The Spirit of Invention

When I was a child my grandfather routinely asked me if I was going to be an engineer when I grew up. Since some of my great-uncles worked on the railroads, I sincerely thought he wanted me to follow in their footsteps. My grandfather died before I clarified exactly what kind of engineer he hoped I would become, but I think he would approve of my interpretation.

I still wasn’t sure what an engineer was when I discovered I wanted to be an inventor. I truly pictured myself alone in my basement toiling on the important but neglected problems of humanity. Seeking help, I joined the Rocky Mountain Inventors’ Congress. They held a conference on invention where I met men carrying whole suitcases filled with clever little mechanical devices. Many of these guys were disgruntled and cranky because the world didn’t appreciate their contributions. One of the speakers, a very successful independent inventor, told of a bankrupt widow whose husband had worked twenty years in isolation and secrecy inventing a mechanical tomato peeler. The tomato peeler had consumed the family savings, and the widow had asked the speaker to salvage the device. With sadness the speaker related the necessity of informing her that tomatoes were peeled in industrial quantities with sulfuric acid. Apparently the inventor had been too narrowly focused to realize that in some cases molecules are more powerful than machines.

I didn’t want to become disgruntled, cranky, or isolated and I didn’t even own a basement. So I went to engineering school and adopted a much easier approach to inventing. I now design products for companies with such basic comforts as R&D budgets, support staff, and manufacturing operations. Along the way I have discovered many ways of nurturing inventiveness. Here are some techniques that seem to work:

Give yourself time to invent. If necessary, steal this time from the unending rote tasks that your employer so readily recognizes and rewards. I try to work on things that have nothing to do with a particular product, have no schedule, and have no one expecting results. I spend time on highly tangential ideas that have little hope for success. I can fail again and again in this daydream domain with no sense of loss.
Get excited. Enjoy the thrilling early hours of a new idea. Stay up all
night, lose sleep, and neglect your responsibilities. Freely explore tan-
gents to your new idea. Digress fearlessly and entertain the absurd.
Invent in the morning or whenever you are most energetic. Save your
“real” work for when you are tired.

Master the fundamentals of your field. The most original and creative
engineers I have known have an astonishing command of undergraduate-
level engineering. Invention in technology almost always stems from the
novel application of elementary principles. Mastery of fundamentals al-
 lows you to consider, discard, and develop numerous ideas quickly, accu-
 rately, and fairly. I believe so much in this concept that I have begun
taking undergraduate classes over again and paying very careful attention.

Honestly evaluate the utility of your new idea at the right time: late
enough not to cut off explorations of alternatives and wild notions, but
early enough that your creativity doesn’t go stale. In this stage you must
ask the hardest questions: “Is this new thing useful to anyone else? Ex-
actly where and how is it useful? Is it really a better solution or just a
clever configuration of parts?” Even if you discover that your creation
has no apparent utility, savor the fun you had exploring it and be thankful
that you don’t have the very hard work of developing it.

Creativity is not a competitive process. It is sad that we engineers are
so inculcated with the competitive approach that we use it even privately.
You must suspend this internal competition because almost all of your
new ideas will fail. This is a fact, but it doesn’t detract a bit from the fun
of inventing.

Now it’s time to get on to a very old and interesting analog design
problem where there is still a great deal of room for invention.

Requirements for Signal Conditioning
in Oscilloscopes

Most of my tenure as an electrical engineer has been spent designing
analog subsystems of digital oscilloscopes. A digital oscilloscope is a
rather pure and wholesome microcosm of signal processing and measure-
ment, but at the signal inputs the instrument meets the inhospitable real
world. The input signal-conditioning electronics, sometimes referred
to as the “front-end” of the instrument, includes the attenuators, high-
impedance buffer, and pre-amplifier. Figure 7–1 depicts a typical front-
end and is annotated with some of the performance requirements.

The combination of requirements makes the design of an oscilloscope
front-end very difficult. The front-end of a 500MHz oscilloscope devel-
op near 1GHz of bandwidth and must have a very clean step response.
It operates at this bandwidth with a 1MΩ input resistance! No significant
resonances are allowed out to 5GHz or so (where everything wants to
resonate). Because we must maintain high input resistance and low ca-
pacitance, transmission lines (the usual method of handling microwave
signals) are not allowed! The designer's only defense is to keep the physical dimensions of the circuit very small. To obtain the 1GHz bandwidth we must use microwave components. Microwave transistors and diodes are typically very delicate, yet the front-end has to withstand ±400V excursions and high-voltage electrostatic discharges. Perhaps the most difficult requirement is high gain flatness from DC to a significant fraction of full bandwidth.

