Differential Signals

Rules to Live By

Douglas Brooks

We generally think of signals propagating through our circuits in one of three modes, single-ended, differential mode, or common mode.

Single ended mode is the mode we are most familiar with. It involves a single wire or trace between a driver and a receiver. The signal propagates down the trace and returns through the ground system[1]

Differential mode involves a pair of traces (wires) between the driver and receiver. We typically say that one trace carries the positive signal and the other carries a negative signal that is both equal to, and the opposite polarity from, the first. Since the signals are equal and opposite, there is no return signal through ground; what travels down one trace comes back on the other.

Common mode signals are typically more difficult to understand. They may involve either single-ended traces or two (or perhaps even more) differential traces. The SAME signal travels along both the trace and its return path (ground) or along both traces in a differential pair. Most of us tend to be unfamiliar with common mode signals because we tend never to intentionally generate them ourselves. They are usually the result of noise being coupled into the circuit from some other (nearby or external) source. Generally, their consequences are neutral, at best, or damaging at worst. Common mode signals can generate noise that interrupts the operation of our circuits, and are a common source of EMI problems.

Advantages: Differential signals have one obvious disadvantage over single-ended signals. They require two traces instead of one, or twice as much board area. But there are several advantages to them.

If there is no return signal through ground, then the continuity of the ground path becomes relatively unimportant. So if we have, for example, an analog signal going to a digital device through a differential pair, we don't have to worry about crossing power boundaries, plane discontinuities, etc. Separation of power systems can be made easier with differential devices.[2]

Differential circuits can be very helpful in low signal level applications. If the signals are VERY low level, or if the signal/noise ratio is a problem, then differential signals effectively double the signal level (+v – (-v) = 2v). Differential signals and differential amplifiers are commonly used at the input stages of very low signal level systems.

Differential receivers tend to be sensitive to the *difference* in the signal levels at their inputs, but they are usually designed to be insensitive to common-mode shifts at the inputs. Therefore, differential circuits tend to perform better than single-ended ones in high noise environments.

Switching timing can be more precisely set with differential signals (referenced to each other) than with single-ended signals (referenced to a less precise reference signal subject to noise at some other point on the board.) The crossover point for a differential pair is very precisely defined (**Figure 1**). The crossover point of a single ended signal between a logical one and a logical zero (for example) is subject to noise, noise threshold, and threshold detection problems, etc.

Key Assumption: There is one very important aspect to differential signals that is frequently overlooked, and sometimes misunderstood, by engineers and designers. Let's start with the two well-known laws that (a) current flows in a closed loop and (b) current is a constant everywhere within that loop.

Consider the "positive" trace of a differential pair. Current flows down the trace and must flow in a loop, normally returning through ground. The negative signal on the other trace must also flow in a loop and would also normally return though ground. This is easy to see if we temporarily imagine a differential pair with the signal on one trace held constant. The signal on the other trace would have to return somewhere, and it seems intuitively clear that the return path would be where the single-ended trace return would be (ground). We say that, with a differential pair, there is no return through ground NOT because it can't happen, BUT because

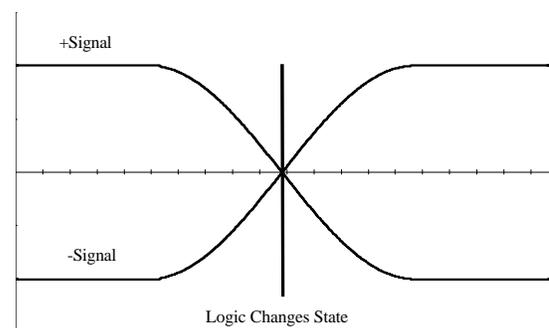+Signal

-Signal

Logic Changes State

Figure 1
Logic level changes state at the precise point where the differential signals cross over.

the returns that do exist are equal and opposite and therefore (sum to zero and) cancel each other out.

This is a VERY important point. If the return from one signal (+i) is exactly equal to, and the opposite sign from, the other signal (-i), then their SUM (+i –i) is zero, and there is no current flowing anywhere else (and in particular, though ground). Now assume the signals are not exactly equal and opposite. Let one signal be +i1 and the other be –i2 where i1 and i2 are similar, but not equal, in magnitude. The sum of their return currents is (i1 – i2). Since this is NOT zero, then this incremental current must be returning somewhere else, presumably ground.

So what, you say? Well let's assume the sending circuit sends a differential pair of signals that are exactly equal and opposite. Then we assume they will still be so at the receiving end of the path. But what if the path lengths are different? If one path (of the differential pair) is longer than the other path, then the signals are no longer equal and opposite during their transition phase at the receiver (**Figure 2**). If the signals are no longer equal and opposite during their transition from one state to another, then it is no longer true that there is no return signal through ground. If there is a return signal through ground, then power system integrity DOES become an issue, and EMI may become a problem.

