Chapter 2

SIGNAL INTEGRITY

Digital systems rely on signaling from drivers to receivers to pass information between their components. Reliable signaling is achieved when the signaling specifications are met under full adverse noise conditions as well as device behavior variations due to both process deviations in device manufacturing and normal changes over the operating temperature. A breakdown in communication leads to glitches where unintended or incorrect data is transferred, a situation called *false signaling*.

Critical data paths within systems often contain safeguards against false signaling, and the primary system for this is data redundancy through parity. Parity can enable the system to detect small-scale signaling failures, while an error correction scheme can be optionally included to correct data faults detected using parity information. However, parity and error correction cannot be relied upon to make a noisy system stable and reliable. In addition, it is often not practical in terms of cost or performance to include parity and error correction on every circuit.

Superimposed on the desired signals are unwanted waveforms (i.e., noise) generated from many sources. The principal sources are crosstalk, impedance mismatch, simultaneous switching noise, and multiple reflections. Each can be independently characterized to facilitate an understanding of the mechanisms that degrade signal quality and to help guide design decisions. In real systems, all act simultaneously and require detailed circuit simulation to obtain good estimates of the total waveform on each signal line.

2.1 Transmission Lines

A transmission line is a two-conductor interconnect (so that it can carry signal frequency components down to DC) that is long compared to the conductor cross section and uniform along its length. Because many interconnects are dominated by long runs over unbroken ground planes (to minimize radiation and EMI susceptibility), they can be accurately modeled as transmission lines, and much of signal integrity analysis is based on them.

If a short length of a transmission line is considered, then the lumped approximation applies and the transmission line can be modeled, as shown in Figure 2.1, with series resistance and inductance and with shunt capacitance and conductance. Applying Kirchhoff's voltage law around the loop, then

$$v(z+\Delta z,t) - v(z,t) = -Ri(z,t) - L\frac{di(z,t)}{dt}. \tag{2.1}$$

Similarly, Kirchhoff's current law applied at $z+\Delta z$ yields

$$i(z+\Delta z,t) - i(z,t) = -Gv(z+\Delta z,t) - C\frac{dv(z+\Delta z,t)}{dt}. \tag{2.2}$$

Divide through by Δz and let $\Delta z \to 0$, then (2.1) and (2.2) transform from difference equations to the differential equations

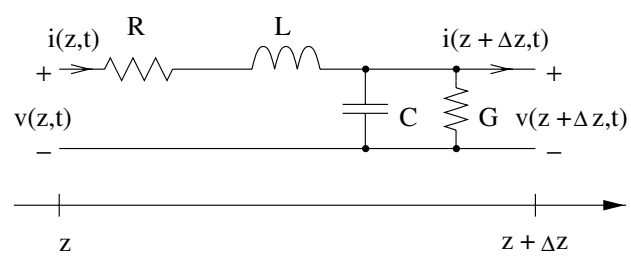


Figure 2.1. Lumped model of a short length of a transmission line.

$$\frac{\partial v(z,t)}{\partial z} = -ri(z,t) - \ell\frac{\partial i(z,t)}{\partial t} \tag{2.3}$$

and

$$\frac{\partial i(z,t)}{\partial z} = -gv(z,t) - c\frac{\partial v(z,t)}{\partial t}, \tag{2.4}$$

where the lumped component values transition to the per-unit-length quantities r, ℓ, c, and g due to normalization by Δz. Simultaneous solution of the transmission line equations (2.3) and (2.4) yields the voltage and current at any point on the transmission line.

2.1.1 Time-Domain Solution

Transmission lines for digital signaling often have low losses. For example, the effect of half an Ohm of loss on a 50Ω transmission line driven with a 50Ω driver is negligible for most applications. To facilitate investigations of the effects of various system imperfections on signal integrity, losses will be neglected. The lossless transmission line equations are recovered from (2.3) and (2.4) by setting $r = g = 0$ yielding

$$\frac{\partial v(z,t)}{\partial z} = -\ell\frac{\partial i(z,t)}{\partial t} \tag{2.5}$$

and

$$\frac{\partial i(z,t)}{\partial z} = -c\frac{\partial v(z,t)}{\partial t}. \tag{2.6}$$

An important property of the lossless transmission line is that pulses propagate undistorted along the length of the line. Consider an arbitrary waveform such as the one in Figure 2.2, where the wave shape is described by the function $f(\tau)$ with τ an independent variable. Since there are no losses and no frequency dependence to ℓ or c, the waveform must move down the line unchanged in shape (see section 2.1.3 for a proof) and can be described mathematically by $f(z,t) = f(\tau) = f(z - \nu t)$, where $\tau = z - \nu t$. A maximum or minimum of the waveform occurs when $\partial\tau/\partial t = 0$, so

$$\frac{\partial \tau}{\partial t} = 0 = \frac{dz}{dt} - \nu,$$

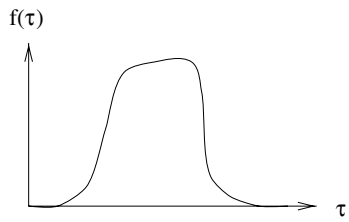


Figure 2.2. Arbitrary waveform for propagation down a transmission line.

or

$$\nu = \frac{dz}{dt},$$

indicating that the maximum or minimum point is moving in the $+z$ direction with velocity ν.

The partial derivatives in (2.5) and (2.6) can be rewritten in terms of τ by noting that

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial \tau}\frac{\partial \tau}{\partial z} = \frac{\partial}{\partial \tau}\frac{\partial (z - \nu t)}{\partial z} = \frac{\partial}{\partial \tau}$$

and

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \tau}\frac{\partial \tau}{\partial t} = -\nu\frac{\partial}{\partial \tau},$$

then

$$\frac{\partial v(\tau)}{\partial \tau} = \ell\nu\frac{\partial i(\tau)}{\partial \tau} \tag{2.7}$$

and

$$\frac{\partial i(\tau)}{\partial \tau} = c\nu\frac{\partial v(\tau)}{\partial \tau}. \tag{2.8}$$

Eliminating $\frac{\partial i(\tau)}{\partial \tau}$ between (2.7) and (2.8) and cancelling $\frac{\partial v(\tau)}{\partial \tau}$ yields

$$\nu = \frac{1}{\sqrt{\ell c}}, \tag{2.9}$$

so the velocity of an arbitrary pulse can be directly computed from the per-unit-length inductance and capacitance of the transmission line. Integrating (2.7) with respect to τ while assuming that no static charge is on the transmission line (so that the integration constant vanishes), then

$$v(\tau) = \sqrt{\frac{\ell}{c}} i(\tau),$$

where (2.9) is used. Therefore, the voltages and currents of an arbitrary waveform on a lossless transmission line are in phase and related by the impedance

$$Z_o = \sqrt{\frac{\ell}{c}},$$

called the *characteristic impedance* of the transmission line.

For a transmission line of length d, the time for a wave to travel the length of the line is called the *delay* or the *time of flight* (TOF) and can be computed as

$$\text{TOF} = d\sqrt{\ell c}. \tag{2.10}$$

The lossless transmission line is completely specified by its characteristic impedance and delay. Note that

$$\ell = Z_o \frac{\text{TOF}}{d} \tag{2.11}$$

and

$$c = \frac{1}{Z_o} \frac{\text{TOF}}{d}. \tag{2.12}$$

The analyses can be repeated with $\tau = z + \nu t$ with identical results, except that the waveform travels in the $-z$ direction with velocity $\nu = 1/\sqrt{\ell c}$ and the voltage and current are related by

$$v(\tau) = -Z_o i(\tau).$$

Effective Dielectric Constant

Since the TOF can be found given the length of a transmission line and the velocity of a wave on it, the velocity is often the unknown parameter that must be found.

