to the solution.

$$H(s) = \left[\left(\frac{s}{\omega_0} \right)^2 + \frac{1}{Q} \left(\frac{s}{\omega_0} \right) + 1 \right]$$

The solutions (locations of the zeros) are given by setting the above to zero. Solving

$$\left(\frac{s}{\omega_0} \right) = -\frac{1}{2Q} \pm \sqrt{\frac{1}{4Q^2} - 1}$$

$$s = \left(-\frac{1}{2Q} \pm \sqrt{\frac{1}{4Q^2} - 1} \right) \omega_0$$

So we get real solutions whenever

$$\frac{1}{Q^2} \geq 4 \text{ or } \frac{1}{Q} \geq 2 \text{ or } Q \leq 0.5$$

And imaginary solutions if Q > 0.5. So Q = 0.5 is called *critical damping*. It corresponds to two coincident zeros, exactly as in classic Type 3 analog compensation, as per our strategy.

Beyond Critical Damping

Note that using digital control, it is easy not only to force the Q of the compensator to values greater than 0.5, but less than 0.5 too. In the latter case, the compensator would be considered over-damped. The solutions to the general second-order equation above, in the case of Q<0.5 would be considered overdamped. If Q>0.5, it is underdamped. The latter occurs the moment an imaginary part enters the solution. We then get "peakiness" (or "dippiness") as sketched in **Figure 5.15** and **Figure 5.19**.

In literature, the damping factor is often used instead of Q. It is ζ=1/2Q. So critical damping corresponds to Q=0.5 or ζ=1.

When Q>0.5, we will start getting two zero locations, corresponding to split zeros. These will be symmetrical about the 0dB-axis (above and below "sea level"). They are complex conjugates of the form a+jb and a-jb.

We will plot these out shortly for all Q.

A Useful Hint: Impedance of a Capacitor

Let us take another function that often appears in electronics in general. It is the impedance of a capacitor! If we look across its terminals we see the series combination of its series inductance (ESL), its series resistance (ESR), and its bulk capacitance (C_{OUT}). So its impedance is (contd.)…

$$Z(s) = (ESR) + \frac{1}{(C_{OUT})s} + (ESL)s$$

We have a constant (frequency-independent) term "ESR", another term which is proportional to frequency (ESL) and another which is inversely proportional to frequency ($1/C_{OUT}$). Just to be more general here, let us write this as

$$F(s) = U + \frac{V}{s} + Ws$$

In our case here, this means that U=ESR, $V = 1/C_{OUT}$ and W = ESL. The reason we are writing this out in a general form is, we see this type of function in different forms everywhere. Simplifying (to our preferred form), we get

$$F(s) = \frac{Us + V + Ws^2}{s} = \frac{\left(\frac{s}{\sqrt{V/W}}\right)^2 + \frac{U}{\sqrt{VW}} \times \left(\frac{s}{\sqrt{V/W}}\right) + 1}{s \times 1/V}$$

The denominator is the equation of an integrator (plot #1 in **Figure 5.13**). The numerator is in the (modified) form of plot #10 in Figure 5.14. The "Q" of the capacitor impedance plot can also be defined as

$$Q_{CAP\_Z} = \frac{1}{U}\sqrt{VW} \equiv \frac{1}{ESR}\sqrt{\frac{ESL}{C_{OUT}}}$$

This being a *series* RCL circuit, the Q is just the reciprocal of the parallel resonant circuit that constituted the LC post filter, where we had

$$Q_{PLANT} = R_{LOAD}\sqrt{\frac{C_{OUT}}{L}}$$

Nevertheless, the analogy is striking.

The analogy carries over to the transfer function of a Type 3 compensator too. Let us look at that too.

$$H(s) \approx \frac{s^2[C_1C_2(R_1+R_3)R_2] + s[C_2(R_1+R_3) + C_1R_2] + 1}{sR_1C_1(sC_2R_3+1)(sR_2C_3+1)}$$

We learned that if we put the two zeros coincident, we get from

$$fz1 = \frac{1}{2\pi(R_1+R_3)C_2}$$

And

$$fz2 = \frac{1}{2\pi R_2 C_1}$$

$$\frac{C_1}{C_2} = \frac{R_1 + R_3}{R_2} \Leftarrow \text{ Condition for coincident zeros}$$

Putting this in the equation for H(s) *we can eliminate C_2* entirely

$$H(s) \approx \frac{s^2\left[(C_1R_2)^2\right] + s[2C_1R_2] + 1}{sR_1C_1\left(s\frac{R_2R_3C_1}{R_1 + R_3} + 1\right)(sR_2C_3 + 1)}$$

As mentioned before, the two poles from the Type 3 function may be one too many. What we unavoidably need is the pole-at-origin (1/s term) and the two zeros (numerator). So let us move p1 and p2 out of the way for now. In digital control we can always bring in additional poles, almost at will. In our case here this is equivalent to shorting R_3 and removing C_3. See what it does to a Type 3 compensator in **Figure 5.20**. Now, assuming we have moved those two poles out of the way, we are left with a simplified Type 3 compensator transfer function

$$H(s) \approx \frac{s^2\left[(C_1R_2)^2\right] + s[2C_1R_2] + 1}{sR_1C_1} = \frac{\left(\frac{s}{1/C_1R_2}\right)^2 + \frac{1}{2}\frac{s}{1/C_1R_2} + 1}{\left(\frac{s}{1/C_1R_1}\right)}$$

Compare this with the general equation involving U, V and W, we see they are similar, except for the following key fact:

The Q of the general function is √(VW) /U and it can have any value virtually, whereas the Q of a Type 3 compensator is 0.5. As mentioned, that is a key limitation of a Type 3 analog compensator.

Other than that, based on the similarity, *we conclude that impedance of a capacitor and gain (transfer) functions of compensators can be visualized very similarly.*

If only we had some way of changing the Q of the compensator to a value of our choice, to induce Q-matching!