A solid grasp of the relationships between the frequency and time domains is essential for the mastery of these design challenges. In the following I will present several examples illustrating the intuitive connections between the frequency magnitude and step responses.

The Frequency and Time Domains

Oscilloscopes are specified at only two frequencies: DC and the -3dB point. Worse, the manufacturers usually state the vertical accuracy at DC only, as if an oscilloscope were a voltmeter! Why is a time domain measuring device specified in the frequency domain? The reason is that bandwidth measurements are traceable to international standards, whereas it is extremely difficult to generate an impulse or step waveform with known properties (Andrews 1983, Rush 1990).

Regardless of how oscilloscopes are specified, in actual practice oscilloscope designers concern themselves almost exclusively with the step response. There are several reasons for focusing on the step response: (1) a good step response is what the users really need in a time domain instrument, (2) the step response conveys at a glance information about a very wide band of frequencies, (3) with practice you can learn to intuitively relate the step response to the frequency response, and (4) the step
response will be used by your competitors to find your weaknesses and attack your product.

Figure 7–2 defines the terms of the frequency and step responses and shows the meaning of flatness error. Response flatness is a qualitative notion that refers roughly to gain errors not associated with the poles that determine the cutoff frequency, or equivalently to step response errors following the initial transition. To assess flatness we generally ignore peaking of the magnitude near the 3dB frequency. We also ignore short-term ringing caused by the initial transition in the step response.

Figure 7–2 illustrates the rough correspondence between the high-frequency portions of the magnitude response and the early events in the step response. Similarly, disturbances in the magnitude response at low frequencies generate long-term flatness problems in the step response.
(Kamath 1974). Thus the step response contains information about a very wide band of frequencies, when observed over a long enough time period. For example, looking at the first ten nanoseconds (ns) of the step conveys frequency domain information from the upper bandwidth of the instrument down to approximately $1/(10\text{ns})$ or 100MHz.

Figure 7–3 shows an RC circuit that effectively models most sources of flatness errors. Even unusual sources of flatness errors, such as dielectric absorption and thermal transients in transistors, can be understood with similar RC circuit models. The attenuator and impedance converter generally behave like series and parallel combinations of simple RC circuits. Circuits of this form often create flatness problems at low frequencies because of the high resistances in an oscilloscope front-end. In contrast, the high-frequency problems are frequently the result of the innumerable tiny inductors and inadvertent transmission lines introduced in the physical construction of the circuit. Notice how in Figure 7–3 the reciprocal nature of the frequency and step responses is well represented.

### High Impedance at High Frequency: The Impedance Converter

Oscilloscopes by convention and tradition have $1\Omega$ inputs with just a few picofarads of input capacitance. The $1\Omega$ input resistance largely determines the attenuation factor of passive probes, and therefore must be accurate and stable. To maintain the accuracy of the input resistance, the oscilloscope incorporates a very high input impedance unity gain buffer (Figure 7–1). This buffer, sometimes called an “impedance converter,” presents more than 100GHz at its input while providing a low-impedance, approximately 50Q output to drive the pre-amp. In a 500MHz oscilloscope the impedance converter may have 1GHz of bandwidth and very carefully controlled time domain response. This section

---

**Figure 7–3.** A simple circuit that models most sources of flatness errors.
shows one way in which these and the many additional requirements of Figure 7–1 can be met (Rush 1986).