In a transmission line where the electric and magnetic fields are completely encased in a dielectric with dielectric constant ϵ_r, then the velocity is

$$\nu = \frac{c_o}{\sqrt{\epsilon_r}}, \tag{2.13}$$

where $c_o = 3 \times 10^8$m/s is the velocity of a wave in a vacuum ($\epsilon_r = 1$) [which is also called the *free-space speed of light*]. For transmission lines such as stripline, dual stripline, embedded microstrip, and coax, the velocity is easily computed once the dielectric constant of the filler material is known.

When the electric and magnetic fields run through two dielectrics, the wave still propagates with some velocity. Generalizing (2.13) yields

$$\nu = \frac{c_o}{\sqrt{\epsilon_{\text{eff}}}},$$

where ϵ_{eff} is an effective dielectric constant. For a transmission line like microstrip, the two dielectrics are air with $\epsilon_r = 1$ and the substrate. The effective dielectric constant must lie between these two, and since most of the field is below the strip in the substrate, the effective dielectric constant must be closer to the dielectric constant of the substrate than to that of air.

Effective dielectric constants are convenient because they offer a handy shortcut in many situations and can be easily estimated for approximate calculations. For example, if $\epsilon_r = 4$ for the substrate in microstrip, then $\epsilon_{\text{eff}} \approx 3$.

For the lossless case, the formulas in (2.11) and (2.12) are easily modified for the effective dielectric constant to be

$$\ell = \frac{Z_o\sqrt{\epsilon_{\text{eff}}}}{c_o} \tag{2.14}$$

and

$$c = \frac{\sqrt{\epsilon_{\text{eff}}}}{Z_o c_o}. \tag{2.15}$$

2.1.2 Directional Independence

The analysis of lossless transmission lines can be carried further to show that two waveforms traveling in opposite directions do not interact. Let v^+ denote a voltage

waveform launched in the $+z$ direction, while v^- indicates one traveling in the $-z$ direction. Due to the linearity of Maxwell's equations, and assuming linear materials, the total voltage must be the superposition of these two, so

$$v(z,t) = v^+(z - \nu t) + v^-(z + \nu t). \tag{2.16}$$

The total current is then

$$i(z,t) = \frac{1}{Z_o} v^+(z - \nu t) - \frac{1}{Z_o} v^-(z + \nu t). \tag{2.17}$$

Substituting these expressions into (2.5) and (2.6) yields

$$\frac{\partial v^+}{\partial z} + \frac{\partial v^-}{\partial z} = -\frac{\ell}{Z_o}\frac{\partial v^+}{\partial t} + \frac{\ell}{Z_o}\frac{\partial v^-}{\partial t}$$

and

$$\frac{\partial v^+}{\partial z} - \frac{\partial v^-}{\partial z} = -cZ_o\frac{\partial v^+}{\partial t} - cZ_o\frac{\partial v^-}{\partial t}.$$

Adding these two results in

$$\frac{\partial v^+}{\partial z} = -\frac{1}{2}\left(\frac{\ell}{Z_o} + cZ_o\right)\frac{\partial v^+}{\partial t} + \frac{1}{2}\left(\frac{\ell}{Z_o} - cZ_o\right)\frac{\partial v^-}{\partial t},$$

while subtracting provides

$$\frac{\partial v^-}{\partial z} = -\frac{1}{2}\left(\frac{\ell}{Z_o} - cZ_o\right)\frac{\partial v^+}{\partial t} + \frac{1}{2}\left(\frac{\ell}{Z_o} + cZ_o\right)\frac{\partial v^-}{\partial t}.$$

These can be further simplified by noting that $\left(\frac{\ell}{Z_o} + cZ_o\right) = 2/\nu$ while $\left(\frac{\ell}{Z_o} - cZ_o\right) = 0$, then

$$\frac{\partial v^+}{\partial z} = -\frac{1}{\nu}\frac{\partial v^+}{\partial t}$$

and

$$\frac{\partial v^-}{\partial z} = +\frac{1}{\nu}\frac{\partial v^-}{\partial t}.$$

Note that these two equations are decoupled, with each equation a function only of v^+ or v^-; therefore, the two waves cannot interact and travel along the transmission line without influencing each other. The waves interact only at boundaries between transmission lines with different impedances or between a transmission line and a component. This property is exploited to facilitate analyses and to construct bounce diagrams.

2.1.3 Frequency-Domain Solution

When losses are significant on transmission lines, the frequency domain is convenient for analytic solutions. After Fourier [4] transformation with (1.4), the lossy transmission line equations in (2.3) and (2.4) become

$$\frac{\partial V(z,\omega)}{\partial z} = -(r + \jmath\omega\ell)I(z,\omega) \tag{2.18}$$

and

$$\frac{\partial I(z,\omega)}{\partial z} = -(g + \jmath\omega c)V(z,\omega), \tag{2.19}$$

where V and I are the Fourier transforms of v and i. Eliminating I from (2.18) and (2.19) provides the second-order differential equation

$$\frac{\partial^2 V}{\partial z^2} - (r + \jmath\omega\ell)(g + \jmath\omega c)V = 0. \tag{2.20}$$

Define $\gamma^2 = (r + \jmath\omega\ell)(g + \jmath\omega c)$, then

$$\gamma(\omega) = \sqrt{(r + \jmath\omega\ell)(g + \jmath\omega c)}, \tag{2.21}$$

and the solution to (2.20) is

$$V(z,\omega) = \mathrm{A}(\omega)e^{-\gamma(\omega)z} + \mathrm{B}(\omega)e^{\gamma(\omega)z}. \tag{2.22}$$

Note in particular that A, B, and γ are functions of ω. γ is called the *complex propagation constant*.

The complex propagation constant can always be written in terms of its real and imaginary parts as

$$\gamma(\omega) = \alpha(\omega) + \jmath\beta(\omega);$$

then (2.22) becomes

$$V(z,\omega) = \mathrm{A}(\omega)e^{-\alpha(\omega)z}e^{-\jmath\beta(\omega)z} + \mathrm{B}(\omega)e^{\alpha(\omega)z}e^{\jmath\beta(\omega)z}.$$

The time-domain voltage can then be found using the inverse Fourier transform given in chapter 1 by (1.5) to obtain

$$\begin{aligned} f(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\mathrm{A}(\omega)e^{-\alpha(\omega)z}e^{-\jmath\beta(\omega)z} + \mathrm{B}(\omega)e^{\alpha(\omega)z}e^{\jmath\beta(\omega)z}\right)e^{\jmath\omega t}d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\Big(\mathrm{A}(\omega)e^{-\alpha(\omega)z}e^{-\jmath(\beta(\omega)z-\omega t)} \\ &\quad + \mathrm{B}(\omega)e^{\alpha(\omega)z}e^{\jmath(\beta(\omega)z+\omega t)}\Big)\,d\omega, \end{aligned} \tag{2.23}$$

where it can be seen in the second form that the total voltage consists of the superposition of many complex exponentials. Each $\mathrm{A}(\omega)e^{-\alpha(\omega)z}e^{-\jmath(\beta(\omega)z-\omega t)}$ represents a weighted sinusoidal wave traveling in the $+z$ direction that is attenuated exponentially with distance. A fixed point on the exponential can be identified when $\beta(\omega)z - \omega t =$ constant. Solving for z and taking the derivative with respect to time yields the phase velocity

$$\nu_p = \frac{dz}{dt} = \frac{\omega}{\beta(\omega)}. \tag{2.24}$$

Since β is a function of frequency, then the phase velocity is a function of frequency. In a similar fashion, each $\mathrm{B}(\omega)e^{\alpha(\omega)z}e^{\jmath(\beta(\omega)z+\omega t)}$ represents a weighted and attenuated wave traveling in the $-z$ direction. Due to their roles in wave propagation, α is called the *attenuation constant* and β is called the *propagation constant*.