We also realize that there will be the equivalent of a "pole-at-origin" in the impedance plot, and it will crossover at frequency ωp0=V Hz, or equivalently fp0=V/2π. From the numerator we will get two zeros at the frequency $\omega_0 = \sqrt{(V/W)}$, or equivalently $\omega_0 = \sqrt{[(V/W)/2\pi]}$. Also the damping will depend on the factor $Q = \sqrt{(VW/U^2)}$. If this is less than or equal to 0.5, we will get a critically or over-damped response. If not, we will get imaginary solutions with corresponding peaky responses. In the latter we will get two split zeros, both on the LHP side, and symmetrical about the real axis, corresponding to the two complex conjugate solutions

$$s = \left(-\frac{1}{2Q} \pm \sqrt{\frac{1}{4Q^2} - 1} \right) \omega_0$$

In terms of our general equation, we just replace Q with $\sqrt{(VW/U^2)}$.

Let us plot these out:

a) Starting with the case of real solutions, i.e. the term in the square root sign being positive, i.e. $Q \leq 0.5$, we have two solutions. Both are along the *real* axis. They coincide when the term in the square root sign is zero, i.e. when Q = 0.5. That is "critical damping" in the context of switchers (in filter theory, critical damping is often defined as Q=0.707!). At Q = 0.5, the two zeros are located at ω_0. As Q is lowered the two zeros spread out on the real axis. Using Mathcad, we plot it out as shown in **Figure 5.24**. As Q is lowered, the two solutions spread out along the real axis. The locations are

$$\left(-\frac{1}{2Q} + \sqrt{\frac{1}{4Q^2} - 1} \right) \omega_0 \quad \text{and} \quad \left(-\frac{1}{2Q} - \sqrt{\frac{1}{4Q^2} - 1} \right) \omega_0$$

The average of these, computed *logarithmically*, is still ω_0, the location where the two zeros were coincident, for Q=0.5.

$$f_{AVG} = 10^{\frac{\log|fz_1| + \log|fz_2|}{2}} = f_0$$

b) Now if Q > 0.5, what happens? We plot this out using Mathcad too. The results are presented in **Figure 5.25**. We can see the complex conjugates, above and below "sea level" (0dB), as mentioned previously. Note that the actual location of the cusp in our real world, say observed on a scope, *correlates to the **radial** distance* away from the origin of the s-plane, and this remains fixed as the zeros separate vertically—for the simple reason that *the break frequency, as determined by the term involving "s^2", has still remained the same*.

We have now a complete visualization of how a function of the following type behaves