Design Rule 1: This brings us to our first design guideline when dealing with differential signals: *The traces should be of equal length.*

There are some people who argue passionately against this rule. Generally, the basis for their argument involves signal timing. They point out in great detail that many differential circuits can tolerate significant differences in the timing between the two halves of a differential signal pair and still switch reliably. Depending on the logic family used, trace length difference of 500 mils can be tolerated. And

these people can illustrate these points very convincingly with parts specs and signal timing diagrams. The problem is …. they miss the point! The reason differential traces must be equal length has almost nothing to do with signal timing. It has everything to do with the assumption that differential signals are equal and opposite and what happens when that assumption is violated. And what happens is this: uncontrolled ground currents start flowing that at the very best are benign but at worst can generate serious common-mode EMI problems.

So, if you are depending on the assumption that your differential signals are equal and opposite, and that therefore there is no signal flowing through ground, a necessary consequence of that assumption is that the differential pair signal lengths ***must*** be equal.

Differential Signals and Loop Areas: If our differential circuits are dealing with signals that have slow rise times, high speed design rules are not an issue. But let's say we are dealing with fast rise time signals. What additional issues then come into play with differential traces?

Consider a design where a differential signal pair is routed across a plane from driver to receiver. Let's also assume that the trace lengths are perfectly equal and the signals are exactly equal and opposite. Therefore, there is no return current path through ground. But there **IS** an induced current on the plane, nevertheless!

Any high-speed signal can (and will) induce a coupled signal into an adjacent trace (or plane). The mechanism is exactly the same mechanism as crosstalk. It is caused by electromagnetic coupling, the combined effects of mutually inductive coupling and capacitive coupling. So, just as the return current for a single-ended signal trace tends to travel on the plane directly under the trace, a differential trace will also have an induced current on the plane underneath it.

But this is NOT a return current. All the return currents have cancelled. So this is purely a coupled noise
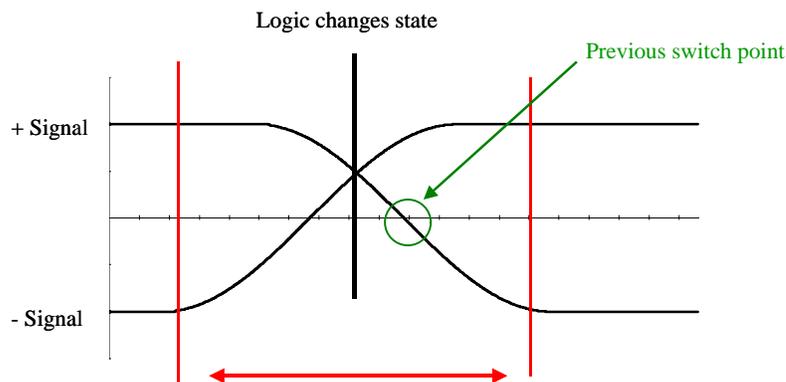

Figure 2
The (-) trace is shorter than in Figure 1, and it is no longer true that the differential signals are equal and opposite over the range indicated by the red arrow. Thus, there will be current flowing through the power system during this time frame.

current on the plane. The question is - if current must flow in a loop, where is the rest of the current flow?

Remember, we have two traces, with equal and opposite signals. One trace couples a signal on the plane in one direction, the other trace couples a signal on the plane in the other direction. These two coupled currents on the plane are equal in magnitude (assuming otherwise good design practices.) So the currents simply flow in a closed loop underneath the differential traces (**Figure 3**). They look like eddy currents. The loop these coupled currents flow in is defined by (a) the differential traces themselves, and (b) the separation between the traces at each end. The loop "area" is defined by these four boundaries.

Design Rule 2: Now it is generally known that EMI is related to loop area[3]. Therefore, if we want to keep EMI under control, we need to minimize this loop area. And the way we do that brings us to our second design rule: *route differential traces closely together*. There are people who argue against this rule, and indeed the rule is not necessary if rise times are slow and EMI is not an issue. But in high-speed environments, the closer we route the differential traces to each other, the smaller will be the loop area of the induced currents under the traces, and the better control over EMI we will have.

It is worthwhile to note that some engineers ask designers to remove the plane under differential traces. Reducing or eliminating the induced current loops under the traces is one reason for this. Another reason is to prevent any noise that might already be on the plane from coupling into the (presumably) low signal levels on the traces themselves.[4]

There is another reason to route differential traces close together. Differential receivers are designed to be sensitive to the *difference* between a pair of inputs, but also to be insensitive to a common-mode shift of those inputs. That means if the (+) input shifts even slightly in relation to the (-) input, the receiver will detect it. But if the (+) and (-) inputs shift together (in the same direction) the receiver is relatively insensitive to this shift. Therefore, if any external noise (such as EMI or crosstalk) is coupled equally into the differential traces, the receiver will be insensitive to this (common mode coupled) noise. The more closely differential traces are routed together, the more equal will any coupled noise be on each trace. Therefore, the better will be the rejection of the noise in the circuit.

