TRANSMISSION LINES AND IMPEDANCE MATCHING

APPENDIX **H**

H.1 Some properties of transmission lines

In §12.9 we introduced transmission lines, which most commonly take the form of *coaxial* cable ("coax"), for example the ubiquitous "BNC cables" (RG-58 cables with male BNC connectors at each end) that are used to run all manner of signals between instruments. As we remarked there, for low-frequency applications it is common (and correct) to think of such a cable simply as a well-shielded wire with ∼30 pF/ft of capacitance. However, at high frequencies (say those for which the cable is at least 1/20 of the wavelength) the behavior is fundamentally different: as a bizarre example, an open-ended cable ironically looks like a *short circuit* at a frequency for which the cable's electrical length is $\lambda/4$. For a 5-foot length of coax, that happens at about 32 MHz. An important consequence is that you can't just hook such a BNC cable from a signal generator to some high-impedance circuit under test and assume that it will provide a nice signal source at the circuit's input; instead you will see huge dips and bumps as you tune the frequency, because the generator sees a load impedance that varies from a short circuit (at odd multiples of 32 MHz) to an open circuit (at even multiples of 32 MHz). Perhaps surprisingly, if you were instead to connect a resistor of exactly 50 Ω across the circuit end of the cable, you would find that it now delivered a constant signal amplitude (equal to half the signal generator's open-circuit output amplitude) as you varied the frequency. This nonintuitive behavior is nicely illustrated in the measured data shown in Figure H.1. And even more nonintuitively, at the driving end of such a "terminated" cable the capacitance disappears entirely – you see instead a pure resistive load of 50 Ω!

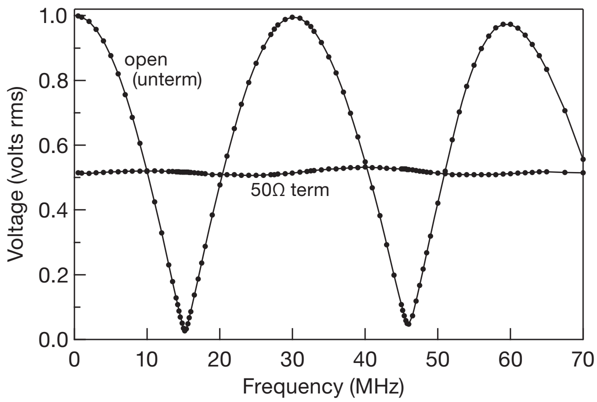

Figure H.1. Measured amplitude at the output connector of a sinewave oscillator of 1 V amplitude (open circuit), under two conditions: driving 10 feet of RG-58 (50 Ω) coaxial cable, open at the far end; and driving the same cable with a 50 Ω resistor connected across the far end.

H.1.1 Characteristic impedance

This simple example illustrates the importance of *termination*: coaxial cable is a form of *transmission line*, with a *characteristic impedance* Z_0 (which is always real: a resistance) that depends on only its physical construction:

$$Z_0 = \sqrt{L/C} = \frac{138}{\sqrt{\varepsilon}} \log_{10} \frac{b}{a} \text{ ohms},$$

where L and C are the inductance and capacitance per unit length, which as indicated depend on only the outer diameter, a, of the inner conductor, the inner diameter, b, of the outer conductor, and the dielectric constant, ε (relative to free space), of the insulating material that separates them. For a wave propagating along a transmission line, Z_0 is the ratio of signal voltage to signal current. The most popular coax line for general purposes is RG-58, with an impedance of 50 Ω (its dimensions are a=0.81 mm, b=2.95 mm, and ε=2.3, for which the above equation gives Z_0=51 Ω). This impedance has become the standard for radiofrequency use, except for video applications where the standard is Z_0=75 Ω; the corresponding popular cable type is called RG-59. In pulse electronics you sometimes see 93 Ω cable (RG-62).

The signal propagates along the cable at a speed

$$v_{\text{wave}} = \frac{c}{\sqrt{\varepsilon}} = \frac{1}{\sqrt{LC}},$$

which is a fraction $1/\sqrt{\varepsilon}$ times c (where c is the speed of light in vacuum, 3×10^8 m/s). The factor $1/\sqrt{\varepsilon}$ is called the *velocity factor*, and ranges from 0.66 (solid polyethylene) to 0.80 (polyethylene foam) for available flexible coaxial cables. In the absence of a dielectric the velocity factor is 1.0, i.e., waves travel at the speed of light in an air-spaced coaxial line. The "electrical length" seen by a propagating signal in a cable of physical length L is larger by the factor $\sqrt{\varepsilon}$, i.e., $L_{\text{elec}}{=}L_{\text{physical}}\sqrt{\varepsilon}$.

Note that the inductance and capacitance of the cable cannot take on any old values, because they are constrained such that their product LC is related to the speed of light. From this it is easy to show that if you know the characteristic impedance and the velocity factor then you can find the capacitance per unit length (or vice versa) by

$$C = \frac{\sqrt{\varepsilon}}{cZ_0}\ \text{Farads/meter}.$$

For example, RG-8 has a specified impedance of 52 Ω, and a velocity factor of 0.66; the above equation then gives C=97.1 pF/m, or 29.6 pF/ft, in good agreement with the specified value of 29.5 pF/ft.

A. Twisted pair and PCB traces

Transmission lines are not *required* to be of coaxial geometry. An extremely popular form of contemporary transmission line is the *twisted pair*, which is just what it sounds like: a pair of insulated wires, gently twisted, and enclosed in an overall insulating jacket (often without any overall shield conductor). These may well be the dominant species of transmission line in our time, because they are the basic stuff of local area networks (LANs). You usually see four twisted pairs bundled into a single unshielded jacket; this is called "UTP" (unshielded twisted pair), and is the most common form of LAN cable. It is available also with a shield (shielded twisted pair, "STP"). Contemporary UTP and STP cables are 100 Ω nominal impedance, and are characterized (in terms of impedance and attenuation) for operation up to 10 Mbps (megabits per second) or 100 Mbps; these are called Category 3 and Category 5, respectively, and appear to differ primarily in the pitch of the twist. Those in the know refer to these casually as cat-3 and cat-5.[1] Ethernet LANs using these data rates are called "10baseT" and "100baseTX," the "T" standing for "twisted"; the corresponding designation for "thinnet" coaxial Ethernet is "10base2."

In high-speed electronics on a printed circuit board it's often necessary to treat connection traces as transmission lines. See the discussion in §*1x.1*, where we described the *microstrip* transmission line, which consists of a thin conducting strip on an outer printed circuit layer, with an underlying ground-plane layer. A popular variant adds a pair of ground-trace shepherds on either side – that's called a *grounded coplanar waveguide* (GCPW). There's also the completely enclosed *stripline* geometry, where the trace(s) are sandwiched between ground-plane layers.

H.1.2 Termination: pulses

A transmission line at low frequencies (wavelength much longer than the cable length) looks simply like a capacitance, typically of order 30 pF/ft. However, at high frequencies, or, equivalently, when dealing with signals with fast rise times, the behavior is different. In order to understand the curious behavior illustrated in Figure H.1, it's helpful to look first at what happens when a simple *pulse* is applied to a length of transmission line. Suppose we connect a fast pulse generator with 50 Ω output impedance (the standard output impedance of signal generators, function generators, and pulse generators) to a length of 50 Ω coax, shorted at the far end. The pulse at first disappears into an impedance Z_0 (thus the signal amplitude is half that of the unloaded generator), but after a round-trip travel time a reflected pulse of opposite polarity returns (Figure H.2). If the signal applied is instead a fast step, the effect of the reflection is to convert the step into a pulse (Figure H.3). An open-ended line produces a reflection of the *same* polarity, with the effects shown in Figure H.4. For arbitrary load resistance R the ratio of reflected to incident wave amplitude (the reflection coefficient) is given by

$$\rho \equiv A_{\text{r}}/A_{\text{i}} = (R-Z_0)/(R+Z_0).$$

Note that a termination resistance of $R{=}Z_0$ produces no reflection. A signal applied to such a terminated line is absorbed by the terminating resistor (as heat) and disappears forever. The signal source sees a loading resistance equal to Z_0. (It is for this reason that we did not have to worry earlier about the reflected pulses reflecting again from the pulse

[1] Higher performance standards include Category 5e ("e" for *enhanced*), and Category 6, propelled by the development of Gigabit Ethernet – literally 1 gigabit/sec over unshielded twisted pair – also known as 1000baseT. To achieve this data rate, all four pairs are used simultaneously with 5-level amplitude encoding.