$$F(s) = U + \frac{V}{s} + Ws$$

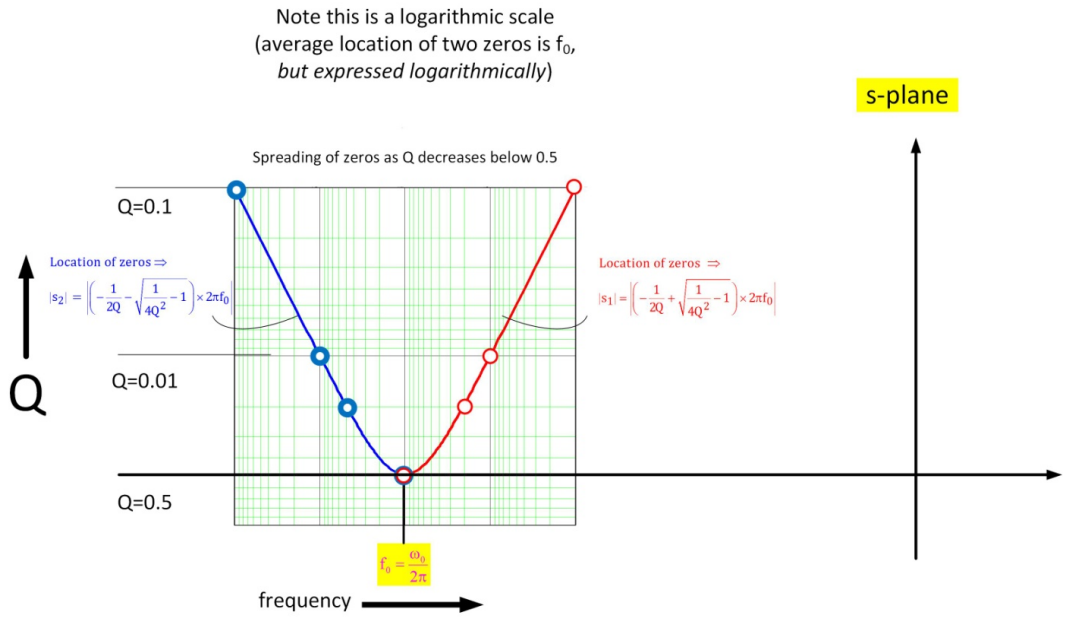


Figure 5.24: Spreading of zeros as Q falls below 0.5

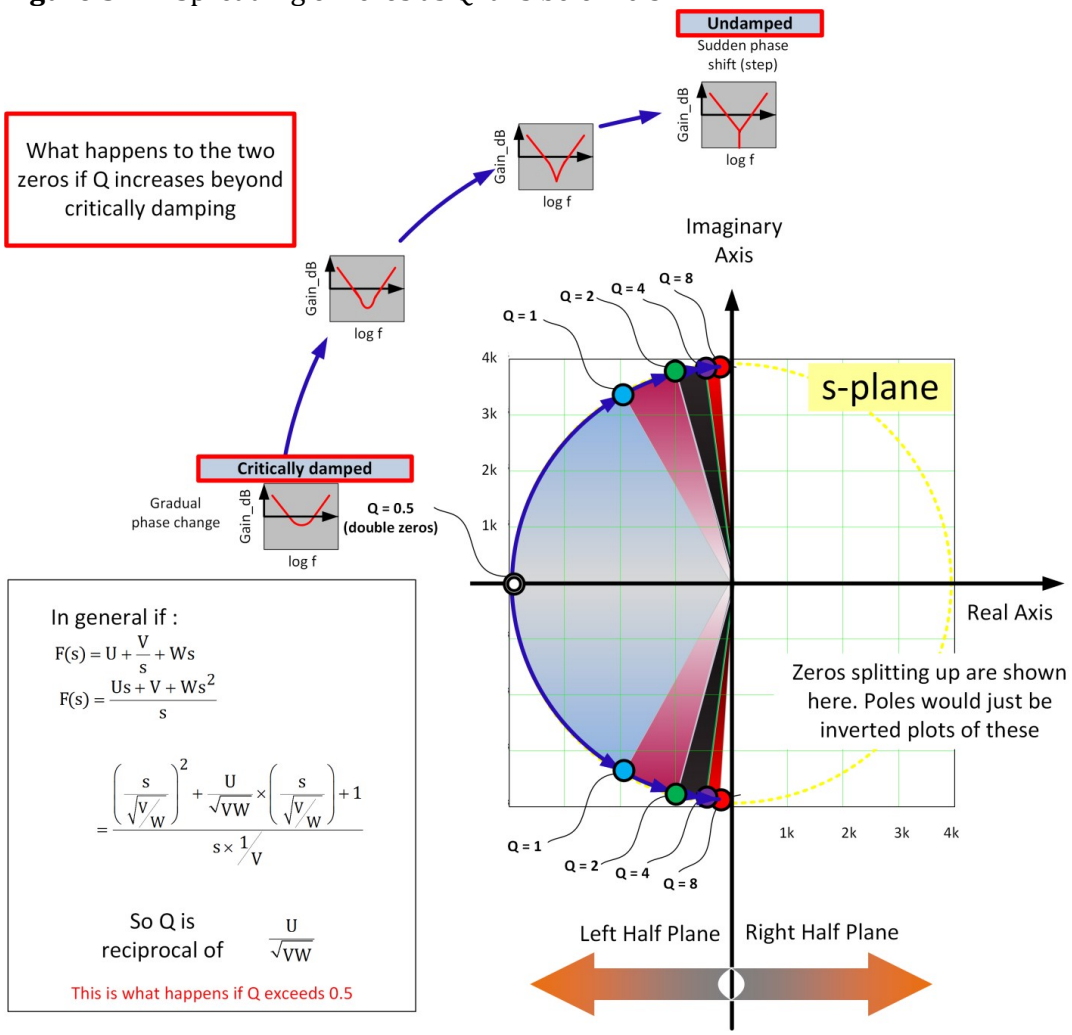


Figure 5.25: How zeros split up as Q exceeds 0.5

Introduction to PID Coefficients

Books have been written on trying to "intuitively" explain proportional, integral, derivative (PID) coefficients. Various statements have been made in literature, about the number of ways the loop can react once an error is detected. For example, it could look at the absolute error (on an instantaneous basis) and react *proportionally* to the error, or look at it over a period of time and this react to its *integral* (over time), or look how fast it changed and react to it based on its *rate of change* (derivative term), and so on. They all are correct. But as pointed out earlier, "intuition" based on mechanical systems is not very meaningful when we come to switchers. That is the reason we are actually going to resort to some *limited math*, based on our previous understanding of *the behavior of certain transfer functions*, to explain "PID coefficients" here. And the results are startling, even from a purely mathematical viewpoint, as we will now see.

In a switcher, we could theoretically combine all three types of responses in parallel, to give us a "PID compensator", as shown in **Figure 5.26**. Its behavior, using Mathcad, is typically presented in **Figure 5.27**. In the time-domain (as measured on a scope), the effect of each is often portrayed as in **Figure 5.28**. For example, a high k_I reduces the DC settling error, as talked about in Part 1. That error, we realized, was largely controlled by the integrator section! And that had a transfer function based on the $1/s$ function (plot #1 in **Figure 5.13)**. So it is no surprise to learn that the integral coefficient, k_I, is essentially the integrator section of the PID compensator!

We must emphasize that implementation of PID coefficients is possible even with analog techniques using multiple op-amps for example, but its implementation using digital techniques is far more precise and easy. Not to mention: *powerful*. Therefore, PID compensators have tended to become synonymous with digital control.
Here is the transfer (gain) function of a typical PID compensator:

$$H(s) = k_P + \frac{k_I}{s} + k_D s$$

where k_P, k_I and k_D are the "PID" (proportional-integral-derivative) coefficients, respectively. We immediately realize that the transfer function of a PID compensator is also analogous to the impedance plot of the capacitor, discussed previously. Clearly, like the impedance equation, this one is also similar to the general function we discussed above:

$$F(s) = U + \frac{V}{s} + Ws$$

And that takes us to the "grand analogy", as discovered one afternoon by Sanjaya…