A silicon field effect transistor (FET) acting as a source follower is the only type of commercially available device suitable for implementing the impedance converter. For 500MHz instruments, we need a source follower with the highest possible transconductance combined with the lowest gate-drain capacitance. These parameters are so important in a 500MHz instrument that oscilloscope designers resort to the use of short-channel MOSFETs in spite of their many shortcomings. MOSFETs with short channel lengths and thin gate oxide layers develop very high transconductance relative to their terminal capacitances. However, they suffer from channel length modulation effects which give them undesirably high source-to-drain or output conductance. MOSFETs are surface conduction devices, and the interface states at the gate-to-channel interface trap charge, generating large amounts of 1/f noise. The 1/f noise can contribute as much noise between DC and 1MHz as thermal noise between DC and 500MHz. Finally, the thin oxide layer of the gate gives up very easily in the face of electrostatic discharge. As source followers, JFETs outperform MOSFETs in every area but raw speed. In summary, short-channel MOSFETs make poor but very fast source followers, and we must use a battery of auxiliary circuits to make them function acceptably in the impedance converter.

Figure 7–4 shows a very basic source follower with the required 1MΩ input resistance. The resistor in the gate stabilizes the FET. Figure 7–5 shows a linear model of a typical high-frequency, short-channel MOSFET. I prefer this model over the familiar hybrid-π model because it shows at a glance that the output resistance of the source is 1/gm. Figure 7–6 shows the FET with a surface-mount package model. The tiny capacitors and inductors model the geometric effects of the package and the surrounding environment. These tiny components are called “parasitics” in honor of their very undesirable presence. Figure 7–7 depicts the parasitics of the very common “0805” surface-mount resistor. This type of resistor is often used in front-end circuits built on printed circuit boards. Package and circuit board parasitics at the 0.1pF and 1nH level seem negligibly small, but they dominate circuit performance above 500MHz.
In oscilloscope circuits I often remove the ground plane in small patches beneath the components to reduce the capacitances. One must be extremely careful when removing the ground plane beneath a high-speed circuit, because it always increases parasitic inductance. I once turned a beautiful 2GHz amplifier into a 400MHz bookend by deleting the ground plane and thereby effectively placing large inductors in the circuit.

Figure 7-5.
A linear model of a BSD22, a typical high-frequency, short-channel MOSFET. The gate current is zero at DC because the controlled current source keeps the drain current equal to the source current.

Figure 7-6.
A MOSFET with SOT-143 surface-mount package parasitics. The model includes the effects of mounting on a 1.6mm (0.063") thick, six-layer epoxy glass circuit board with a ground plane on the fourth layer from the component side of the board.
Parasitics have such a dominant effect on high-frequency performance that 500MHz oscilloscope front-ends are usually built as chip-and-wire hybrids, which have considerably lower parasitics than standard printed circuit construction. Whether on circuit boards or hybrids, the bond wires, each with about 0.5 to 1.0nH inductance, present one of the greatest difficulties for high-frequency performance. In the course of designing high-frequency circuits, one eventually comes to view the circuits and layouts as a collection of transmission lines or the lumped approximations of transmission lines. I have found this view to be very useful and with practice a highly intuitive mental model.

Figure 7–8 shows the magnitude and step responses of the simple source follower, using the models of Figures 7–5 through 7–7. The bandwidth is good at 1.1GHz. The rise time is also good at 360ps, and the 1% settling time is under 1ns!

Our simple source follower still has a serious problem. The high drain-to-source conductance of the FET forms a voltage divider with the source resistance, limiting the gain of the source follower to 0.91. The pre-amp could easily make up this gain, but the real issue is temperature stability. Both transconductance and output conductance vary with temperature, albeit in a self-compensating way. We cannot comfortably rely on this self-compensation effect to keep the gain stable. The solution is to bootstrap the drain, as shown in Figure 7–9. This circuit forces the drain and source voltages to track the gate voltage. With bootstrapping, the source follower operates at nearly constant current and nearly constant terminal voltages. Thus bootstrapping keeps the gain high and stable, the power dissipation constant, and the distortion low.

There are many clever ways to implement the bootstrap circuit (Kimura 1991). One particularly simple method is shown in Figure 7–10. The BF996S dual-gate, depletion-mode MOSFET is intended for use in television tuners as an automatic gain controlled amplifier. This device acts like two MOSFETs stacked source-to-drain in series. The current source shown in Figure 7–10 is typically a straightforward bipolar transistor current source implemented with a microwave transistor. An ap-
proximate linear model of the BF996S is shown in Figure 7–11. The BF996S comes in a SOT-143 surface-mount package, with parasitics, as shown in Figure 7–6.