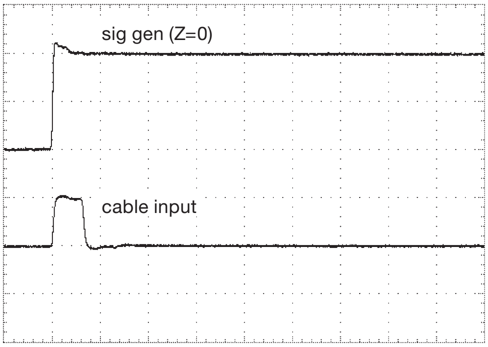

Figure H.3. ’Scope trace of a step waveform applied to an 8-foot length of RG-58A/U (solid polyethylene dielectric, velocity factor of 66%), shorted at the end. Vertical; 1 V/div; horizontal; 40 ns/div.

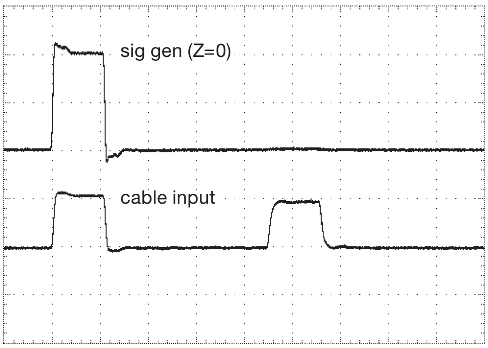

Figure H.4. ’Scope trace showing reflection from open-ended coax line. Same parameters as those in Figure H.2B.

generator – its 50 Ω source impedance swallows any signals returning from an improperly terminated cable, which is the reason most signal sources are standardized at 50 Ω impedance.)

A. Series termination

This last point – that returning (reflected) signals are completely absorbed if the signal source’s impedance matches the line – leads to a nice technique called *series termination* (or *back termination*), frequently used for high-speed logic signals (and in other situations where the load has a high input impedance). Look at Figure H.5: a signal source in series with a resistor (equal to the line’s impedance) drives a transmission line whose far end is unterminated (i.e., open). Now imagine a step input of amplitude V_{sig} at the signal source; it propagates down the line at half-amplitude, then reflects back from the far end at the full V_{sig} amplitude. Although any point along the line sees a two-step waveform, the surprising fact is that the waveform seen at the far end makes a single step from zero to the full

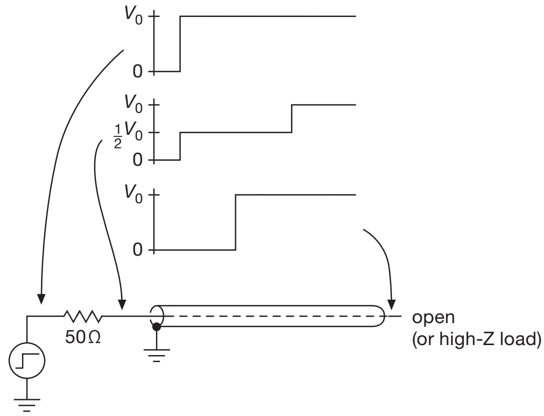

Figure H.5. Series termination of an open-ended line: the line is fed from a signal source of matched impedance; the half-sized step at the input propagates to the far end, from which it reflects in-phase to produce a returning step equal in amplitude to the *unloaded* amplitude of the generator. A high-impedance load at the far end sees only a single full-sized step.

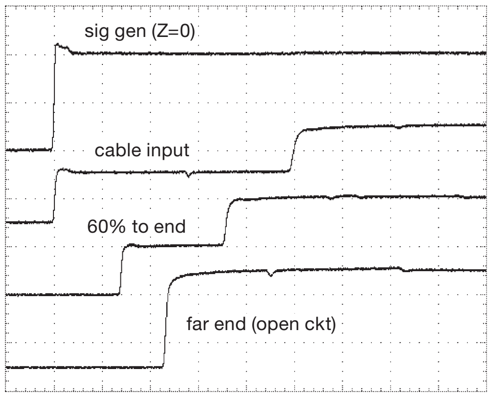

Figure H.6. ’Scope trace showing waveforms at cable input, midpoint, and far end, for series-terminated step input. The generator’s zero-impedance signal is also shown; it was set to produce a 2 V step into an open circuit. The cable is 60 ft of RG-58/U (velocity factor of 66%), tapped at 36 ft with a high-impedance voltage probe. Vertical; 1 V/div; horizontal; 40 ns/div.

V_{sig}; at that place the half-sized incident wave arrives at the same time that the half-sized reflected wave departs. This interesting behavior is demonstrated in the ’scope traces in Figure H.6.

You can use this technique for sending CMOS logic signals through a length of coax: three paralleled 74HC buffers (for low source impedance, roughly 15 Ω) in series with a 33 Ω resistor nicely drives lengths of RG-174 or RG-316 (thin 50 Ω coax line), connected to the receiving gate at the far end *without termination*. The receiving gate sees

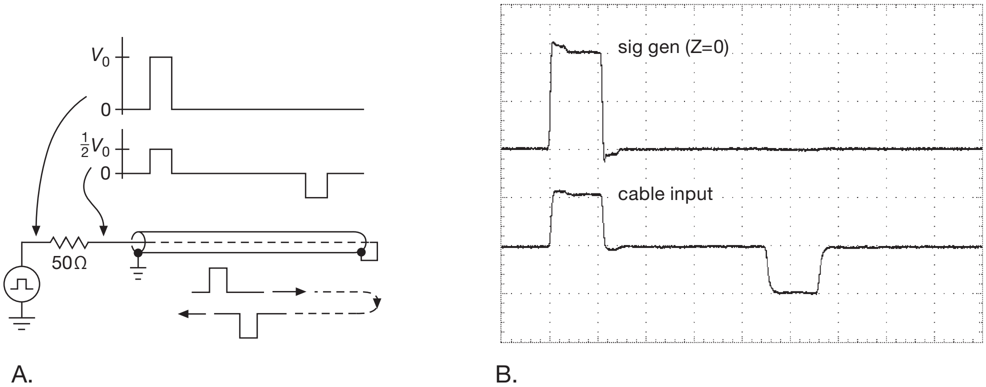

Figure H.2. A. A pulse driving a length of shorted transmission line reflects off the short and returns as a pulse of opposite polarity. B. Scope trace taken with 70 ft of RG-8/U with foam dielectric (velocity factor of 78%), shorted at end. Vertical; 1 V/div; horizontal; 40 ns/div. For this and following figures we used a high-impedance 'scope probe to avoid introducing additional transmission-line effects.

full-swing logic signals. This technique is often preferable to the matched-load alternative – directly driving a line that is terminated in 50 Ω – because with series termination the driver sees a load resistance twice as high (100 Ω in this case), and that only for the round-trip duration of the signal (after which the load becomes an open-circuit).[2] For very fast logic signals (e.g., ECL 100K, or contemporary high-speed CMOS processors, memory, and peripherals), it may be necessary to treat a circuit trace of just a few inches as a transmission line. Typically printed circuit board (PCB) trace impedances are in the range of 50 Ω to 100 Ω, but can be tailored to a specific impedance by proper choice of trace width and spacing above the ground plane; this specialty art is known as *microstrip* technique,[3] useful both for fast digital signals and for signals at frequencies above about 100 MHz (UHF and microwave).

This is all very nice in theory – but in practice you have to contend with sources of fast digital signals that are not matched to the line impedance. This happens often on digital PCBs where the digital output ports of speedy microprocessors and FPGAs are poorly matched to the PCB trace impedances. For example, a line driven by a signal of source impedance $Z_0/2$ produces lots of ringing at both ends of an unterminated line, including 33% overshoot at the far end; such ringing can produce false clocking.

To illustrate what this looks like, we drove a 2.4 m length of unterminated RG-58/U (Z_0=50 Ω) coax with a step input, probing both the input and output signals.