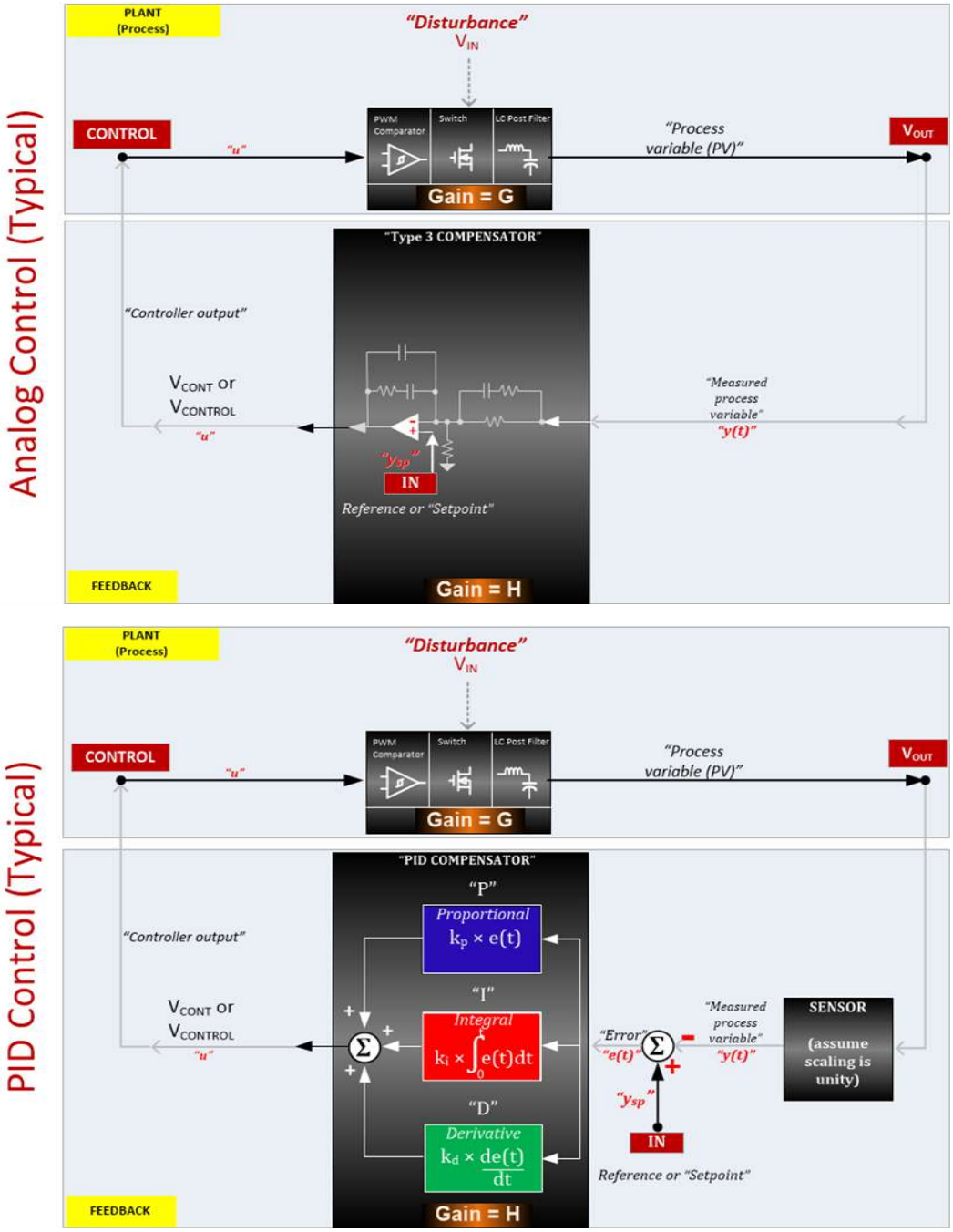


Figure 5.26: PID control

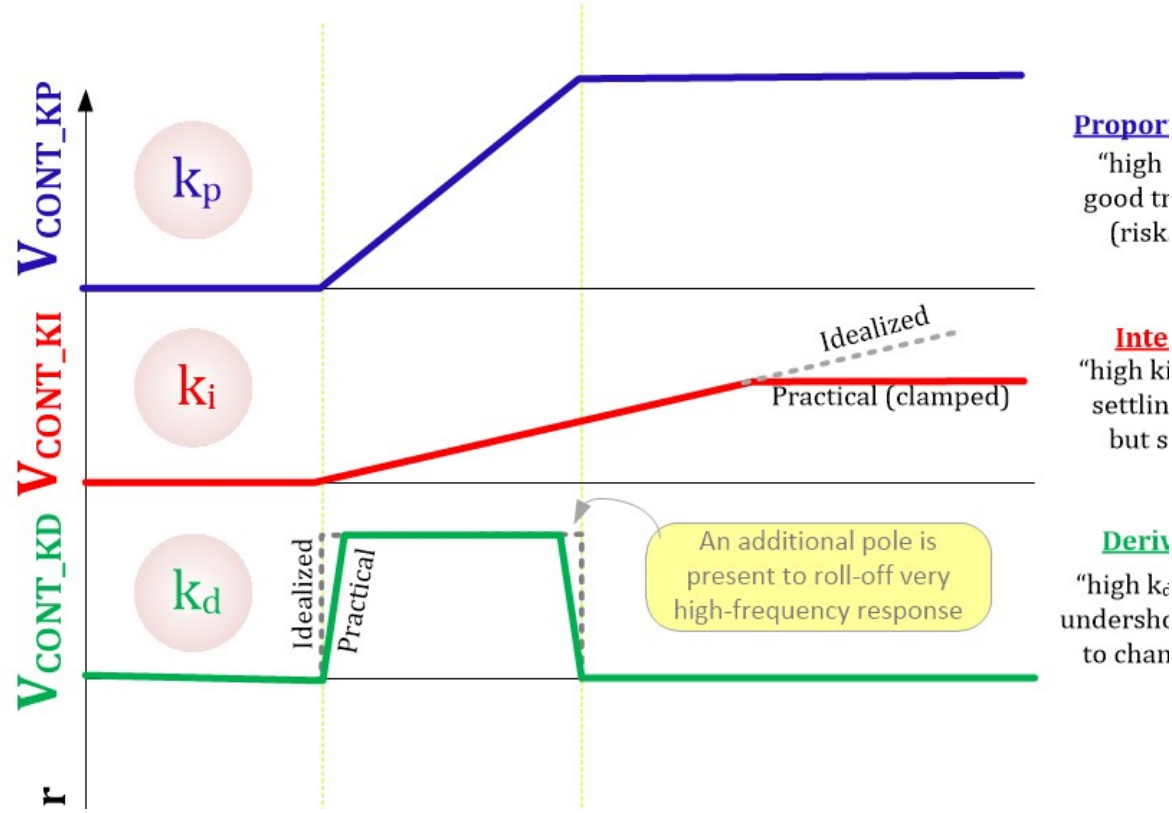


Figure 5.27: PID response characteristics

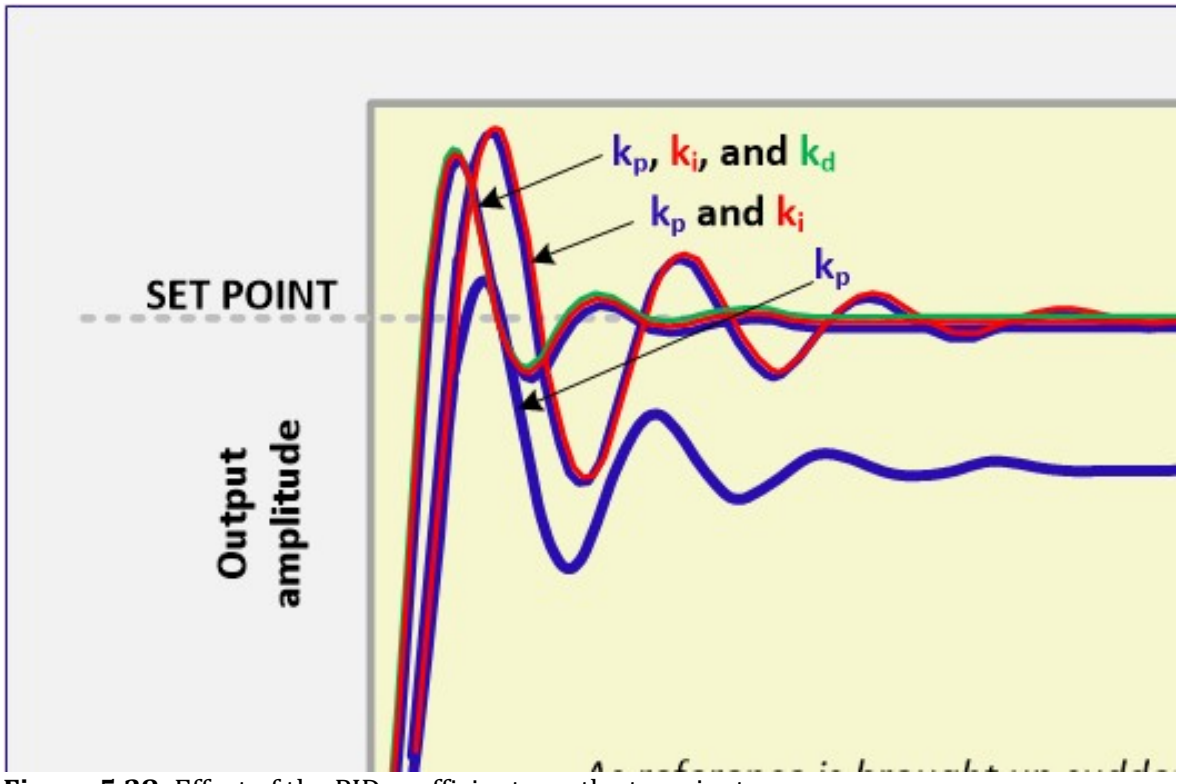


Figure 5.28: Effect of the PID coefficients on the transient response

The Grand Analogy

See **Figure 5.29** in detail now. It presents the analogy between the transfer function using basic digital (PID) compensators, and the impedance plot of a capacitor! In **Figure 5.30**, we present a numerical example to show that:

If we fix k_I and k_D, i.e. V and W, or $1/C_{OUT}$ and ESL, respectively, and only vary k_P, i.e. ESR or U, we manage to tune Q! The "dippiness" of the curve changes, but not the resonant/break frequency.

And since we can vary Q by simply varying k_P (or U), and in the process not changing the location of the "two zeros" ($\omega_0 = \sqrt{(V/W)}$, we can get the Q of the compensator ($\sqrt{(VW)}/U$) to match the Q of the plant ($R_{LOAD}\sqrt{(C_{OUT}/L)}$, this ensuring proper cancelation of the LC pole, no conditional stability concerns, and much reduced ringing in the output under large-signal transients.

This is the scheme of things we are focusing on. We have a potentially far more powerful compensation technique in our hands today, thanks to digital techniques.