The phase velocity and attenuation of the sinusoidal waves are different at every frequency, so the shape of the time-domain waveform must change as it moves down the line. Attenuation is stronger for higher-frequency components, so the waveform tends to spread, or *disperse*, with distance as low-frequency components take over. For this reason, the effects of losses on wave shape is called *dispersion*. Sometimes dispersion is attributed to the source of the losses, such as conductor loss dispersion or dielectric loss dispersion.

Because each frequency component has a different phase velocity, the wave velocity is not equal to the phase velocity of any given component. For lossy lines,

the wave velocity is found by finding the TOF of a pulse between two points and dividing into the distance traveled.[1]

The characteristic impedance can be found by solving for I in (2.18) with the general voltage solution in (2.22) to obtain

$$I(z,\omega) = \sqrt{\frac{g + \jmath\omega c}{r + \jmath\omega\ell}} \left(\mathrm{A}(\omega) e^{-\gamma(\omega)z} - \mathrm{B}(\omega) e^{\gamma(\omega)z} \right). \tag{2.25}$$

The characteristic impedance for a lossy transmission line is then

$$Z_o(\omega) = \sqrt{\frac{r + \jmath\omega\ell}{g + \jmath\omega c}}, \tag{2.26}$$

and it is, in general, frequency-dependent.

The directional components of voltage and current are clearly apparent with a change of notation in (2.22) to

$$V(z,\omega) = V^+(\omega) e^{-\gamma(\omega)z} + V^-(\omega) e^{+\gamma(\omega)z}, \tag{2.27}$$

then the current from (2.25) is

$$I(z,\omega) = \frac{1}{Z_o(\omega)} \left[V^+(\omega) e^{-\gamma(\omega)z} - V^-(\omega) e^{+\gamma(\omega)z} \right]. \tag{2.28}$$

These two expressions are often a good starting point in solving circuit problems involving transmission lines.

Low-Loss Transmission Lines

When losses on a transmission line are small, additional mathematical manipulation can yield good insight into waveform propagation on transmission lines. The complex propagation constant can be rearranged as

$$\begin{aligned} \gamma &= \sqrt{(r + \jmath\omega\ell)(g + \jmath\omega c)} \\ &= \sqrt{\jmath\omega\ell(1 + \frac{r}{\jmath\omega\ell})\jmath\omega c(1 + \frac{g}{\jmath\omega c})} \\ &= \jmath\omega\sqrt{\ell c}\sqrt{1 + \frac{r}{\jmath\omega\ell}}\sqrt{1 + \frac{g}{\jmath\omega c}}. \end{aligned} \tag{2.29}$$

1. For narrow-band signals, the group velocity can be computed from the frequency dependence of the phase velocity to find the velocity of a wave packet. Digital signals are baseband, so the group velocity is not applicable since the narrow-band assumption does not apply.

For low losses, the approximations

$$\begin{aligned} \sqrt{1+\tfrac{r}{\jmath\omega\ell}} &\approx 1+\tfrac{1}{2}\tfrac{r}{\jmath\omega\ell}, \quad \tfrac{r}{\omega\ell} \ll 1 \\ \sqrt{1+\tfrac{g}{\jmath\omega c}} &\approx 1+\tfrac{1}{2}\tfrac{g}{\jmath\omega c}, \quad \tfrac{g}{\omega c} \ll 1 \end{aligned} \tag{2.30}$$

can be applied to (2.29) to obtain

$$\begin{aligned} \gamma &\approx \jmath\omega\sqrt{\ell c}\left(1+\frac{1}{2}\frac{r}{\jmath\omega\ell}+\frac{1}{2}\frac{g}{\jmath\omega c}-\frac{1}{4}\frac{rg}{\omega^2\ell c}\right) \\ &\approx \jmath\omega\sqrt{\ell c}+\frac{1}{2}\left(\sqrt{\frac{c}{\ell}}r+\sqrt{\frac{\ell}{c}}g\right), \end{aligned} \tag{2.31}$$

where the term $\frac{1}{4}\frac{rg}{\omega^2\ell c}$ is dropped as a negligible second-order small term. The attenuation and propagation constants for low losses are then given by

$$\alpha = \frac{1}{2}\left(\sqrt{\frac{c}{\ell}}r+\sqrt{\frac{\ell}{c}}g\right) \tag{2.32}$$

and

$$\beta = \omega\sqrt{\ell c}, \tag{2.33}$$

and the phase velocity is

$$\nu_p = \frac{1}{\sqrt{\ell c}}.$$

Note that the attenuation constant and the phase velocity are frequency-independent, so all frequency components of the waveform travel together with uniform attenuation. Therefore, the waveform propagates down the transmission line with no change in wave shape except for reduction in amplitude. In this case, the wave velocity does equal the phase velocity. This propagation is dispersionless.

Substituting the attenuation and propagation constants in (2.32) and (2.33) for the low-loss line into (2.23) yields

$$\begin{aligned} f(t) &= \frac{1}{2\pi}e^{-\alpha z}\int_{-\infty}^{\infty} A(\omega)e^{-\jmath\omega\sqrt{\ell c}z}e^{\jmath\omega t}d\omega \\ &\quad +\frac{1}{2\pi}e^{+\alpha z}\int_{-\infty}^{\infty} B(\omega)e^{+\jmath\omega\sqrt{\ell c}z}e^{\jmath\omega t}d\omega \\ &= e^{-\alpha z}a(t-\sqrt{\ell c}z)+e^{+\alpha z}b(t+\sqrt{\ell c}z), \end{aligned} \tag{2.34}$$

where $a(t) = \mathcal{F}^{-1}[A(\omega)]$ and $b(t) = \mathcal{F}^{-1}[B(\omega)]$. The time-shifting theorem for Fourier transforms,

$$\mathcal{F}[f(t-\xi)] = \mathcal{F}[f(t)]e^{-\jmath\omega\xi},$$

is also utilized. The result in (2.34) generalizes to low-loss lines the directional independence established in section 2.1.2 for lossless lines. In addition, the result shows that waveforms travel unchanged except for amplitude on low-loss lines.

Conductor Loss-Dominated Transmission Lines

Losses are often dominated by conductors, so while r is often significant, g is often not. Setting $g = 0$ in the lossy expressions for Z_o and γ yields

$$Z_o = \sqrt{\frac{r + \jmath\omega\ell}{\jmath\omega c}} \tag{2.35}$$

and

$$\gamma = \sqrt{(r + \jmath\omega\ell)\jmath\omega c}. \tag{2.36}$$

At low frequencies, $r \gg \omega\ell$, so

$$Z_o \approx \sqrt{\frac{r}{\omega c}}\, e^{-\jmath\pi/4}$$

and

$$\gamma \approx \sqrt{\omega r c}\, e^{\jmath\pi/4}.$$

For both Z_o and γ, the real and imaginary parts are equal (ignoring the sign) with a square root dependency on frequency. These results are very different from those obtained from the lossless transmission line model, yet the lossless model is often used in signal integrity work where baseband digital signals include significant low-frequency content.