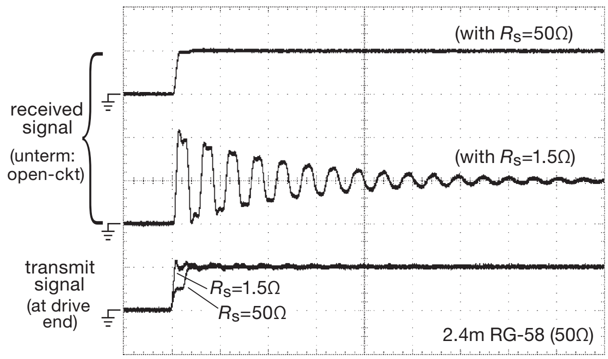

Figure H.7. Signals seen at the far end of an unterminated 8ft length of 50 Ω line when driven with a unit step from a 50 Ω source (top trace) and from a low-impedance source (middle trace). The corresponding signals at the driving end are shown at bottom. Horizontal; 100 ns/div.

We did this under two conditions: (a) when driven with a 50 Ω series termination at the input (i.e., a "back-terminated" line); and (b) when driven with a low-impedance (R_s=1.5 Ω) voltage step.

Figure H.7 shows the results, where the drive signal's open-circuit amplitude (call it V_{OC}) is displayed as one vertical division. The matched series termination generates a clean received step to V_{OC} (and an input waveform with initial step first to $V_{OC}/2$, then to V_{OC} after a round-trip delay, just as seen in expanded scale in Figure H.6). By contrast, the low-Z drive imposes a full V_{OC} step at the input, which is first seen at the far end at $2V_{OC}$ (because the non-inverted reflected wave doubles the amplitude of the arriving wave), subsequently brought nearly back to zero (one round-trip time later) by the inverted wave reflected from

[2] See §12.10, where several methods of driving cables with logic levels are described and illustrated.

[3] If you sandwich the signal-carrying conductors between a pair of ground planes, you've got a *stripline*; see §*1x.1*.

the low-impedance driver. This alternating pattern continues, damped both by cable loss and resistive loss in the $1.5\,\Omega$ driver. This signal is a mess! That's why it's worth some effort to match driver impedances to the characteristic impedance of lines whose length (given the signal rise times) qualifies them as transmission lines.[4]

B. A robust logic link

The sort of overshoot behavior seen in Figure H.7 can destroy logic circuits at the receiving end. You see this vulnerability particularly in the laboratory, where a low-impedance logic signal (or pulse-generator output) travels through a length of coax to a logic input on some instrument. The latter is often unterminated, to keep the input impedance relatively high (to prevent heavy loading and attenuation for a signal source unable to drive $50\,\Omega$).

If you want to make your own designs bulletproof against this hazard, you can add a few components as shown in Figure H.8. At the receiving end, series resistor R_2 limits the current, safely clamped by the logic gate's internal protection diodes; the "speed-up capacitor" C_1 prevents loss of speed (1 kΩ into a typical input and wiring capacitance of 10 pF is an *RC* time constant of 10 ns, a near-eternity in the frenetic world of digital logic). It's always a good idea to include an input pull-down resistor (R_1) to ensure a defined logic level when the input is disconnected. In an abundance of caution we've added resistor R_3, whose job is to prevent a large positive overdrive from forcing the V_+ rail to a positive voltage that can damage other ICs; this could normally be omitted, but it would be a good idea in an instrument with an inviting BNC connector on the front and that contains only a few ICs powered from a small regulator (like a 78L05 – see §9.3.2, Figure 9.6, and Table 9.1) whose dc output can be easily overdriven.

At the driver end the parallel connection of several logic gates generates a drive impedance down in the neighborhood of 5–10 Ω, which the added series resistor R_o brings up to a driving resistance close to the cable's $50\,\Omega$ characteristic impedance. That's what you want, of course, and that alone is enough to give you peace of mind as a respectable series-terminated driver. But it never hurts to protect the receiver end, as just discussed: you never know when someone will drive it with a pulse generator, inadvertantly set to deliver *negative* pulses, or 20 V positive pulses (as happened in our laboratory recently).

[4] An analogous effect in long-distance power transmission lines is known as the Ferranti effect; it is said that overvoltages caused by the Ferranti effect, if not properly compensated, can cause damage to power-line switch gear.

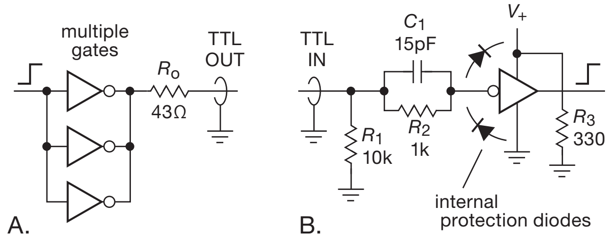

Figure H.8. A. Simple logic-level driver for series-terminated cable; B. Logic-level receiver protected against overdrive. Connect these together to form a complete signal path.

There's further discussion of cable driving and logic interfacing in Chapter 12, beginning at §12.10.

H.1.3 Termination: sinusoidal signals

We have been talking about signals propagating along transmission lines, for clarity using the particular case of pulses or voltage steps. Of course, a *sinewave* applied to a length of cable also produces reflections, unless of course the cable is properly terminated. The effect is to alter the input current, for an applied input voltage, in a way that depends on the (mismatched) load impedance Z_L, and also on the ratio of the signal's wavelength in the cable λ to the physical length of the cable l. The final effect is to produce an input impedance (complex in general) given by

$$Z_{in} = Z_0 \frac{Z_L + jZ_0 \tan(2\pi l/\lambda)}{Z_0 + jZ_L \tan(2\pi l/\lambda)}.$$

From this one can see that:

(a) a matched termination ($Z_L = Z_0 =$ [usually] $50\,\Omega$) results in an input impedance equal to the characteristic impedance of the cable, independent of length or frequency;

(b) a quarter-wave line inverts the load impedance, i.e., $Z_{in} = Z_0^2/Z_L$;

(c) a half-wave line preserves the load impedance, i.e., $Z_{in} = Z_L$;

(d) a short length of open-circuited line $l \ll \lambda$ looks capacitive, viz., $Z_{in} \approx -j/\omega C'$, where C' (the effective capacitance) is the constant l/cZ_0;

(e) a short length of short-circuited line $l \ll \lambda$ looks inductive, viz., $Z_{in} \approx j\omega L'$, where L' (the effective inductance) is the constant $Z_0 l/c$.

The impedance-changing properties of transmission lines can be used to match impedances, though any such scheme will be frequency dependent; when you hear words like "stubs," you're dealing with transmission-line impedance matching. Virtuosos in this area make heavy use of network analyzers, and they will try to dazzle you with their handsome “Smith Charts” (which are well beyond the humble scope of this book[5]).

When you have sinusoidal signals – with reflections – on a transmission line, you generate *standing waves*. That is, you can picture the net result of waves propagating in both directions (at the same frequency) as the sum of a nonpropagating (hence “standing”) wave and some additional propagating wave. For example, an open-ended line produces a reflected wave of full amplitude; the result is a pure standing wave of the same frequency and twice the amplitude, with a maximum amplitude of oscillation at the open end (and repeating every half-wavelength), and complete nulls (“nodes” – places with no voltage) midway between. For a shorted-end line a similar thing happens, but the reflected wave is of opposite amplitude, producing a null at the far end (and repeating every half-wavelength), with maxima in between. (You get the same pattern if you tie a length of clothesline to a fence, then wiggle the end up and down at the right rate.) With a smaller termination mismatch you don’t get complete cancellation anywhere.

Standing waves are not necessarily bad (though they are never[6] *good*!). But they do increase both the peak voltages and the resistive losses (see next section), relative to a matched line, for the same power transmitted. They are ordinarily seen as the symptom of a mismatched line. So in communications systems people try to minimize the *standing wave ratio* (abbreviated SWR, or sometimes VSWR – for *voltage* standing wave ratio – pronounced “VIZ-wahr”), which is defined as the ratio of the maximum amplitude to minimum amplitude:

$$\mathrm{VSWR} = \frac{V_{\max}}{V_{\min}} = \frac{A_\mathrm{f}+A_\mathrm{r}}{A_\mathrm{f}-A_\mathrm{r}},$$

where V is the ac (signal) voltage, measured at points along the line; and A_f and A_r are the amplitudes of the forward and reflected waves, respectively. Voltage measurements along a cable with no standing waves will give a constant amplitude (hence VSWR=1.0).