Figure 7–12 shows the frequency and step responses of the bootstrapped source follower. The bootstrapping network is AC coupled, so
Signal Conditioning in Oscilloscopes and the Spirit of Invention

Figure 7-10. Bootstrapping the drain with a dual-gate MOSFET.

Figure 7-11. Linear model of the BF996S dual-gate, depletion MOSFET.

it does not boost the gain at DC and low frequencies. The response therefore is not very flat, but we can fix it later. From 1kHz to 100MHz the gain is greater than 0.985 and therefore highly independent of temperature. The 1% settling time is very good at 1.0ns.

Several problems remain in the bootstrapped source follower of Figure 7-10. First, the gate has no protection whatever from overvoltages and electrostatic discharges. Second, the gate-source voltage will vary drastically with temperature, causing poor DC stability. Third, the 1/f noise of the MOSFET is uncontrolled. The flatness (Figure 7-12) is very poor indeed. Finally, the bootstrapped source follower has no ability to handle large DC offsets in its input.

Figure 7-13 introduces one of many ways to build a “two-path” impedance converter that solves the above problems (Evel 1971, Tektronix 1972). DC and low frequencies flow through the op amp, whereas high frequencies bypass the op amp via C1. At DC and low frequencies, feed-
Mid-band gain is 0.9875
Low freq. gain is 0.904

Figure 7-12.
The magnitude and step responses of the bootstrapped source follower.

Figure 7-14 shows the flatness details of the two-path impedance converter. Feedback around the op amp has taken care of the low-frequency gain error exhibited by the bootstrapped source follower (Figure 7-12). The gain is flat from DC to 80MHz to less than 0.1%. The “wiggle” in the magnitude response occurs where the low- and high-frequency paths cross over.

There are additional benefits to the two-path approach. It allows us to design the high-frequency path through C1 and the MOSFET without regard to DC accuracy. The DC level of the impedance converter output is independent of the input and can be tailored to the needs of the preamplifier. Although it is not shown in the figures, AC coupling is easily implemented by blocking DC to the non-inverting input of the op amp.
Thus we avoid putting an AC coupling relay, with all its parasitic effects, in the high-frequency path.

There are drawbacks to the two-path impedance converter. The small flatness errors shown in Figure 7–14 never seem to go away, regardless of the many alternative two-path architectures we try. Also, C1 forms a capacitive voltage divider with the input capacitance of the source follower. Along with the fact that the source follower gain is less than unity, this means that the gain of the low-frequency path may not match that of the high-frequency path. Component variations cause the flatness to vary further. Since the impedance converter is driven by a precision high-impedance attenuator, it must have a very well-behaved input impedance that closely resembles a simple RC parallel circuit. In this regard the most common problem occurs when the op amp has insufficient speed and fails to bootstrap R1 in Figure 7–13 to high enough frequencies.
The overdrive recovery performance of a two-path amplifier can be abysmal. There are two ways in which overdrive problems occur. If a signal is large enough to turn on one of the protection diodes, C1 charges very quickly through the low impedance of the diode (Figure 7-13). As if it were not bad enough that the input impedance in overdrive looks like 270pF, recovery occurs with a time constant of 270pF · 4.7MΩ, or 1.3ms! Feedback around the op amp actually accelerates recovery somewhat but recovery still takes eons compared to the 400ps rise time! Another overdrive mechanism is saturation of the source follower. When saturation occurs, the op amp integrates the error it sees between the input and source follower output, charging its 6.8nF feedback capacitor. Recovery occurs over milliseconds. The seriousness of these overdrive recovery problems is mitigated by the fact that with careful design it can take approximately ±2V to saturate the MOSFET and ±5V to activate the protection diodes. Thus, to overdrive the system, it takes a signal about ten times the full-scale input range of the pre-amp.

I apologize for turning a simple, elegant, single transistor source follower into the “bootstrapped, two-path impedance converter.” But as I stated at the beginning, it is the combination of requirements that drives us to such extremes. It is very hard to meet all the requirements at once with a simple circuit. In the next section, I will extend the two-path technique to the attenuator to great advantage. Perhaps there the two-path method will fully justify its complexity.