Rule 2 Consequence: Again assuming a high-speed environment, if differential traces are routed closely to each other (to minimize the loop area underneath them) then the traces will couple into each other. If the traces are long enough that termination becomes an issue, this coupling impacts the calculation of the correct termination impedance[5]. Here's why:
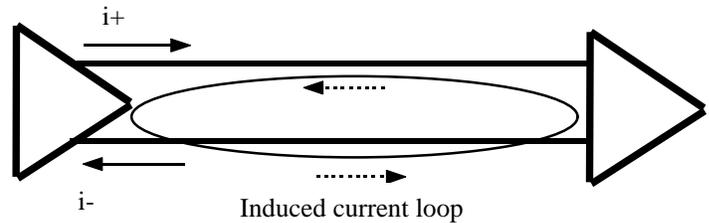

i+

i-

Induced current loop

Figure 3
Even if the differential signals are exactly equal and opposite, so that there is no return current through the power system, there will still be an induced current flowing in a closed loop on the plane under the traces.

Consider a differential pair of traces, Trace 1 and Trace 2. Let's say they carry signals V1 and V2, respectively. And since they are differential traces, V2 = -V1. V1 causes a current i1 along Trace 1 and V2 causes a current i2 along trace 2. The current necessarily is derived from Ohm's Law, $I = V/Z_o$, where Z_o is the characteristic impedance of the trace. Now the current carried by Trace 1 (for example) actually consists of i1 and also k*i2, where k is proportional to the coupling between Trace 1 and 2. It can be shown that the net effect of this coupling is an *apparent* impedance along Trace 1 equal to

$Z = Z_o – Z_{12}$

where Z_{12} is caused by the mutual coupling between Trace 1 and Trace 2 [6].

If Trace 1 and 2 are far apart, the coupling between them is very small, and the correct termination of each trace is simply Z_o, the characteristic impedance of the single-ended trace. But as the traces come closer together, and the coupling between them increases, then the impedance of the trace reduces proportional to this coupling. THAT means the proper termination of the trace (to prevent reflections) is $Z_o – Z_{12}$, or something less than Z_o. This applies to both traces in the differential pair. And since no return current flows through ground (or so it is assumed) then the terminating resisters are connected in series between Traces 1 and 2, and the correct terminating impedance is calculated as $2(Z_o – Z_{12})$. This value is often given the name "differential impedance."[7]

Design Rule 3: Differential impedance changes with coupling, which changes with trace separation. Since it is always important that the trace impedance remain constant over the entire length, this means that the coupling must remain constant over the entire length. And this leads to our third rule: *the separation between the two traces (of the differential pair) must remain constant over the entire length.*

Note that these differential impedance impacts are merely *consequences* of Design Rule 2. There is nothing really inherent about them at all. The reason we want to route differential traces close together has to do with EMI and noise immunity. The fact

that this has an impact on the correct termination of "long" traces, and this in turn has an impact on the uniformity of trace separation, is simply a consequence of routing the traces close together for EMI control.[8]

Conclusion: Differential signals have several advantages, three of which can be (a) effective isolation from power systems, (b) noise immunity, and (c) improvement in S/N ratios. Isolation from power systems (and in particular from system ground(s)) depends on the assumption that the signals on the differential traces are truly equal and opposite. This assumption may not be correct if the trace lengths of the individual traces of the differential pair are not evenly matched. Noise immunity often depends on close coupling of the traces. This, in turn, has an impact on the value of the proper termination of the traces to prevent reflections, and generally also requires that, if the traces must be close coupled, their separation must also be constant over their entire length.

Footnotes:

1. In truth the signal can return through either or both the ground OR power system. I will use the singular term "ground" throughout this article simply for convenience.
2. Optically coupled devices are another approach to solving this same type of problem.
3. See "Loop Areas: Close 'Em Tight," January, 1999
4. I know of no definitive studies that either support or refute this practice.
5. There are many references throughout the industry on impedance controlled traces. See, for example, "PCB Impedance Control: Formulas and Resources," March, 1998; "Impedance Terminations: What's the Value?" March, 1999; and "What Is Characteristic Impedance" by Eric Bogatin, January, 2000, p. 18.
6. See "Differential Impedance: What's the Difference," August, 1998
7. For an interesting discussion about how to terminate BOTH the differential mode and common mode components of a pair of traces, see "Terminating Differential Signals on PCBs," Steve Kaufer and Kellee Crisafulu, March, 1999, p. 25
8. The reason this doesn't happen with other closely routed traces, those subject to crosstalk for example, is that other traces don't have a coupling between them that is perfectly correlated --- i.e. equal and opposite. If the coupled signals are simply randomly related to each other, the *average* coupling is zero and there is no impact on the impedance termination.