In **Figure 5.31** we present the full analogy between a capacitor's impedance and a PID compensator. Such capacitor impedance curves are available in most capacitor datasheets. But we have now seen that the PID compensator is really no different in terms of *visualizing* matters! We do not need pages on building physical intuition. For once, the math speaks louder. And once we understand that math, our ability to manipulate PID coefficients, to do what our typical Type 3 analog compensator was doing, *and much more*, goes up exponentially. Digital techniques just help us immensely in implementing these compensators based on PID coefficients.

	F(s)	Z(s)	H(s)
constant	U	ESR	k_P
inv. prop. to freq.	V	$1/C_{OUT}$	k_I
prop. to freq.	W	ESL	k_D
Quality factor Q	$\frac{\sqrt{VW}}{U}$	$\frac{1}{ESR}\sqrt{\frac{ESL}{C_{OUT}}}$	$\frac{\sqrt{k_I k_D}}{k_P}$
crossover of pole-at-origin, fp0	$V/2\pi$	$1/2\pi C_{OUT}$	$k_I/2\pi$
location of (center) of two zeros f0	$(V/W)^{1/2}/2\pi$	$1/2\pi(ESL\times C_{OUT})^{1/2}$	$1/2\pi(k_D/k_I)^{1/2}$

$$Z(s) = ESR + \frac{1}{C_{OUT}\times s} + ESL\times s$$

$$F(s) = U + \frac{V}{s} + Ws$$

$$H(s) = k_P + \frac{k_I}{s} + k_D s$$

Other Definitions:

Critical damping Q = 0.5

Damping factor $\zeta = 1/2Q$

Critical damping factor $\zeta_{CRIT} = 1$

Quality factor $Q = 1/2\zeta$

Figure 5.29: The grand analogy in the form of a table

	Z(s)	F(s)	H(s)
constant	R	U	k_P
inv. prop. to freq.	1/C	V	k_I
prop. to freq.	L	W	k_D
damping factor δ	$\frac{U}{\sqrt{VW}}$	$R\sqrt{\frac{C}{L}}$	$\frac{k_P}{\sqrt{k_I k_D}}$
crossover of pole-at-origin, fp0	V/2π	1/2πC	$k_I/2\pi$
location of (center) of two zeros f0	(V/W)½ /2π	$1/2\pi(LC)^{1/2}$	$1/2\pi(k_D/k_I)^{1/2}$

$$Z(s) = R + \frac{1}{Cs} + Ls$$
$$F(s) = U + \frac{V}{s} + Ws$$
$$H(s) = k_P + \frac{k_I}{s} + k_D s$$

R is the ESR of the cap here! L is its ESL

Other Definitions:

Critical damping factor $\delta_{CRIT} = 2$

Classic damping factor $\zeta = \delta/2$

Critical classic damping factor $\zeta_{CRIT} = 1$

Quality factor $Q = 1/\delta = 1/2\zeta$

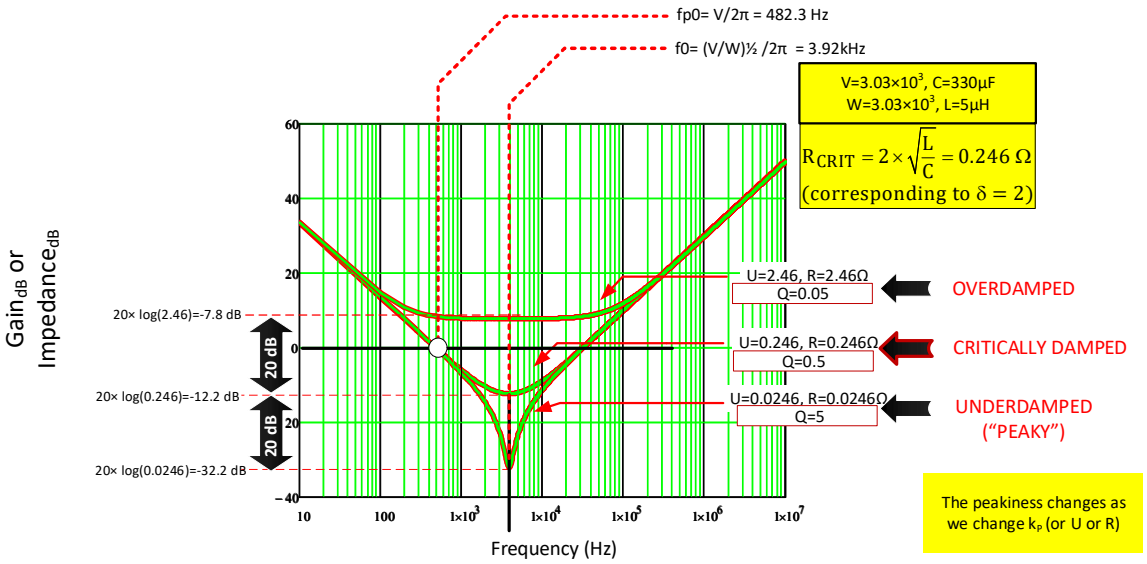


Figure 5.30: Varying the k_P (Q) only

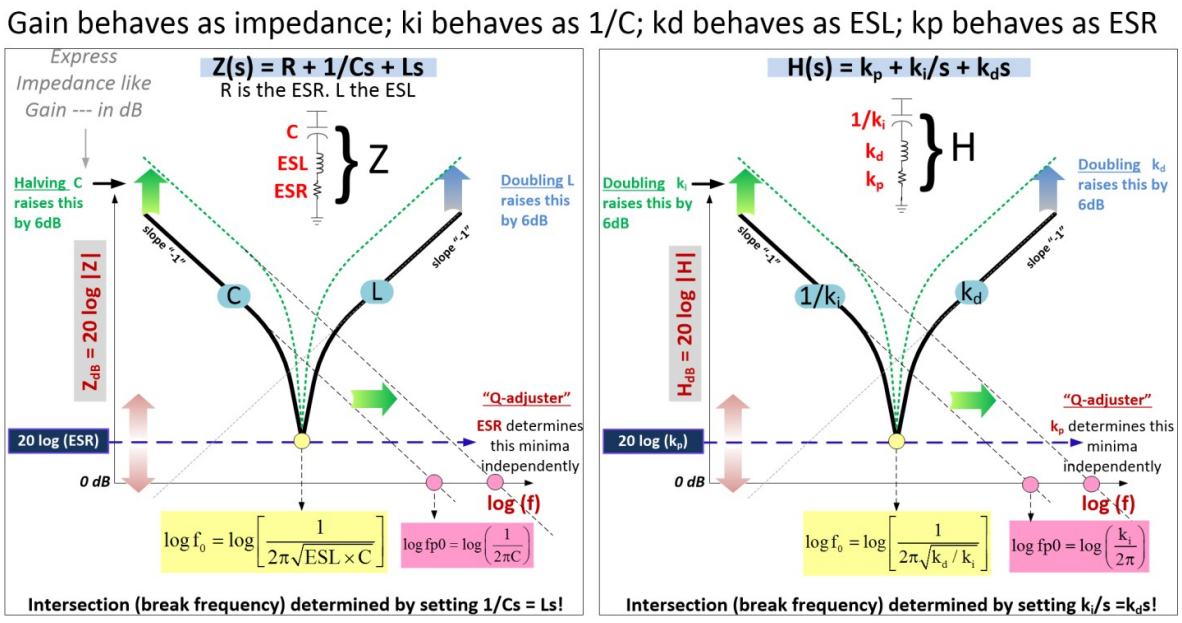


Figure 5.31: The full analogy between a capacitor impedance and a PID compensator