An approximate metric is available for determining if lossless transmission line modeling is appropriate. At low frequencies, the lumped approximation applies,

and the lossy transmission line is simply a loop of wire that can be modeled as a series RLC circuit. For losses to be negligible, the RLC circuit must be strongly underdamped, a condition that occurs when

$$R \ll 2\sqrt{\frac{L}{C}}.$$

Using per-unit-length quantities for a transmission line with length d, then

$$d \ll \frac{2}{r}\sqrt{\frac{\ell}{c}}. \tag{2.37}$$

If the line length is sufficiently short, then lossless transmission line modeling can be appropriate.

To demonstrate the line length dependence, consider a 50Ω transmission line with $\epsilon_{\text{eff}} = 3$ and 10Ω/m of loss. The per-unit-length quantities from (2.14) and (2.15) are 289nH/m and 115pF/m, respectively. The line length limit from (2.37) is $d \ll 10$m. Simulation results for lossy and lossless versions of this transmission line are shown in Figure 2.3. The results show that the lossless model is accurate for line lengths less than 1m, so the prediction from (2.37) works very well for this case. Board-level signal integrity work deals with lines less than a meter long, so lossless transmission line modeling can be appropriate if the edge rates are sufficiently slow. Waveforms with fast edge rates experience significant frequency-dependent losses due to the skin effect, so frequency-dependent lossy simulations can be required (see chapter 10, section 10.3).

At high frequencies, $r \ll \omega\ell$, then

$$Z_o \approx \sqrt{\frac{\ell}{c}}\left(1 - \jmath\frac{1}{2}\frac{r}{\omega\ell}\right)$$

and

$$\gamma \approx \jmath\omega\sqrt{\ell c}\left(1 - \jmath\frac{1}{2}\frac{r}{\omega\ell}\right).$$

In this case, the real part of Z_o is constant and dominates the imaginary part, while for γ the imaginary part dominates with a simple linear dependence on frequency. When high-frequency losses are negligible, the characteristic impedance and propagation constant simplify to that of the lossless transmission line.

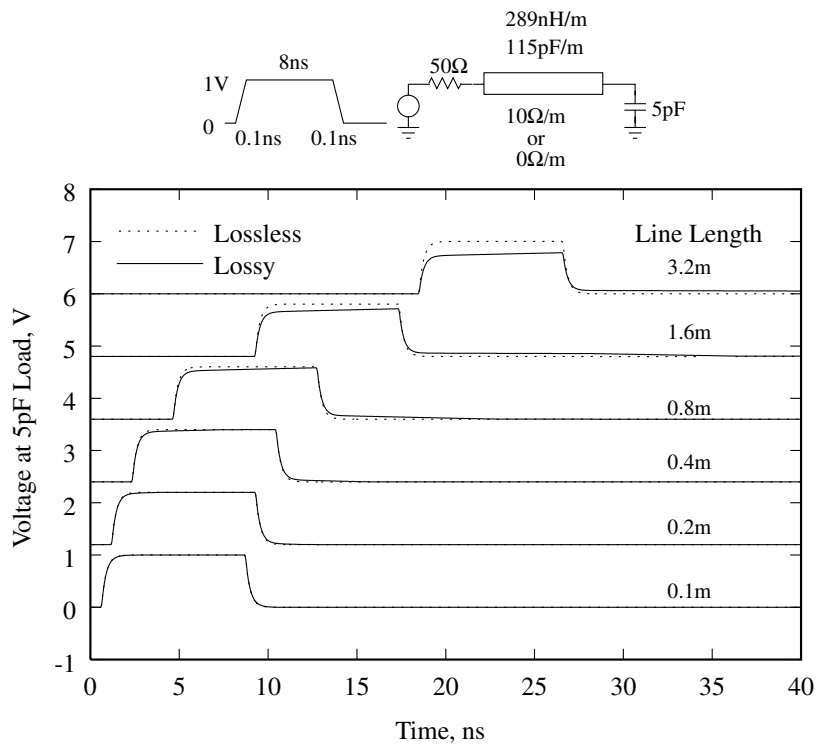


Figure 2.3. The accuracy of the lossless transmission line model depends on the line length.

General Formulas Using Real Math

Without assumptions on r and $\omega\ell$, complex math is required for calculating Z_o and γ, so calculator-based calculations using (2.35) and (2.36) are somewhat difficult. Convenient formulas based only on real math are possible after some manipulation. First,

$$Z_o = \sqrt{\frac{\ell}{c}}\sqrt{1 - \jmath\frac{r}{\omega\ell}},$$

then the real part is

$$\mathrm{Re}[Z_o] = \frac{1}{2}\sqrt{\frac{\ell}{c}}\left(\sqrt{1-\jmath\frac{r}{\omega\ell}}+\sqrt{1+\jmath\frac{r}{\omega\ell}}\right).$$

The term in parenthesis can be rearranged as

$$\begin{aligned}\sqrt{1-\jmath\frac{r}{\omega\ell}}+\sqrt{1+\jmath\frac{r}{\omega\ell}} &= \pm\sqrt{\left(\sqrt{1-\jmath\frac{r}{\omega\ell}}+\sqrt{1+\jmath\frac{r}{\omega\ell}}\right)^2}\\ &= \pm\sqrt{2\left(1+\sqrt{1+\left(\frac{r}{\omega\ell}\right)^2}\right)}.\end{aligned}$$

The real part of the characteristic impedance is then

$$\mathrm{Re}[Z_o] = \sqrt{\frac{\ell}{c}}\sqrt{\frac{1}{2}\left(1+\sqrt{1+\left(\frac{r}{\omega\ell}\right)^2}\right)},$$

where the positive root is taken to represent a passive structure. For the imaginary part of the characteristic impedance,

$$\mathrm{Im}[Z_o] = \frac{1}{2\jmath}\sqrt{\frac{\ell}{c}}\left(\sqrt{1-\jmath\frac{r}{\omega\ell}}-\sqrt{1+\jmath\frac{r}{\omega\ell}}\right).$$

The term in parenthesis can be rearranged as

$$\begin{aligned}\sqrt{1-\jmath\frac{r}{\omega\ell}}-\sqrt{1+\jmath\frac{r}{\omega\ell}} &= \pm\sqrt{\left(\sqrt{1-\jmath\frac{r}{\omega\ell}}-\sqrt{1+\jmath\frac{r}{\omega\ell}}\right)^2}\\ &= \pm\sqrt{2\left(1-\sqrt{1+\left(\frac{r}{\omega\ell}\right)^2}\right)},\end{aligned}$$

then

$$\mathrm{Im}[Z_o] = -\sqrt{\frac{\ell}{c}}\sqrt{\frac{1}{2}\left(-1+\sqrt{1+\left(\frac{r}{\omega\ell}\right)^2}\right)},$$

where the negative sign is taken to match up to the low-loss approximation above. The complex propagation constant can be similarly addressed to find that

$$\mathrm{Re}[\gamma] = \alpha = \omega\sqrt{\ell c}\sqrt{\frac{1}{2}\left(-1+\sqrt{1+\left(\frac{r}{\omega\ell}\right)^2}\right)},$$

and

$$\mathrm{Im}[\gamma] = \beta = \omega\sqrt{\ell c}\sqrt{\frac{1}{2}\left(1+\sqrt{1+\left(\frac{r}{\omega\ell}\right)^2}\right)}.$$