The VSWR is a real number, between 1.0 (perfect match, no reflected wave) and ∞ (“perfect mismatch,” reflected wave amplitude equal to incident wave amplitude). In terms of the reflection coefficient ρ, the VSWR is just

$$\mathrm{VSWR} = \frac{1+|\rho|}{1-|\rho|}.$$

For a purely resistive mismatch we can use our earlier expression for ρ to find that

$$\mathrm{VSWR} = \begin{cases} R/Z_0 & \text{if } R > Z_0 \\ Z_0/R & \text{if } Z_0 > R. \end{cases}$$

From the VSWR you can find the magnitude (but not the phase) of the reflection coefficient:

$$|\rho| = \frac{\mathrm{VSWR}-1}{\mathrm{VSWR}+1}.$$

The VSWR is measured with a directional power meter. Knowing the values of forward and reflected power, you know from the defining equation above that

$$\mathrm{VSWR} = \frac{1+\sqrt{P_\mathrm{r}/P_\mathrm{f}}}{1-\sqrt{P_\mathrm{r}/P_\mathrm{f}}}.$$

H.1.4 Loss in transmission lines

In the real world of non-ideal transmission lines, things aren’t quite as nice as we’ve led you to believe. Real transmission lines are *lossy*, meaning that signals are attenuated as they travel down the line; they are also slightly *dispersive*, meaning that signals of different frequency travel with slightly different speeds. The loss is *frequency dependent*: its value (often specified as attenuation in “dB per 100 ft”) increases proportional to $\sqrt{f}$; i.e., a quadrupling of frequency doubles the loss (in dB) of a given length of line. This happens because the loss is dominated by the *skin effect*: when alternating current flows through a conductor, the current is not uniform throughout the bulk – it is confined to an outer layer (called the *skin depth*) of thickness $\delta{=}(\pi\sigma f)^{1/2}$, where σ is the conductivity and f is the frequency.[7] Because the skin depth decreases inversely with the square root of frequency, you have to quadruple the frequency to double the resistance (halve the skin depth),

[5] Smith charts are treated in the excellent reference *Fields and Waves in Communication Electronics* by Ramo, Whinnery, and Van Duzer (Wiley, 1994), as well as in the insightful and refreshing *Radio-Frequency Electronics* by Hagen (Cambridge University Press, 2009).

[6] Well, *hardly ever!* For operation over a narrow frequency range you sometimes exploit the impedance-changing properties of mismatched lines, which necessarily have standing waves. Examples are (a) the use of open or shorted lengths of line as high-Q capacitors or inductors, (b) a shorted half-wave or an open quarter-wave line used as an RF bypass capacitor, (c) an open half-wave or a shorted quarter-wave line used as an RF choke, (d) matching two different impedances (cables, signal sources, or loads) by interposing a quarter-wave section of transmission line whose characteristic impedance is the geometric mean of the two impedances being matched (this is analogous to a quarter-wave anti-reflection coating in optics), (e) the use of a slotted line and high-impedance probe for a direct measurement of wavelength, and (f) the use of transmission lines to make “ring hybrids” and “rat-race hybrids.” We thank Jon Hagen and Darren Leigh for these applications of “good standing waves.”

[7] To be precise, the current density decreases exponentially, falling to $1/e$ (37%) of its surface value at a depth equal to δ.

which is equivalent (in terms of loss) to doubling the length of the line. This explains the approximate slope of the attenuation curves for transmission lines (Figure H.9), where lines of larger diameter have lower losses. Dielectric losses contribute additional attenuation at the highest frequencies. The curves shown are for a matched line; if there are reflections (i.e., if the VSWR is greater than 1.0) then the loss, *for a given net power transmitted down the line*, will be greater.

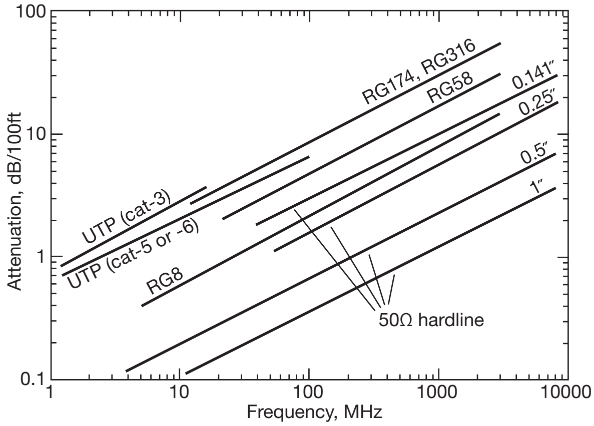

Figure H.9. Attenuation (dB/100 ft) as a function of frequency for several representative cable types.

It's useful to realize that skin-depth effects are significant even at low frequencies – current at the powerline frequency of 60 Hz is confined to a surface layer of roughly 1 cm in copper, for which the skin depth (in centimeters) at room temperature is given by $\delta(\text{Cu})=6.6/\sqrt{f}$; you don't reduce power transmission losses much by using wire thicker than that. At radio frequencies the skin depth is so shallow (e.g., $\delta=10\,\mu\text{m}$ at 40 MHz) that you can make low-loss connections, inductors, and so on, by silver plating a poor conductor. A common technique for shielding lightweight instruments and computers is to apply a thin metallic plating to a plastic enclosure. Figure H.10 plots the skin depth in copper conductors from 10 Hz to 10 GHz.

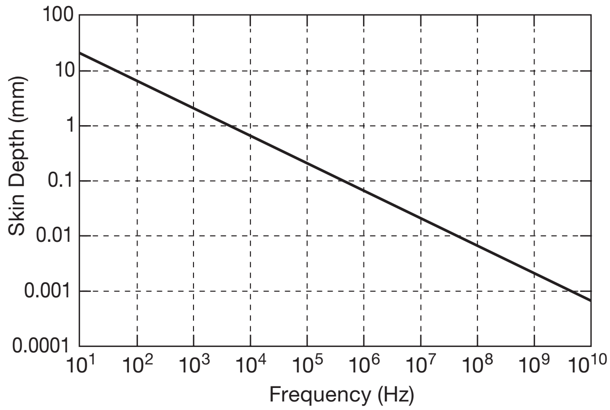

Figure H.10. Skin depth in copper as a function of frequency.

H.2 Impedance matching

Because you get reflections from unterminated (or incorrectly terminated) transmission lines, it's obviously a good idea to make sure you match impedances when you use coax lines whose electrical length is a significant fraction (at least 1/20, say) of the wavelength of the highest frequencies you're using. Stated in terms of time rather than frequency, you have to start worrying about termination when the round-trip propagation time is about 20% of the signal rise time.

We've seen already that a simple way to do this is to terminate the line in its characteristic impedance (resistance), for example 50 Ω for most coax lines. Termination is not required at *both* ends, because a terminated end swallows any incident signals. Thus you can use a mismatched signal-source impedance to drive a line terminated at the far end; or, as we saw above, you can "series terminate" the driven end of a line whose far end is unterminated (i.e., fed to a load of much higher impedance than the line). The usual practice, however, is to terminate *both* ends in the line's characteristic impedance (a conservative instinct that ensures a minimum of reflections). For example, you usually use a 50 Ω cable to pipe the signal from a synthesizer or signal generator to a 50 Ω matched load at the far end; if your load is high impedance, you place a 50 Ω resistor across it (or use a 50 Ω coax feedthrough termination).

Sometimes you need to match a line to a load (or source) of a different impedance; for example, you might want to measure the properties of some 75 Ω video cable (that's the impedance that the video industry has chosen, much to the chagrin of the rest of the electronics community), and all you've got are 50 Ω test instruments. Or, you might want to match the output impedance of a high-frequency amplifier to a length of cable that goes to an antenna.

This brings us to the subject of matching networks. In the following subsections we treat (a) resistive (lossy) networks for broadband impedance matching and attenuation, (b) transformer (lossless) broadband matching, and (c) reactive (lossless) narrowband matching.

H.2.1 Resistive (lossy) broadband matching network

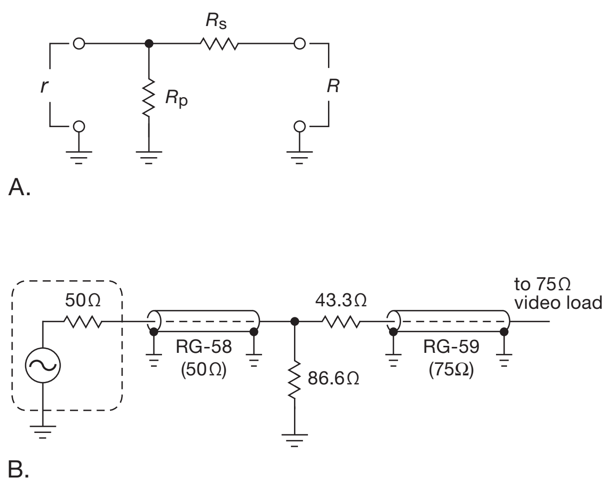

Figure H.11. A. A resistive L-network can match any pair of real (resistive) impedances; the parallel resistor R_p goes across the port of lower impedance r. B. Example of matching a 50 Ω signal source and cable to a 75 Ω video cable and load (with a loss of 5.72 dB).