Finally: Q-mismatch Resolved

In **Figure 5.32** we present a summary of all the equations we need to implement Q-matching using digital compensators. First, we simply set k_I, based on the desired crossover, using the well-known equation (with k_I introduced)

$$fp0 \equiv \frac{k_I}{2\pi} = \frac{V_{RAMP}}{V_{IN}} f_{CROSS}$$

Once we know k_I, we can fix k_D, by setting the two "coincident" zeros of the PID compensator at the LC pole. So

$$f_0 = \frac{1}{2\pi\sqrt{k_D / k_I}} = f_{LC} = \frac{1}{2\pi\sqrt{LC}}$$

Then for any given load, since we know the Q of the plant, if we are sensing the load, we can appropriately set k_P using Q-matching:

$$Q_{COMP} = \frac{\sqrt{k_I k_D}}{k_P} = Q_{PLANT} = R_{LOAD}\sqrt{\frac{C_{OUT}}{L}}$$

And that sums up our technique for fine-tuning PID coefficients, to achieve excellent transient response, which was eventually validated on the bench too.

$$H_{PID} = k_p + \frac{k_i}{s} + k_d s \equiv \frac{\left[\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q_{PID}} \times \left(\frac{s}{\omega_0}\right) + 1\right]}{s/k_i}$$

(Provides a tunable double-zero and an integrator)

$$G_{LC} = \frac{1}{\left[\left(\frac{s}{\omega_{LC}}\right)^2 + \frac{1}{Q_{LC}} \times \left(\frac{s}{\omega_{LC}}\right)\right.}$$

Location of double pole

$$Q_{PID} = \frac{\sqrt{k_i \times k_d}}{k_p}$$

Q of PID compensator

$$Q_{PLANT} = R_{LOAD}\sqrt{\frac{C_{OUT}}{L_{BUCK}}}$$ Q o

$$fp0 = \frac{\omega p0}{2\pi} = \frac{k_i}{2\pi}$$

Pole at origin crossover

$$f_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi\sqrt{k_d / k_i}}$$

Location of double zero

$$H_{} = \frac{R_{LOWER}}{}$$

$$G_{PWM} = \frac{1}{V_{RAMP}}$$

Figure 5.32: Full equation set for implementing Q-matching

Bench Validation

On Sept 17, 2015, this technique was partially implemented by Sanjaya, on a vendor's latest generation digital controller (ZMDI, later IDT, then Renesas). Partially, because it was not possible to continuously feed in a k_P value based on load, given the architecture of the device.

However, we did set the k_P to tune out the LC pole almost completely at a fixed half-maximum (2.5A) load. A load transient from 0 to 2A and from 0 to 5A was carried out. In each case, the part's existing features were successively activated to see the improvement. Finally, all the existing features were *deactivated and only the PID coefficients based on our new method described above were inputted*. The overshoot/undershoot was almost a factor of 2 better than their "state of the art" achieved so far over the years. Which bears out the Q-mismatch discovery and our solution to it.

Alternative Ways to write PID coefficients

Some authors prefer to talk in terms of time constants. It is more difficult to get a feel for and manipulate. Nevertheless, for completeness sake it is mentioned here.

$$G(s) = k_p + \frac{k_i}{s} + k_d s$$

Often rewritten as

$$G(s) = k_p\left(1 + \frac{1}{s\tau_i} + s\tau_d\right)$$

$$\text{So, } k_i = \frac{k_p}{\tau_i}, k_d = k_p\tau_d$$

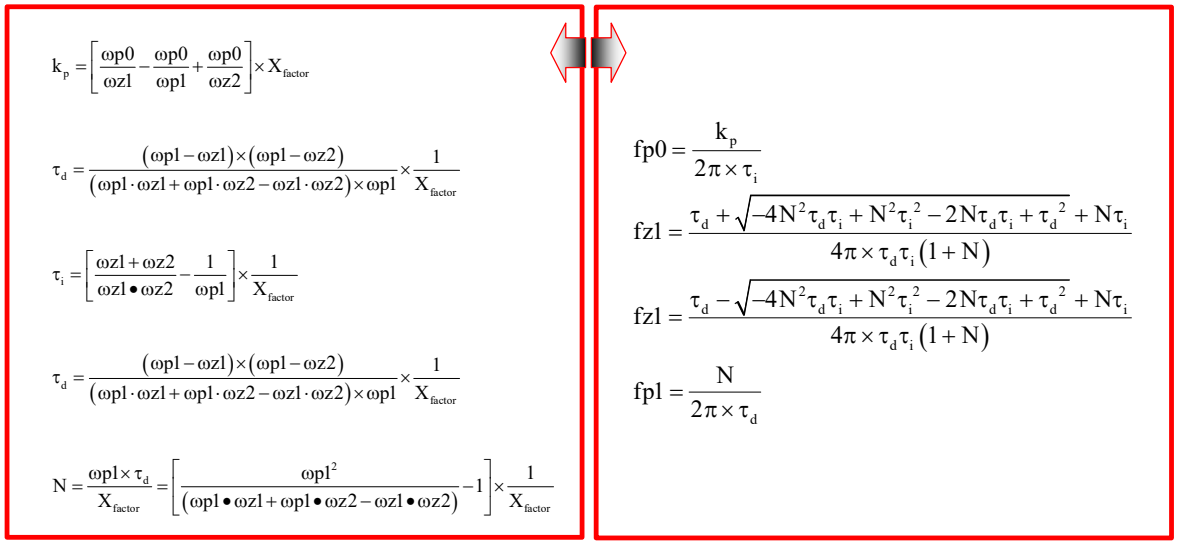
Modified by extra pole in last term

$$G(s) = k_p\left(1 + \frac{1}{s\tau_i} + \frac{s\tau_d}{1 + \frac{s\tau_d}{N}}\right)$$

So now we have an extra pole to kill the ESR-zero if we desire, or for lowering the phase margin judiciously by placing it around f_{CROSS}, as Lloyd Dixon had suggested.