2.1.4 Impedance Boundaries

Waves traveling in opposite directions couple at the interfaces between transmission lines with different characteristic impedances. Suppose multiple lossless or lossy transmission lines are cascaded. At the interface between the nth and $(n+1)$th line, the voltage and the current must be continuous. For the voltage,

$$v_n^+ + v_n^- = v_{n+1}^+ + v_{n+1}^-,$$

while for the current,

$$\frac{v_n^+ - v_n^-}{Z_{o,n}} = \frac{v_{n+1}^+ - v_{n+1}^-}{Z_{o,n+1}}.$$

Eliminating v_{n+1}^- yields

$$v_{n+1}^+ = \frac{1}{2}\left(1+\frac{Z_{o,n+1}}{Z_{o,n}}\right)v_n^+ + \frac{1}{2}\left(1-\frac{Z_{o,n+1}}{Z_{o,n}}\right)v_n^-, \tag{2.38}$$

while eliminating v_{n+1}^+ finds

$$v_{n+1}^- = \frac{1}{2}\left(1-\frac{Z_{o,n+1}}{Z_{o,n}}\right)v_n^+ + \frac{1}{2}\left(1+\frac{Z_{o,n+1}}{Z_{o,n}}\right)v_n^-. \tag{2.39}$$

The waveforms in both directions on each line are mixed at the boundary, and waveforms are generated as necessary to fulfill the boundary condition.

For example, if the only wave that exists is v_n^+, after hitting the boundary a reflected wave

$$v_n^- = -v_n^+\frac{1-\frac{Z_{o,n+1}}{Z_{o,n}}}{1+\frac{Z_{o,n+1}}{Z_{o,n}}} = v_n^+\frac{Z_{o,n+1}-Z_{o,n}}{Z_{o,n+1}+Z_{o,n}}$$

is formed according to (2.39), yielding the classic reflection coefficient formula

$$\Gamma = \frac{v_n^-}{v_n^+} = \frac{Z_{o,n+1}-Z_{o,n}}{Z_{o,n+1}+Z_{o,n}}. \tag{2.40}$$

Similarly, a transmitted wave is formed according to (2.38), leading to the transmission coefficient

$$T = \frac{v_{n+1}^{+}}{v_n^{+}} = \frac{2Z_{o,n+1}}{Z_{o,n+1} + Z_{o,n}}.$$

Note that $v_{n+1}^{-} = 0$ throughout as it is assumed that the only wave introduced is v_n^{+}.

One of the primary techniques for maximizing the quality of signals on an interconnect is to eliminate or minimize the magnitude of impedance discontinuities to minimize the generation of reflections. When two components of an interconnect, such as a driver and a transmission line, have the same impedance, then they are said to be *matched* and will produce no reflection.

2.2 Ideal Point-to-Point Signaling

Point-to-point signaling can achieve the highest signal quality to support the fastest clock rates. The basic setup for CMOS signaling is shown in Figure 2.4a, where a driver signals a receiver through a three-conductor transmission line. The receiver is represented by capacitive parasitics to both power and ground.

To facilitate discussions on signaling, the setup can be simplified with a few strong but reasonable assumptions. If the power (V_{DD}) and ground (V_{SS}) connections have sufficient bypass capacitance, they can be wired in parallel. In this case, V_{DD} is said to be an AC ground. If the driver pulls up and down with equal strength, it can be linearized and modeled as a Thevenin equivalent circuit. With these assumptions, the schematic in Figure 2.4a can be replaced with that in Figure 2.4b.

Energy injected into the system must be dissipated to prevent it from accumulating as noise. Resistance is required at the source, at the load, or at both. The driver has some resistance that varies depending on the application. Often, drivers with very low impedance are used to drive large capacitive loads on short interconnects, such as with memory buses. Higher impedances can be used for lighter loads or when timing is more relaxed. Reflections at the driver can be minimized

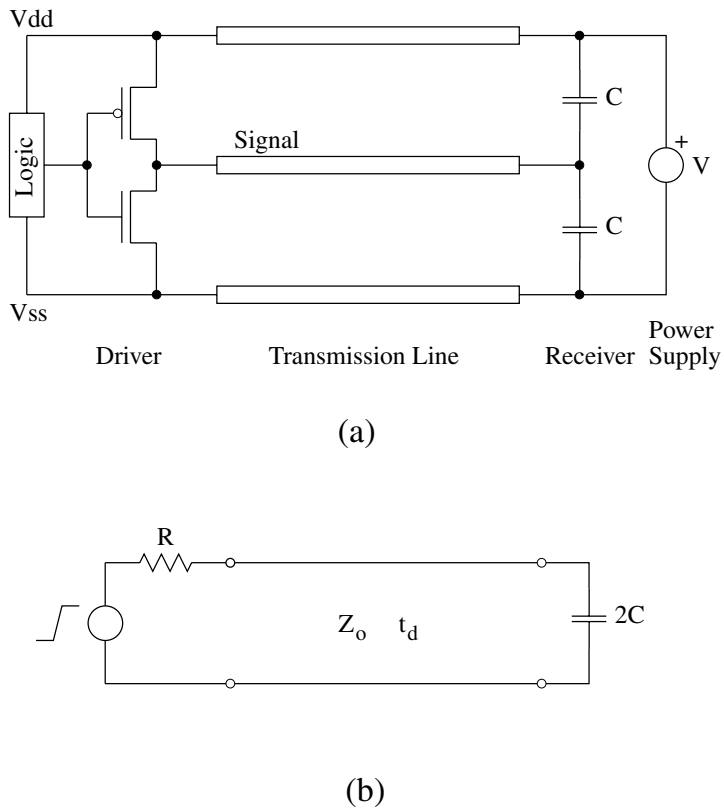


Figure 2.4. CMOS point-to-point signaling: (a) full model, (b) model with infinite bypass, symmetric and linear driver.

by designing its impedance, R, to be equal to the transmission line impedance, Z_o. This configuration is called *source matched*, or *source termination*.

At the load, the small receiver capacitance acts nearly as an open circuit that reflects the entire signal with a reflection coefficient of almost +1. These reflections can be virtually eliminated by adding a resistor in parallel to the load with a value of Z_o. This configuration is called *parallel termination*.

For a point-to-point interconnect, either source or parallel termination is sufficient, although some demanding applications use both. After a short discussion on

edge rates, the cases of source plus parallel termination and of source termination alone are considered.

2.2.1 Fast and Slow Edges

Generally, there are two situations to consider. For slow rise times, the transmitted wave has sufficient time to reflect off the load and return to the driver before the driver has completed its transition. The load then modifies the impedance seen by the driver and affects its switching characteristics. In other words, the driver "feels" the load while it switches. For this situation to occur, the edge rate must be greater than twice the TOF.

For fast rise times with edge rates less than twice the TOF, the driver completes its transition before any of the transmitted wave can reflect from the load and return. During the logic transition, the driver sees only the transmission line's characteristic impedance. In this situation, the loading does not affect the driver's switching behavior.

The waveforms on an interconnect are quite different depending on whether the edges are fast or slow. The switching characteristics of the driver can also vary substantially depending on the TOF.