You can easily figure out that two resistors (in the form of an "L-network," Figure H.11A) is all it takes to match a pair of impedances r and R (assumed resistive, as all cables are); both sides are happy – each sees a matched load. The values of the matching resistors are

$$R_p = r\sqrt{\frac{X}{X-1}},$$
$$R_s = r\sqrt{X(X-1)},$$

where r is the smaller impedance and X is the ratio of impedances: $X=R/r$. Taking the earlier example, you can match a 50 Ω test instrument to a 75 Ω coax line (the common variety is called RG-59) by putting 86.6 Ω across the 50 Ω port and connecting it to the cable through a 43.3 Ω series resistor (Figure H.11B).

The good news is that such a resistive L-network is frequency independent; the bad news is that it's lossy. It's easy to show that the loss is

$$\text{loss} = 20\log_{10}\left(\frac{\sqrt{X}}{X+\sqrt{X(X-1)}}\right)\ \text{dB}.$$

For example, the 50Ω-to-75Ω L-network above has a transmission loss of 5.72 dB for signals going in either direction. With a resistive match you have to accept this attenuation (this is sometimes called a *minimum loss pad*). We'll see below how to make lossless matching networks with transformers or with networks of L's and C's ("reactive matching networks").

As you might imagine, you can do even *worse*, in terms of loss, with a network containing more resistors! In particular, you can add another resistor, making either a "T" network or a "Pi" network, that matches two resistive impedances to each other, with loss greater than the minimum loss we found above. Although this isn't something you ordinarily want to do, there is a variation on that theme that is often useful; namely, a resistive attenuation network between a pair of already-matched impedances.

H.2.2 Resistive attenuator

In radiofrequency circuits you sometimes need to attenuate a signal level – for example, to avoid overdriving a stage of gain. In other situations you need to use a resistive attenuator to provide some isolation between an impedance-sensitive component like an amplifier, mixer, or cable, say, and a component that is not impedance matched; an example of the latter is a filter, which typically is impedance matched in its passband, but reflective (a severe mismatch) in its stopband. Some amplifiers will oscillate if their output directly drives a sharp filter.

The solution to these problems is a resistive impedance-matched attenuator. The two topologies are T and Pi, named for their appearance on a schematic diagram (Figure H.12). It is not difficult to derive the resistor values:

$$R_p = \frac{1+x}{1-x}R_0,$$
$$R_s = \frac{1-x^2}{2x}R_0, \qquad (\text{Pi}-\text{network}),$$

and

$$R_s = \frac{1-x}{1+x}R_0,$$
$$R_p = \frac{2x}{1-x^2}R_0, \qquad (\text{T}-\text{network}),$$

where x is given by the attenuation: $\text{atten(dB)}=-20\log_{10}x$ (or, equivalently, $x=10^{-\text{atten(dB)}/20}$), and Z_0 (assumed resistive) is the impedance at both input and output. Tabulated values for 50 Ω source and load impedances are given in Table H.1.

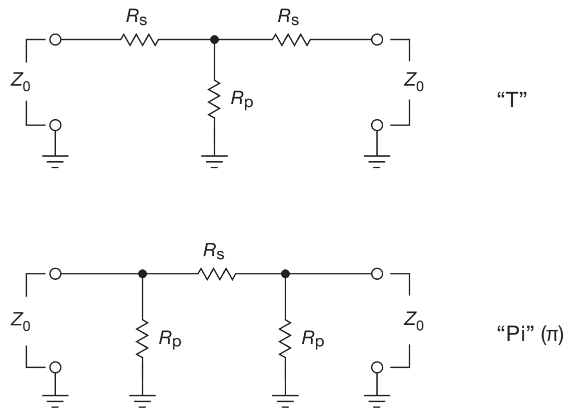

Figure H.12. Resistive T and Pi attenuators for equal input and output impedances.

Table H.1 50Ω T and Pi Attenuators

Attenuation (dB)	Pi R_P (Ω)	Pi R_S (Ω)	T R_P (Ω)	T R_S (Ω)
0	∞	0	∞	0
0.25	3.47k	1.44	1.74k	0.72
0.50	1.74k	2.88	868	1.44
0.75	1.16k	4.32	578	2.16
1.00	870	5.77	433	2.88
1.25	696	7.22	346	3.59
1.50	581	8.68	288	4.31
1.75	498	10.1	247	5.02
2.0	436	11.6	215	5.73
2.5	350	14.6	171	7.15
3	292	17.6	142	8.55
4	221	23.9	105	11.3
5	178	30.4	82.2	14.0
6	150	37.4	66.9	16.6
7	131	44.8	55.8	19.1
8	116	52.8	47.3	21.5
9	105	61.6	40.6	23.8
10	96.3	71.1	35.1	26.0
15	71.6	136	18.4	34.9
20	61.1	248	10.1	40.9
25	56.0	443	5.64	44.7
30	53.3	790	3.17	46.9
35	51.8	1.41k	1.78	48.3
40	51.0	2.50k	1.00	49.0
45	50.6	4.45k	0.56	49.4
50	50.3	7.91k	0.32	49.7
55	50.2	14.1k	0.18	49.8
60	50.1	25.0k	0.10	49.9

Resistor values for T and Pi attenuators for use with 50 Ω source and load. The values shown can be scaled for use at some other impedance, assuming equal input and output impedances.

H.2.3 Transformer (lossless) broadband matching network

If the unavoidable loss of a resistive matching network is unimportant in your application, that certainly is the simplest method. However, in many applications it is essential to minimize loss – for example, in a communications transmitter or in low-level circuits whose performance is limited by amplifier or thermal noise.

In that case you can use either a transformer or a reactive matching network; neither method provides coupling all the way down to dc, however. Transformer coupling is relatively broadband, but it is limited in impedance ratios; reactive matching, by contrast, flexibly matches impedances (including reactive impedances), but only around some chosen center frequency. We treat reactive matching in the next subsection.

Transformers for use at signal frequencies are similar in principle to ordinary ac power transformers, that is, they use a pair of windings that are coupled magnetically, and whose turns ratio is the desired voltage ratio. The impedance ratio is then the square of the turns ratio. (For example, a transformer with a 1:4 primary:secondary turns ratio, driven with a 50 Ω signal source, would present an output impedance of 800 Ω, and should be loaded with that resistance.) However, because of the higher signal frequencies, it is necessary to use magnetic cores that do not have significant conductive paths for eddy currents. At audio frequencies the solution is to use the same sort of laminated metal stacks used in power transformers, but with much thinner laminations. At still higher frequencies, laminated cores are replaced either by powdered iron cores or by completely nonconducting magnetic "ferrite" materials. Because of the devastating effects of parasitic capacitance and inductance, transformers for use at high radio frequencies (say above 10 MHz) generally are constructed from transmission lines (coaxial or parallel) wound around a magnetic core.

Impedance-matching transformers are widely available commercially, though for special applications you may need to design and wind your own. At audio frequencies many manufacturers offer miniature impedance matching transformers with "telephone" bandwidths (200 Hz–4 kHz), or full audio bandwidth (20 Hz–20 kHz); impedances go from loudspeaker and microphone impedances (8–600 Ω) up to "hi-*Z*" values of 10k–50kΩ. There's further discussion in §8.10.

A nice series of radiofrequency transformers is made by North Hills, including models that transform 50 Ω or 75 Ω to impedances up to 1200 Ω; these cover the frequency range between 20 Hz and 100 MHz, with a typical frequency range of 1000:1 or more for a given transformer. For higher frequencies you can get broadband matching transformers from Mini-Circuits, covering the range of 4 kHz to 2 GHz with impedance ratios from 1:1 to 16:1, and with frequency ranges of 1000:1 or more for a given transformer. These are constructed with transmission-line techniques.

It is worth noting that transformer coupling provides *galvanic isolation*: input and output circuits need not share the same ground. This is particularly useful when you need to send a signal (or distribute a "house clock") between instruments whose individual cases are each grounded through their power cords or enclosure rails. We have seen several instances in which instrument "ground," in the same laboratory, differed by as much as *several volts* of 60 Hz ac. Here an isolated 50 Ω:50 Ω broadband transformer is ideal, for example the Mini-Circuits FTB1-6 (10 kHz–125 MHz) or the North Hills 0016PA (20 Hz–20 MHz).