The conversion tables to go back and forth between the PID coefficients and the pole-zero locations are provided in **Figure 5.33**. Note however that these equations (from Chris Basso's APEC2012 seminar, but corrected a bit) emulate a Type 3 compensator, As a result, placing the two zeros at the same location, as per our usual strategy for canceling the LC pole, leads to a Q of the compensator of around 0.5 (critical damping). We have seen that is a major limitation of standard analog compensators, so in fact this equation set fails to exploit the ability to shape the peakiness by varying k_P, as in digital compensators. To try and correct this, the X-factor was introduced. It basically multiplies the k_P of the limited equation set by that factor.

Basically, if it is set to unity, we get all the equations in related literature. However, the X-factor basically allows you to change the k_P to a new k_P, while keeping all the other PID coefficients, k_I and k_D unchanged, thereby fixing the location of the coincident zeros. In this manner we can just tweak the Q of the compensator, to match the Q of the plant, thereby ensuring Q-matching, which is clearly the secret of our excellent bench performance.



$$k_p = \left[\frac{\omega p0}{\omega z1} - \frac{\omega p0}{\omega p1} + \frac{\omega p0}{\omega z2}\right] \times X_{factor}$$

$$\tau_d = \frac{(\omega p1 - \omega z1) \times (\omega p1 - \omega z2)}{(\omega p1 \cdot \omega z1 + \omega p1 \cdot \omega z2 - \omega z1 \cdot \omega z2) \times \omega p1} \times \frac{1}{X_{factor}}$$

$$\tau_i = \left[\frac{\omega z1 + \omega z2}{\omega z1 \bullet \omega z2} - \frac{1}{\omega p1}\right] \times \frac{1}{X_{factor}}$$

$$\tau_d = \frac{(\omega p1 - \omega z1) \times (\omega p1 - \omega z2)}{(\omega p1 \cdot \omega z1 + \omega p1 \cdot \omega z2 - \omega z1 \cdot \omega z2) \times \omega p1} \times \frac{1}{X_{factor}}$$

$$N = \frac{\omega p1 \times \tau_d}{X_{factor}} = \left[\frac{\omega p1^2}{(\omega p1 \bullet \omega z1 + \omega p1 \bullet \omega z2 - \omega z1 \bullet \omega z2)} - 1\right] \times \frac{1}{X_{factor}}$$

$$fp0 = \frac{k_p}{2\pi \times \tau_i}$$

$$fz1 = \frac{\tau_d + \sqrt{-4N^2\tau_d\tau_i + N^2\tau_i^2 - 2N\tau_d\tau_i + \tau_d^2} + N\tau_i}{4\pi \times \tau_d\tau_i(1+N)}$$

$$fz1 = \frac{\tau_d - \sqrt{-4N^2\tau_d\tau_i + N^2\tau_i^2 - 2N\tau_d\tau_i + \tau_d^2} + N\tau_i}{4\pi \times \tau_d\tau_i(1+N)}$$

$$fp1 = \frac{N}{2\pi \times \tau_d}$$

Figure 5.33: Alternative representation of PID coefficients with unique X-factor method included

Conclusion

Based on the discussion so far, here is a summary of "Digital versus Analog" in control loop implementations:

a) Analog compensators are very tricky to adjust. Digital compensators can be tweaked more easily, and logically.
b) Analog compensators are very prone to component tolerances, temperature, standard-value availability and so on. Digital compensators can be set precisely.
c) Analog compensation throws up one or two more poles than perhaps necessary. We struggle to locate them correctly, without affecting the other poles and zeros, since they are all so intertwined. Digital loops allow us to introduce additional poles or zeros completely separately, and on demand.
d) Analog compensators cannot properly "kill" the LC double pole, because of the nature of the zeros they provide (limited Q). Digital compensators, *if used properly*, can enable the removal of conditional instability issues, leading to much lower output ringing during load and line transients.

We have come a long way in understanding the link between analog and digital control loops, as applied to switchers. At this stage it is perhaps unnecessary to confront the additional sampling-related issues of digital control, as reflected in the z-plane etc. We do not think that helps in either the understanding of control loops, or their implementation from a systems-level design perspective. The Q-mismatch issue seems to have been recognized, and fixed, leading to best results.

Appendix to Chapter 5

(Mathcad)

$fp0 = 1.34 \times 10^4$ $fp1 = 5.606 \times 10^5$ $fz1 = 1.186 \times 10^4$ $fz2 = 1.186 \times 10^4$ REQUIREMENTS

These are based on LC pole position, ESR zero and desired fcross. Corresponding angular frequencies

$\omega p0 := 2 \cdot \pi \cdot fp0$ $\omega z1 := 2 \cdot \pi \cdot fz1$

$\omega p1 := 2 \cdot \pi \cdot fp1$ $\omega z2 := 2 \cdot \pi \cdot fz2$

Angular frequencies

Calculation of PID coefficients based on Maniktala's design table

$ki := 2 \cdot \pi \cdot fp0$ $ki = 8.419 \times 10^4$

Use $f_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi\sqrt{k_d / k_i}}$

$kd := ki \cdot \frac{1}{4 \cdot \pi^2 \cdot fLC^2}$ $kd = 1.517 \times 10^{-5}$

Introduce our concept of Q-matching here

$Qplant := R \cdot \sqrt{\frac{C}{L}}$ $Qplant = 406.761$

More accurate form of Q is

$Qplant := \frac{\sqrt{L \cdot C}}{\left[\left[\frac{L}{R} + DCR \cdot C \cdot \left(1 + \frac{ESR}{R}\right)\right] + (ESR \cdot C)\right]}$ $Qplant = 42.282$

Desired kp is therefore $kp := \frac{\sqrt{ki \cdot kd}}{Qplant}$

If self-consistent, these are the values of the PID coefficients.

$kp = 0.027$ $ki = 8.419 \times 10^4$ $kd = 1.517 \times 10^{-5}$

Figure 5.34: Mathcad Spreadsheet for calculating PID coefficients