2.2.2 Source and Parallel Termination

With both ends of the interconnect damping reflections, the signal quality on the transmission line is excellent. Since there are no reflections on the line, the driver always sees a constant load of Z_o, and the behavior of the interconnect is mostly independent of the edge rate.[2]

A voltage divider is formed between the driver impedance and the transmission line's characteristic impedance. With a good source match, the voltage launched onto the transmission line is just half the source voltage. The driver must continuously supply current to hold this voltage. In other words, the static logic condition dissipates power, and this is the main tradeoff for parallel termination. As a rule,

2. The predominant exception is for heavy capacitive loading, but parallel termination would not generally be applied in this case.

enhanced signal quality can always be purchased with higher power dissipation. For example, this signaling approach is taken with ECL, a logic family well known for clean signals and high power dissipation.

SPICE simulations of the topology in Figure 2.4b are shown in Figures 2.5 and 2.6 for both fast and slow rise times. The 50Ω lossless transmission line has a delay of 1ns. The source ramps from 0V to 1V through 50Ω of source resistance in either 1ns or 4ns for the fast and slow cases, respectively. The capacitive load is 5pF with a 50Ω resistor in parallel for the parallel termination case.

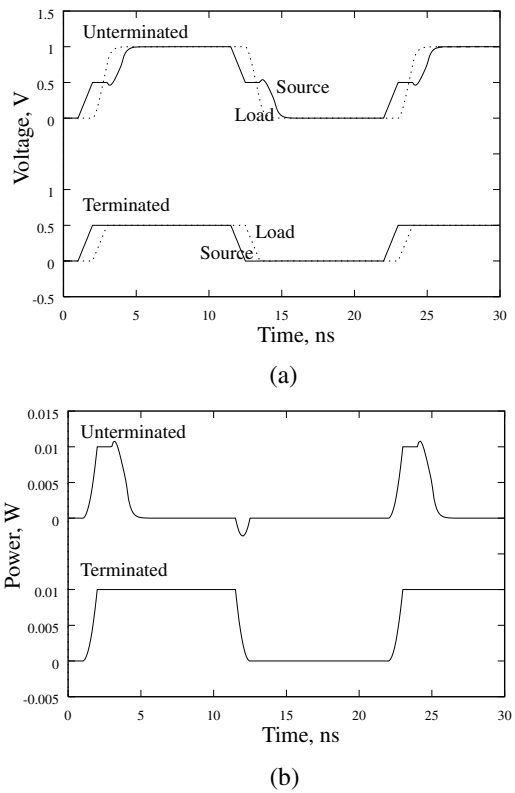


Figure 2.5. Signaling with fast edges on source terminated and on source plus parallel terminated transmission lines; 1ns rise and fall times with 1ns delay: (a) voltage, (b) power.

With both source and parallel termination, the simulations show no difference between the fast and slow cases except for the edge rates. Power is constantly expended while holding logic high. Since all resistances equal the characteristic impedance of the transmission line, the peak voltage on the line is just 0.5V. Signal quality is excellent.

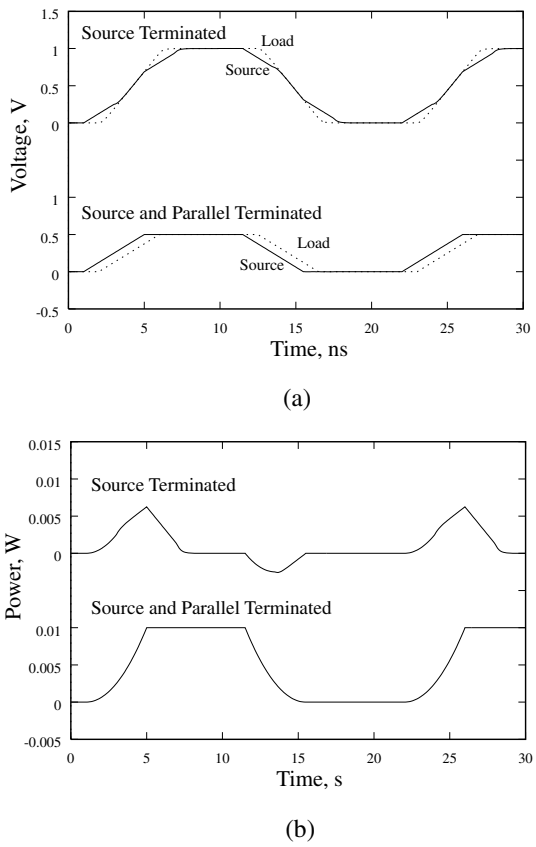


Figure 2.6. Signaling with slow edges on source terminated and on source plus parallel terminated transmission lines; 4ns rise and fall times with 1ns delay: (a) voltage, (b) power.

2.2.3 Source Termination Only

Without parallel termination at the load, the +1 reflection coefficient at the load causes the voltage to essentially double. The large reflected wave travels back to the driver where it is absorbed by the matched source impedance. Due to the presence of the reflected waveform, the signal integrity along the transmission line is not good, but it can be quite good at the load, which is where it matters for a point-to-point net. Once the load capacitance is charged, then the driver no longer needs to supply current to hold the receiver in the logic high state; therefore, the static logic condition does not dissipate power.

Unterminated nets are common with both TTL and CMOS to minimize power dissipation. With good source match, the signal integrity at the load can be quite good. However, the reflected wave is signficant and can cause difficulties on more complex topologies (such as multidrop nets). Also, a bidirectional bus using source termination can be slower because it must wait for the signal to return and terminate in the source impedance before the bus can be turned around into receive mode.

For the fast edges in Figure 2.5, the voltage at the output of the driver is the same for both the unterminated and terminated loads for a time span of 2ns (2×TOF). At that time, the reflection from the unterminated capacitive load arrives at the driver and terminates. Power dissipation is the same for the two cases for the first 2ns, then it drops off for the unterminated case because the load capacitor becomes charged and blocks DC current flow.

For the slow edges in Figure 2.6, the reflected wave returns to the driver while it is still transitioning. The voltage at the source is the sum of the source voltage and the reflected voltage, so the shape is more complex. The waveform shape at the load is still a simple ramp. Power dissipation is reduced because the driver does not have to launch a complete waveform onto the transmission line before the capacitive load begins charging.

2.3 Nonideal Signaling

Ideal signaling for CMOS is represented by point-to-point controlled-impedance nets with matched source impedance. For very high-speed signaling situations, point-to-point nets are required to achieve the required performance. In most cases, point-to-point nets are a luxury that cannot be afforded due to pin counts, board space, power dissipation, cost, and so on. In these cases, many factors arise that disrupt the signal quality, and these factors must be controlled to guarantee system performance. In nonideal situations, the basic signaling requirements that must be met are the switching incidence, which is strongly related to the receiver threshold, and the avoidance of false signaling.

2.3.1 Synchronous vs. Asynchronous

Inputs to a device are either asynchronous or synchronous. Asynchronous inputs are sensitive to voltage at all times, and they are commonly used for clocks, resets, and interrupts. Synchronous inputs sample voltage only at times specified by a clock, and they are commonly used for data, address, mode, and configuration signals. For asynchronous receivers, any time the voltage falls below V_{IL} or rises above V_{IH} for a sufficient time, a logic transition will be registered. If noise causes the voltage level to inadvertently approach valid logic levels, then a system fault can occur. For synchronous receivers, voltage levels only matter for times around clock transitions, so synchronous circuits are much more tolerant of poor signal integrity than are asynchronous circuits. A conservative design approach is to treat all signals as asynchronous.