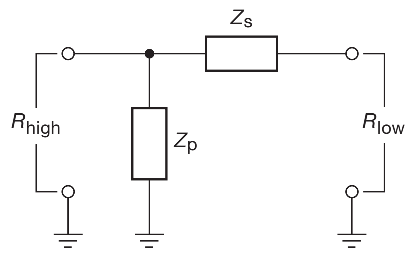

Figure H.13. Lossless reactive impedance-matching L-network.

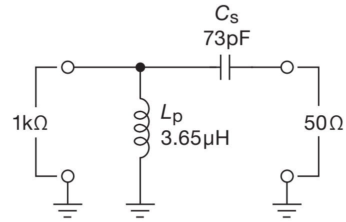

Figure H.14. Example of lossless network to match a 1 kΩ source impedance to a 50 Ω load, at a center frequency of 10 MHz.

H.2.4 Reactive (lossless) narrowband matching networks

You can match *any* pair of impedances, real or complex, with just two reactive components. The resulting match is perfect only at a single frequency, but "good enough" over some modest band of frequencies. This can be considered an alternative to (broadband) transformer matching, with considerably greater flexibility in target impedances. It is worth noting that a lossless match between impedances that are not purely real (i.e., that have a reactive component) will always be narrowband.

The simplest reactive matching network is an L-network with one inductor and one capacitor (Figure H.13). You can choose either the inductor or the capacitor as the parallel element, but the network must have the parallel reactance located across the port with the larger impedance. The design procedure is straightforward, and is nicely motivated and explained in Hagen (see Appendix N). It amounts to choosing the parallel reactance to produce (in combination with the higher port's impedance R_{high}) the correct lower resistance R_{low} as its real part, then using the series reactance to cancel the resulting reactance.

The procedure goes like this.

1. Calculate the quantity

$$Q_{EL} = \sqrt{\frac{R_{high}}{R_{low}} - 1}$$

(this will be twice the actual Q value, or frequency selectivity, of the matching network).

2. Now select the form of parallel reactance (i.e., inductor or capacitor), and set the magnitude of its reactance equal to R_{high}/Q_{EL} at the center frequency. In other words, $L_{parallel}=R_{high}/2\pi f Q_{EL}$ or $C_{parallel}=Q_{EL}/2\pi f R_{high}$, respectively.
3. Finally, set the magnitude of the series reactance (i.e., capacitor or inductor, respectively) equal to $Q_{EL}R_{low}$ at the center frequency. In other words, $C_{series}=1/2\pi f Q_{EL} R_{low}$ or $L_{series}=Q_{EL}R_{low}/2\pi f$, respectively.

As an example, let us match a 1000 Ω source (an amplifier output) to a 50 Ω load (an antenna) at a frequency of 10 MHz. We find Q_{EL}=4.36, and, choosing a parallel inductor at the input, $L_{parallel}$=3.65 µH and C_{series}=73 pF (Figure H.14). The Q of the resultant coupled network is equal to $Q_{EL}/2$, roughly $Q \approx 2$; its bandwidth is thus about 50% between half-power points, though the match is perfect only at the center frequency. Note that the Q rises with increasing impedance ratio, and that you have no control over it. The network becomes a narrow bandpass filter for very large impedance-transformation ratios.

If you want a higher Q, you can get it by adding another reactive component, to form a Pi (or T) network. You can think of this as a pair of L-networks, going to an intermediate impedance that is much lower (or higher) than either port impedance. Each L-section then has an impedance ratio greater than the final transformation, hence the higher Q value. You might think narrow bandwidth is bad, but in fact it is often desirable in communications circuits where you want to suppress out-of-band signal energy.

Alternatively, you can get lower Q than the simple L-network gives you by cascading a pair of Ls – a "double L." Here the impedance transformation is taken in two half-steps – each section's ratio is smaller than the final ratio, hence a lower Q.

H.3 Lumped-element delay lines and pulse-forming networks

The continuous transmission line with inductance and capacitance per unit length of L/l and C/l, respectively, can be approximated by an array of N discrete series inductors L and shunt capacitors C (Figure H.15). It is easy to show that the resultant circuit approximates a transmission line whose propagation time per element is $\tau_i = \sqrt{LC}$, and whose impedance is $Z_0 = \sqrt{L/C}$; the total propagation time is $t_p = N\tau_i = N\sqrt{LC}$.

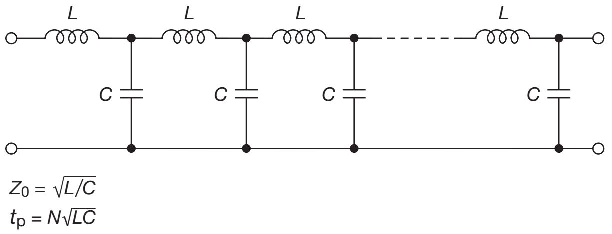

Figure H.15. Lumped-element delay line formed with an array of equal-value capacitors and inductors.

You can make a delay line this way, approximating a long transmission line. Roughly speaking, such a discrete approximation to a continuous transmission line will preserve details of the waveform only down to time scales of τ_i, or $1/N$th of the total propagation time. For example, a $1\,\mu$s lumped delay line with 20 LC sections will swallow details shorter than about 50 ns. It is a lowpass filter that attenuates frequencies above $f = 1/2\pi\sqrt{LC}$.

Lumped-element delay lines were used in early analog oscilloscopes to allow time for the sweep to begin before the (delayed) signal reached the deflection circuitry; this let you see the triggering event (and a bit before). Later analog scopes used lengths of a helical-conductor coaxial line for the same purpose. Take a look at §H.4.3 for a bit more on this fascinating application.

Lumped delay lines are useful as *pulse-forming networks*, as shown in Figure H.16. Here the parallel capacitance of a set of LC delay sections is charged to a high positive voltage; then the "center conductor" of the coax equivalent is switched to ground with a high-voltage switching element such as a thyratron. The common terminal (analogous to the coax "shield") then produces a negative voltage pulse, of duration equal to *twice* the delay-line propagation time; its source impedance is just that of the delay line. This can drive a load directly; often it is converted to a different amplitude (and perhaps opposite polarity) with a pulse transformer, as shown. Pulse-forming networks find use in radar and other applications in which the pulse voltage and/or duration are inconvenient to produce with the analogous transmission-line circuit.

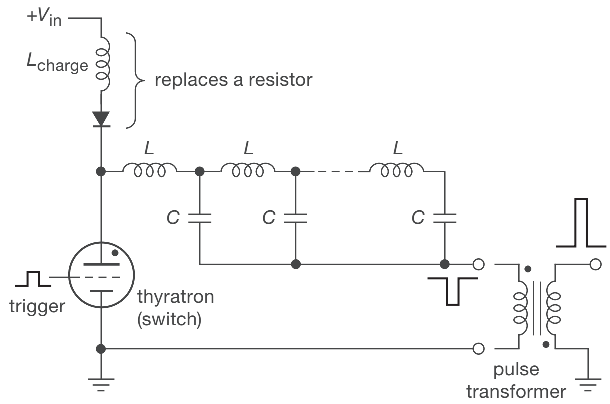

Figure H.16. Pulse-forming network for producing a high-voltage pulse of high energy. The thyratron is a special type of vacuum tube, containing a small amount of hydrogen or other gas, designed for switching really high voltages and high currents (10s of kV, 1000s of amperes: 10s of megawatts!). The inductor–diode fragment shown is a way to implement efficient "resonant charging" of a capacitor from a fixed dc voltage.

Shown also in this diagram is the method of "resonant charging," in which an inductor L_{charge} plus diode replaces the conventional charging resistor for the purpose of charging a capacitor C_{total} (the N capacitors in parallel). This has several benefits: (a) the charging wastes no energy, whereas resistive charging wastes exactly 50%; (b) it is complete after a time equal to half the period of the resonant circuit formed by L_{charge} and C_{total}; and (c) it charges the capacitor(s) to twice the supply voltage. Resonant charging is a clever technique, and it is used also in switching converters and power supplies, flashlamp circuits, and elsewhere.