The Attenuator

I have expended a large number of words and pictures on the impedance converter, so I will more briefly describe the attenuator. I will confine myself to an introduction to the design and performance issues and then illustrate some interesting alternatives for constructing attenuators. The purpose of the attenuator is to reduce the dynamic range requirements placed on the impedance converter and pre-amp. The attenuator must handle stresses as high as ±400V, as well as electrostatic discharge. The attenuator maintains a 1MΩ input resistance on all ranges and attains microwave bandwidths with excellent flatness. No small-signal microwave semiconductors can survive the high input voltages, so high-frequency oscilloscope attenuators are built with all passive components and electromechanical relays for switches.

Figure 7-15 is a simplified schematic of a 1MΩ attenuator. It uses two stages of the well-known “compensated voltage divider” circuit. One stage divides by five and the other by 25, so that division ratios of 1, 5, 25, and 125 are possible. There are two key requirements for the attenuator. First, as shown in Figure 7-3, we must maintain \( R_1C_1 = R_2C_2 \) in the \( \pm 5 \) stage to achieve a flat frequency response. A similar requirement holds for the \( \pm 25 \) stage. Second, the input resistance and capacitance at each stage must match those of the impedance converter and remain very
nearly constant, independent of the switch positions. This requirement assures that we maintain attenuation accuracy and flatness for all four combinations of attenuator relay settings.

Dividing by a high ratio such as 125 is similar to trying to build a high-isolation switch; the signal attempts to bypass the divider, causing feed-through problems. If we set a standard for feedthrough of less than one least-significant bit in an 8-bit digital oscilloscope, the attenuator must isolate the input from the output by \(20\log_{10}(125 \cdot 2^8) = 90\text{dB}\). I once spent two months tracking down such an isolation problem and traced it to wave guide propagation and cavity resonance at 2GHz inside the metallic attenuator cover.

Relays are used for the switches because they have low contact impedance, high isolation, and high withstanding voltages. However, in a realm where 1mm of wire looks like a transmission line, the relays have dreadful parasitics. To make matters worse, the relays are large enough to spread the attenuator out over an area of about 2 × 3cm. Assuming a propagation velocity of half the speed of light, three centimeters takes 200ps, which is dangerously close to the 700ps rise time of a 500MHz oscilloscope. In spite of the fact that I have said we can have no transmission lines in a high-impedance attenuator, we have to deal with them anyway! To deal with transmission line and parasitic reactance effects, a real attenuator includes many termination and damping resistors not shown in Figure 7–15.

Rather than going into extreme detail about the conventional attenuator of Figure 7–15, it would be more interesting to ask if we could somehow eliminate the large and unreliable electromechanical relays. Consider the slightly different implementation of the two-path impedance converter depicted in Figure 7–16. The gate of the depletion MOSFET is self-biased by the 22MΩ resistor so that it operates at zero gate source voltage. If the input and output voltages differ, feedback via the op amp and bipolar current source reduces the error to zero. To understand this circuit, it helps to note that the impedance looking into the source of a self-biased FET is very high. Thus the collector of the bipolar current source sees a
high-impedance load. Slight changes in the op amp output can therefore produce significant changes in the circuit output.

The impedance converter of Figure 7–16 can easily be turned into a fixed attenuator, as shown in Figure 7–17. As before, there is a high-frequency and a low-frequency path, but now each divides by ten. There is an analog multiplier in the feedback path to make fine adjustments to the low-frequency gain. The multiplier matches the low- and high-frequency paths to achieve a high degree of flatness. A calibration procedure determines the appropriate gain for the multiplier.

Now we can build a complete two-path attenuator with switched attenuation, as shown in Figure 7–18 (Roach 1992). Instead of cascading attenuator stages, we have arranged them in parallel. In place of the two double-pole double-throw (DPDT) relays of Figure 7–15, we now need only two single-pole single-throw (SPST) relays. Note that there is no need for a switch in the $\div 100$ path because any signal within range for
A two-path attenuator and impedance converter using only two SPST electromechanical relays. The protection diodes and some resistors are omitted for clarity.

The \( \pm 1 \) or \( \pm 10 \) path is automatically in range for the \( \pm 100 \) path. The switches in the low-frequency feedback path are not exposed to high voltages and therefore can be semiconductor devices.