2.3.2 Switching Incidence

Driver impedances are often not matched to the characteristic impedance of the interconnect, and higher driver impedances are often used to save power and silicon area. Due to voltage division, higher driver impedances launch lower amplitude voltage waves towards the receiver. If the voltage at the receiver is sufficient when

the voltage wave arrives, then the receiver switches on the first incidence of the wave, and signaling is said to be *first incidence*.

If the first incidence is insufficient, then the voltage must build to a sufficient level to trigger the receiver. The voltage wave reflects off of the receiver, propagates back to the source, reflects there, and propagates back to the receiver. If the voltage is sufficient at this second arrival, then the receiver switches, and the signaling is said to be *second incidence*. If the signal is again insufficient, then switching may occur at the third incidence, fourth incidence, and so on. Anything other than first incidence switching involves a time penalty of two TOFs per incidence. Note that a matched source impedance requires first incidence switching because there is no reflected wave from the source.

Second incidence switching is demonstrated with the SPICE simulation in Figure 2.7. On the first incidence, the voltage at the receiver is just

$$V_{1st} = 5\frac{50}{50+250} \times 2 = 1.67V,$$

where the factor of 2 accounts for voltage doubling at the small capacitive load. A second incidence is required to take the voltage above V_{IH} (2V). First incidence switching requires either a drop in V_{IH} below 1.67V (plus some noise margin) or a

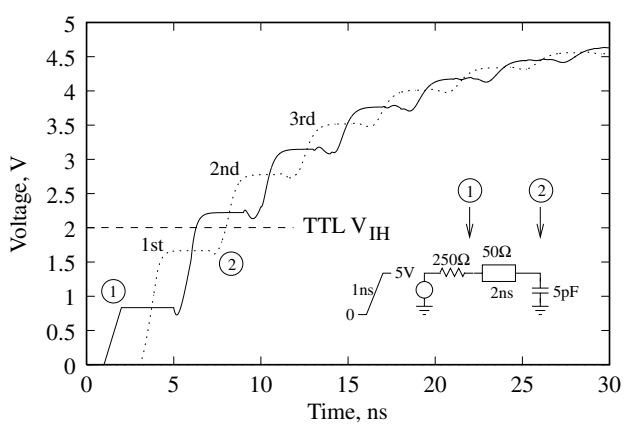


Figure 2.7. Example SPICE simulation showing second incidence switching.

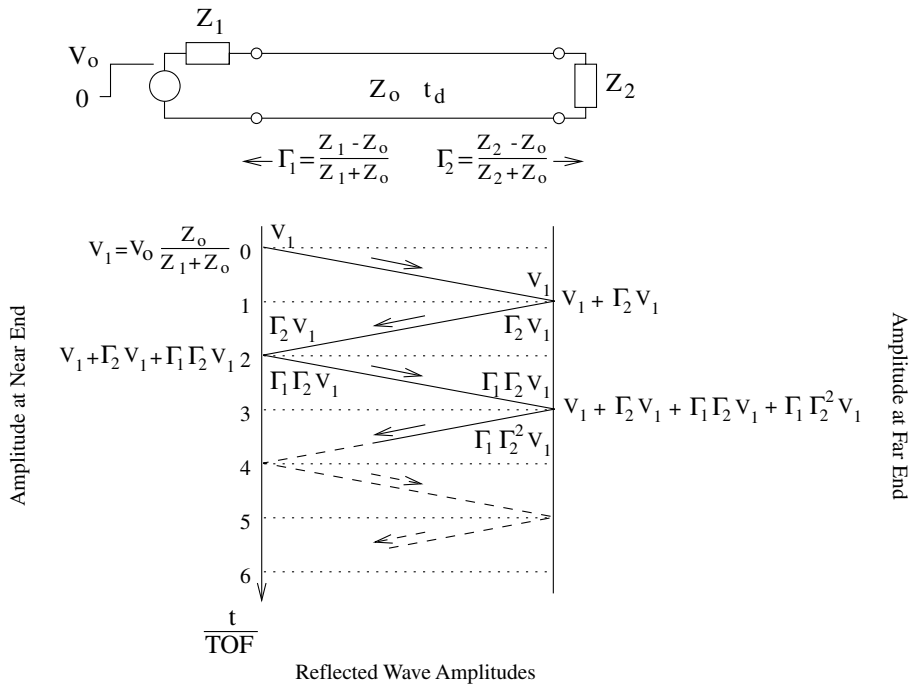


Figure 2.8. Bounce diagram for arbitrary reflection coefficients.

decrease in the driver impedance. For example, a driver impedance of 158.3Ω would produce first incidence switching with a noise margin of 0.4V.

Multiple reflections on a lossless transmission line can be visualized with the help of a bounce diagram, which tracks the amplitude of a reflected wave as it bounces off of imperfect terminations. A bounce diagram for arbitrary terminations is shown in Figure 2.8. The pulse from the voltage source is divided across the source impedance and the transmission line's characteristic impedance to launch a pulse onto the transmission line. One TOF later, the wave arrives at the far end termination. The reflected amplitude is given by the reflection coefficient Γ_2 and the total voltage is the sum of the incident and reflected waves. This total is tallied to the right. After another TOF, the wave arrives back at the source, where a reflected wave is generated with amplitude determined by Γ_1. The total voltage is the incident plus reflected waves plus the voltage that was already there. This total

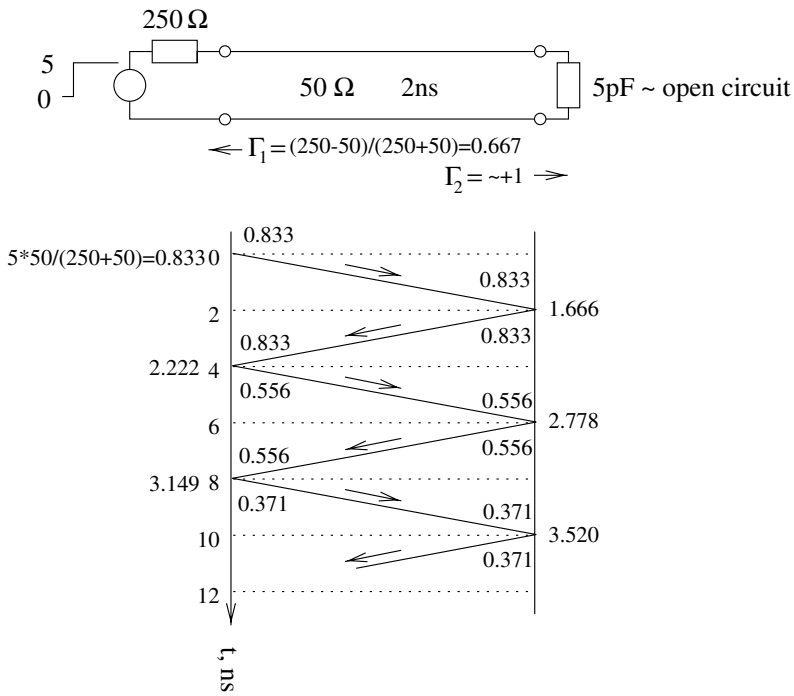


Figure 2.9. Bounce diagram for the circuit in Figure 2.7.

is tallied to the left. The process continues at each interface. Running totals down the sides of the bounce diagram show the voltage amplitudes as a function of time in increments of the TOF.

As a numerical example, a bounce diagram for the circuit in Figure 2.7 is constructed in Figure 2.9. To simplify the calculations, the load capacitance is assumed to be vanishingly small so that the reflection coefficient at the load is simply +1. The amplitudes and timing closely match the SPICE simulation in Figure 2.7.