H.4 Epilogue: ladder derivation of characteristic impedance

In a formal course on "waves" you are usually subjected to an analysis that uses Maxwell's equations to derive the relationship between the traveling $\vec{E}$ field and $\vec{B}$ field, from which follows the relation between voltage and current, and hence impedance. As a bonus you get the capacitance and inductance per unit length, and the speed of propagation.

But there's a nice "circuit" way to convince yourself that a properly terminated transmission line looks like a pure resistance (equal in value to its "characteristic impedance," for example 50 Ω), namely to model it as a discrete *LC* ladder (Figure H.17) consisting of little increments of length δx, each having a series inductance $\mathscr{L}\,\delta x$ and shunt capacitance $\mathscr{C}\,\delta x$, where $\mathscr{L}$ and $\mathscr{C}$ are the inductance and capacitance per unit length of the coaxial line. There are $l/\delta x$ of these in the whole length l of the line.

H.4.1 First method: terminated line

We start[8] by writing down the impedance looking into the last section of a terminated line ($\mathbf{Z}_1$ in Figure H.17), which will give us a condition on the ratio $\mathscr{L}/\mathscr{C}$ in order for $\mathbf{Z}_1$ to equal, approximately, R_0. Then we'll see that the impedance $\mathbf{Z}_{\text{in}}$ looking into the whole ladder converges exactly to R_0 as we convert the discrete ladder approximation to a continuous transmission line by taking the limit as δx goes to zero.

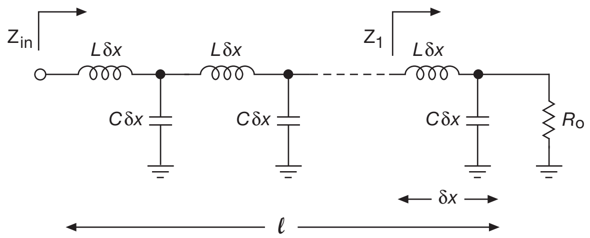

Figure H.17. *LC* ladder model of a transmission line of length l. We're ultimately interested in the limit $\delta x \to 0$, where the number of sections $N=l/\delta x \to \infty$.

Let's do it. $\mathbf{Z}_1$ is just the impedance of $\mathscr{L}\,\delta x$ in series with the parallel impedance of R_0 and $\mathscr{C}\,\delta x$:

$$\mathbf{Z}_1 = j\omega\mathscr{L}\delta x + \frac{R_0\cdot(-j/\omega\mathscr{C}\,\delta x)}{R_0 - j/\omega\mathscr{C}\,\delta x}$$

$$= j\omega\mathscr{L}\delta x + \frac{R_0}{1+j\omega\mathscr{C}R_0\delta x}$$

$$\approx R_0 + j\omega\delta x(\mathscr{L} - R_0^2\mathscr{C}),$$

where in the last step we've kept only the first term of the binomial expansion, i.e., $1/(1+\varepsilon) \approx 1-\varepsilon$.

The second term vanishes when $\sqrt{\mathscr{L}/\mathscr{C}}=R_0$, which is the formula for the characteristic impedance of a transmission line. But... not so fast – that term would have vanished anyway as $\delta x \to 0$; however, our transmission line would have vanished as well! What we need to do is to include the next-order binomial terms, and see what happens as we let δx go to zero, while holding the total line length l constant; the number of sections then increases, as $N=l/\delta x$.

You can do the math. You'll find that the next two terms add contributions[9] of order δx^2 and δx^3, making $\mathbf{Z}_1$ look like

$$\mathbf{Z}_1 \approx \left\{R_0 + \mathscr{O}(\delta x^2)\right\} + j\left\{\omega\delta x(\mathscr{L} - R_0^2\mathscr{C}) + \mathscr{O}(\delta x^3)\right\}$$

and so, cascading N such sections (where $N=l/\delta x$), with the condition $\sqrt{\mathscr{L}/\mathscr{C}}=R_0$, the higher order terms vanish as $\mathscr{O}(\delta x)$ and $\mathscr{O}(\delta x^2)$ (for the real and imaginary parts, respectively) as $\delta x \to 0$. Thus the input impedance of a transmission line of length l, terminated in its characteristic impedance (resistance) R_0, is pure resistive and equals R_0.

H.4.2 Second method: semi-infinite line

Here's a clever method[10] that doesn't require approximation or worries about convergence. The idea is to look into one end of a lumped-element transmission line that extends to infinity (Figure H.18), noticing that it looks just the same if we step one notch to the right. So, calling the (complex) input impedance $\mathbf{Z}_0$, we have, simply,

$$\mathbf{Z}_0 = j\omega L + \mathbf{Z}_0 \parallel \mathbf{Z}_\mathbf{C} = j\omega L + \frac{\mathbf{Z}_0\cdot(-j/\omega C)}{\mathbf{Z}_0 - j/\omega C}.$$

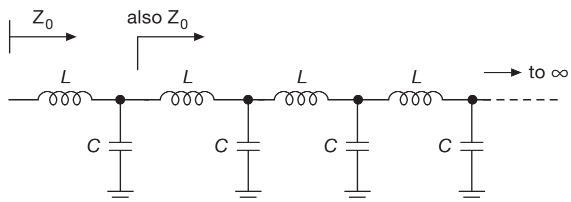

Figure H.18. Semi-infinite *LC* ladder model makes for an easy calculation.

8 This treatment was inspired by Hagen's *Radio Frequency Electronics*; see Appendix N.

9 They are $-R_0^3\omega^2C^2\delta x^2 + jR_0^4\omega^3C^3\delta x^3$, if you want to check your math (or ours!).

10 Suggested to us by Jene Golovchenko.

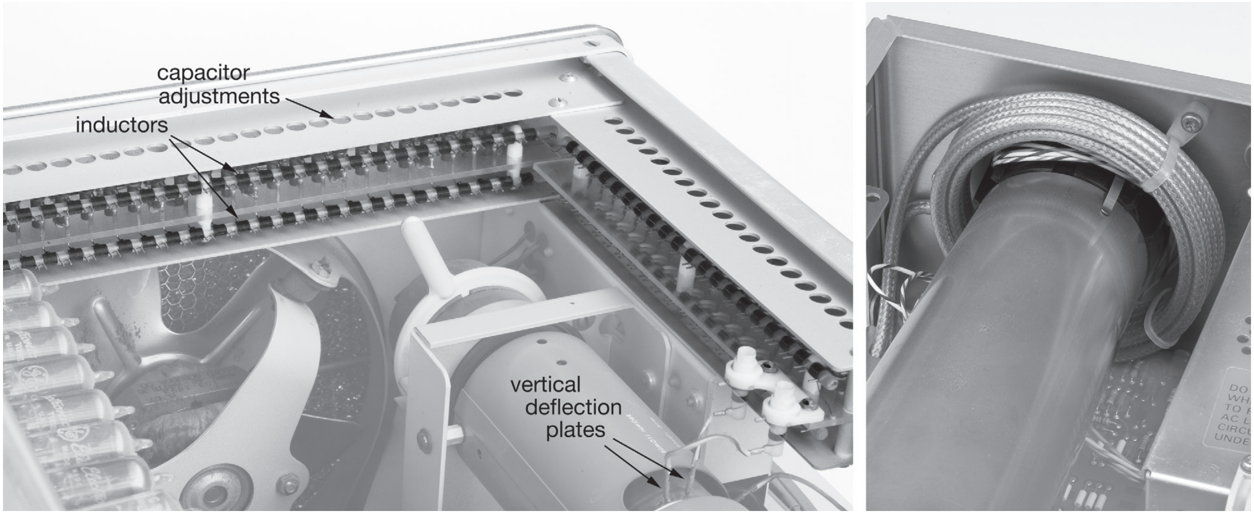

Figure H.19. Analog oscilloscopes used delay lines in the signal path, so you could see the triggering event. The 1959-vintage vacuum-tube Tektronix 545A (left) had a 30 MHz bandwidth, and used a two-channel lumped-element 200 ns delay line consisting of 50 inductor pairs and 50 (adjustable!) trimmer capacitors; (Figure H.20); their 1982 vintage solid-state 2213 'scope (right) had 60 MHz bandwidth, and used a 2.5 m length of spiral-conductor coaxial cable for its 100 ns delay line.