A number of advantages accrue from the two-path attenuator of Figure 7-18. The SPST relays are simpler than the original relays, and the high-frequency path is entirely AC coupled! The relays could be replaced with capacitive switches, eliminating the reliability problems of DC contacts. One of the most important contributions is that we no longer have to precisely trim passive components as we did in Figure 7-15 to make \( R_1C_1 = R_2C_2 \). This feature eliminates adjustable capacitors in printed circuit (PC) board attenuators and difficult laser trimming procedures on hybrids. With the need for laser trimming eliminated, we can build on inexpensive PC board attenuators that formerly required expensive hybrids.
We can take the new attenuator configuration of Figure 7–18 further. First observe that we can eliminate the +10 relay in Figure 7–18, as shown in Figure 7–19. The diodes are reverse biased to turn the +10 path on and forward biased to turn it off. Forward biasing the diodes shorts the 1pF capacitor to ground, thereby shunting the signal and cutting off the +10 path. The input capacitance changes by only 0.1pF when we switch the +10 path.

Now we are down to one electromechanical relay in the +1 path. We can eliminate it by moving the switch from the gate side of the source follower FET to the drain and source, as shown in Figure 7–20. In doing so we have made two switches from one, but that will turn out to be a good trade. With the +1 switches closed, the drain and source of the FET are connected to the circuit and the +1 path functions in the usual manner. The protection diodes are biased to ±5V to protect the FET.

To cut off the +1 path, the drain and source switches are opened, leaving those terminals floating. With the switches open, a voltage change at
the input drives the gate, source, and drain of the FET through an equal change via the 20pF input capacitor and the gate-drain and gate-source capacitances. Since all three terminals of the FET remain at the same voltage, the FET is safe from overvoltage stress. Of course, the switches must have very low capacitance in the open state, or capacitive voltage division would allow the terminals of the FET to see differing voltages. In ±100 mode, the floating FET will see 40V excursions (eight divisions on the oscilloscope screen at 5V per division) as a matter of course. For this reason the ±1 protection diodes must be switched to a higher bias voltage (±50V) when in the ±10 and ±100 modes. The switches that control the voltage on the protection diodes are not involved in the high-frequency performance of the front-end and therefore can be implemented with slow, high-voltage semiconductors.

Can we replace the switches in the drain and source with semiconductor devices? The answer is yes, as Figure 7–21 shows. The relays in the drain and source have been replaced by PIN diodes. PIN diodes are made with a p-type silicon layer (P), an intrinsic or undoped layer (I), and an n-type layer (N). The intrinsic layer is relatively thick, giving the diode high breakdown voltage and extremely low reverse-biased capacitance. A representative packaged PIN diode has 100V reverse breakdown and only 0.08pF junction capacitance. To turn the ±1 path of Figure 7–21 on, the switches are all set to their “±1” positions. The PIN diodes are then forward biased, the bipolar transistor is connected to the op amp, and the FET is conducting. To turn the path off, the switches are set to their “±10,100” positions, reverse-biasing the PIN diodes. Since these switches

![Figure 7-21. Using PIN diodes to eliminate the relays in the ±1 path.](image-url)
are not involved in the high-frequency signal path, they too can be built with slow, high-voltage semiconductors.

The complete circuit is now too involved to show in one piece on the page of a book, so please use your imagination. We have eliminated all electromechanical switches and have a solid-state oscilloscope front-end. Although I had a great deal of fun inventing this circuit, I do not think it points the direction to future oscilloscope front-ends. Already research is under way on microscopic relays built with semiconductor micro-machining techniques (Hackett 1991). These relays are built on the surface of silicon or gallium arsenide wafers, using photolithography techniques, and measure only 0.5mm in their largest dimension. The contacts open only a few microns, but they maintain high breakdown voltages (100s of volts) because the breakdown voltages of neutral gases are highly nonlinear and not even monotonic for extremely small spacing. The contacts are so small that the inter-contact capacitance in the open state is only a few femtofarads (a femtofarad is 0.001 picofarads). Thus the isolation of the relays is extraordinary! Perhaps best of all, they are electrostatically actuated and consume near zero power. I believe micro-machined relays are a revolution in the wings for oscilloscope front-ends. I eagerly anticipate that they will dramatically improve the performance of analog switches in many applications. Apparently, even a device as old as the electromechanical relay is still fertile ground for a few ambitious inventors!

References