Multiplying through by the denominator of the last term and rearranging terms, we get a quadratic equation for $\mathbf{Z}_0$:

$$\mathbf{Z}_0^2 - j\omega L\mathbf{Z}_0 - L/C = 0,$$

with the solution

$$\mathbf{Z}_0 = j\omega L \pm \frac{\sqrt{4L/C - \omega^2 L^2}}{2}.$$

Now for the *coup de grâce*: we let the individual segments shrink toward zero while keeping the full line length. The individual L's and C's go to zero, but their ratio remains constant. Only the $4L/C$ term survives, giving us the (real) impedance $\mathbf{Z}_0 = \sqrt{L/C}$. No approximations!

H.4.3 Postscript: lumped-element delay lines

The clever designers of analog oscilloscopes, back in the dark ages of electronics, found a way to get the horizontal trace going *before* the triggering event, namely by delaying the displayed signal by ~100 ns, using a delay line.[11] The designers of early vacuum-tube 'scopes (like the legendary Tektronix 545A) used a lumped-element transmission line like the one in Figure H.19 to achieve the delay (200 ns in this case) that otherwise would have required more than 100 feet of cable; presumably they also felt some satisfaction in exploiting the theory they learned in an electronics course they had taken years earlier.[12] Later analog 'scopes replaced the lumped-element delay line with a cleverly engineered coax delay line with a pair of spiral inner conductors; its larger inductance per unit length increased the signal delay,[13] and, happily, increased the characteristic impedance as well. The photograph shows an example, where the cable has been stuffed into an empty space at the rear, conveniently wrapped about the CRT.[14] Figure H.21 reveals the inner secrets of this elegant differential-pair delay line, whose crisscrossing counterwound helical "center conductor" produces a delay of 12 ns/ft. And Figure H.22

[11] Digital 'scopes finesse this problem by using digital memory to store some pretrigger digitized samples.

[12] The capacitors in Figure H.19's lumped-element delay line are in fact connected to the *midpoint* of each inductor, as can be seen in the official schematic of Figure H.20. It turns out that this is a more efficient implementation of a finite lumped line, as explained to us over a dinner of fine conversation and Persian cuisine by Larry Baxter ("Mr. Capacitive Sensors"; see the book by the same name).

[13] By a factor of πnD, approximately, where n is the number of turns per unit length and D is the spiral diameter. Because the coarse-pitch spiral is closely surrounded by the shield, you can think of the signal as propagating, corkscrew fashion, *along* the spiral; that approximation then gives you this simple expression, without the need to calculate the inductance and capacitance per unit length.

[14] That's *cathode ray tube*, for those born in this millennium, and thus deprived of the opportunity to admire one.

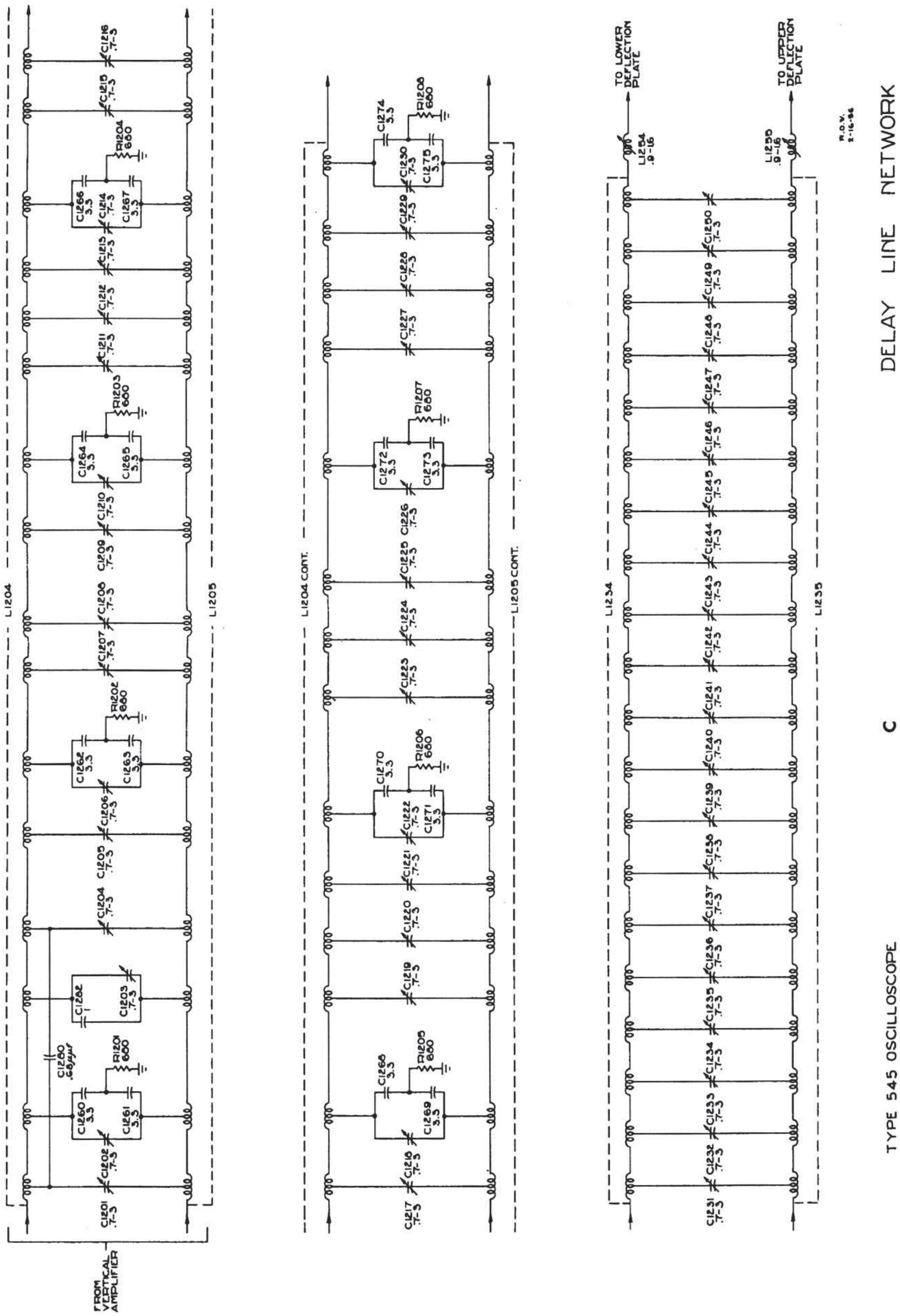

Figure H.20. Working hard to achieve perfection: Tektronix used fifty inductor pairs and fifty trimmers for the delay line in their type 545 ’scope (pictured in Figure H.19) from the 1950s. Reproduced with permission of Tektronix Inc.

shows the observed delay when a differential triangular pulse-pair waveform, launched into one end, is received at the other end.

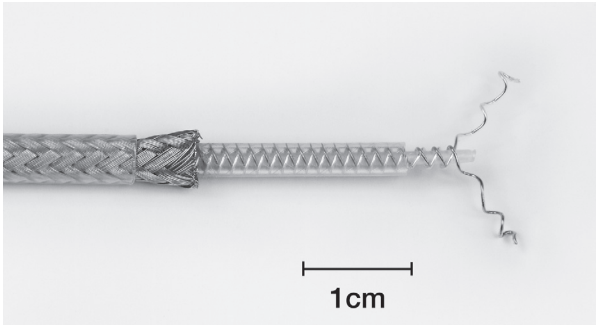

Figure H.21. We borrowed a Tektronix 2213 carcass from Brian Shaban, and, after a bit of surgery, look what we found inside! We measured a 100 ns delay, and a differential impedance of 155 Ω, for the 8.5 ft cable that used to live in this 'scope. The counterwound helix consists of two insulated 30 ga wires, with a pitch of 1.125 mm and an average diameter of 2mm, insulated from the surrounding braided shield of 3.25 mm inside diameter.

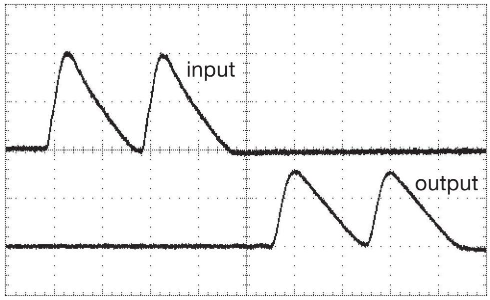

Figure H.22. We launched this analog waveform down the cable of Figure H.21 (well, what remained of it, anyway) to see the signal delay of ∼95 ns. The delayed signal is of good fidelity, with a test signal bandwidth somewhat greater than that of the 60 MHz 'scope. Vertical: 1 V/div; horizontal: 20 ns/div.