2.4 Discontinuities

Except for very high-performance systems, high-quality point-to-point nets are rare, and in practice, signal lines experience real-world discontinuities where the

impedance changes. As shown in section 2.1.4, any change in impedance alters the signal while generating a reflected wave. A few of the basic discontinuities are series inductance, shunt capacitance, capacitive loads, and impedance steps.

Many interconnects introduce series inductance, such as wirebonds, connectors, and vias, in PCBs and packages. Note that the complete current loop must be considered. For example, a signal wirebond located far from a power or ground wirebond creates a far more inductive loop than one located next to a power or ground wirebond. For very high-speed edges, when the lumped approximation fails, series inductive discontinuities can be modeled as high-impedance transmission lines.

Shunt capacitance results whenever a signal line passes near other metal. For example, a via through a hole in a ground plane results in series inductance for the via plus shunt capacitance from the via to plane. For very high-speed edges, shunt capacitive discontinuities can be modeled as low-impedance transmission lines.

Unterminated signaling results in a capacitive load at the far end due to the loading of the receiver. The capacitive load creates a reflection while affecting timing due to RC charging times.

Impedance steps occur when the characteristic impedance of the transmission line changes. For example, a 70Ω microstrip PCB line connected to a 50Ω coaxial cable introduces a significant impedance step.

These basic discontinuities are examined in detail. A general result is that discontinuities have characteristic time constants that can be used to estimate the impact of a discontinuity on an interconnect.

2.4.1 Laplace Transform

Analytical time-domain analysis requires the use of the Laplace transform. The Laplace transform pair is defined as

$$F(s) = \int_{0^-}^{\infty} f(t)e^{-st}\, dt$$

and

$$f(t) = \frac{1}{2\pi \jmath} \int_{\sigma - \jmath\infty}^{\sigma + \jmath\infty} F(s) e^{st} ds.$$

The Laplace transform pair is denoted symbolically as $F(s) = \mathcal{L}[f(t)]$ and $f(t) = \mathcal{L}^{-1}[F(s)]$. An important fact from Laplace transform theory is that

$$\mathcal{L}\left[\frac{df(t)}{dt}\right] = sF[s] - f[0^-].$$

For interconnect studies, the Laplace transforms of step functions are needed. For the unit step function

$$u(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases},$$

$$\mathcal{L}[u(t)] = \frac{1}{s}. \tag{2.41}$$

The unit ramp function is defined as

$$r(t) = \begin{cases} 0, & t < 0 \\ t/t_o, & 0 \leq t \leq t_o \\ 1, & t > t_o \end{cases},$$

for which

$$\mathcal{L}[r(t)] = \frac{1 - e^{-st_o}}{t_o s^2}.$$

The unit exponential ramp is

$$e(t) = (1 - e^{-t/t_o})u(t),$$

which is highly convenient because it is expressible without limits and because it captures the exponential rolloff typically seen in circuits. Its Laplace transform is

$$\mathcal{L}[e(t)] = \frac{1}{s} - \frac{1}{s + 1/t_o}. \tag{2.42}$$

2.4.2 Capacitive Load

The voltage-current relationship for a capacitor is $i = C\,dv/dt$. In terms of traveling waves, using (2.27) and (2.28) with $z = 0$, the relationship is

$$\frac{v^+}{Z_o} - \frac{v^-}{Z_o} = C\left(\frac{dv^+}{dt} + \frac{dv^-}{dt}\right).$$

Taking the Laplace transform and collecting terms leads to

$$V^- = \frac{1 - cZ_o s}{1 + cZ_o s} V^+. \tag{2.43}$$

For the unit step input, $v^+(t) = u(t)$, so substituting (2.41) into (2.43) and computing the inverse Laplace transform for the circuit initially at rest yields

$$v^- = (1 - 2\,e^{-t/\tau})u(t),$$

where $\tau = CZ_o$. The total voltage is then

$$v = v^+ + v^- = 2(1 - e^{-t/\tau})u(t),$$

which shows the familiar exponential charging of a capacitive load with voltage doubling for large t/τ. Because of the capacitive loading, the waveform at the load does not rise to one-half its full value ($v = 1/2 \times 2 = 1$) until $t = \tau \ln 2$, so the capacitive load introduces a delay adder of $\tau \ln 2$.

For the exponential step, $v^+(t) = e(t)$, and repeating the above process yields

$$v^- = \left(1 + \frac{2e^{-t/\tau} - (\frac{t_o}{\tau} + 1)e^{-t/t_o}}{\frac{t_o}{\tau} - 1}\right) u(t).$$

For very fast rise times, $t_o \ll \tau$, and

$$v^- \approx (1 - 2\,e^{-t/\tau})u(t),$$

which is the same result as for the unit step input. For very slow rise times, $t_o \gg \tau$, and

$$v^- \approx (1 - \,e^{-t/t_o})u(t),$$

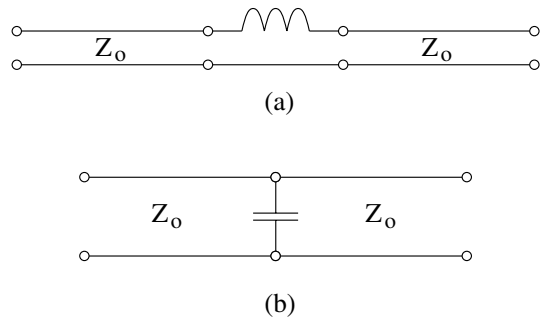


Figure 2.10. Common discontinuities on transmission lines: (a) series inductance, (b) shunt capacitance.

which is the same as the incident wave, so the capacitive load acts like a perfect open circuit with a reflection coefficient of +1.

In general, the effect of a discontinuity can be determined by comparing the rise time to the discontinuity's time constant, τ. If the rise time is short compared to τ, then the behavior is like that for the unit step. On the other hand, if the rise time is long compared to τ, then the behavior is that for the idealized version of the discontinuity.

2.4.3 Series Inductance

For a wave incident from the left of the series inductance discontinuity in Figure 2.10a, the voltage and current are

$$v_1 = v_1^+ + v_1^-$$

and

$$i_1 = \frac{1}{Z_o} v_1^+ - \frac{1}{Z_o} v_1^- .$$

On the right of the inductor, assume that the line is long so there is no reflected wave, then the voltage and current there are

$$v_2 = v_2^+$$

and

$$i_2 = \frac{1}{Z_o} v_2^+.$$

Applying Kirchhoff's voltage law yields

$$v_1^+ + v_1^- = L\frac{di}{dt} + v_2^+, \tag{2.44}$$

while continuity of current requires

$$\frac{1}{Z_o} v_1^+ - \frac{1}{Z_o} v_1^- = \frac{1}{Z_o} v_2^+. \tag{2.45}$$

Eliminating v_1^+ and substituting $i = \frac{1}{Z_o} v_2^+$ yields

$$v_1^- = \tau \frac{dv_2^+}{dt}, \tag{2.46}$$

where $\tau = \frac{1}{2}\frac{L}{Z_o}$ is the time constant of the discontinuity.

Approximate Solution for Weak Reflections

For small inductances causing weak reflections, the incident wave passes through almost unchanged, then $\frac{dv_2^+}{dt} \approx \frac{dv_1^+}{dt}$ and

$$v_1^- \approx \tau \frac{dv_1^+}{dt}. \tag{2.47}$$

The reflected wave is a pulse with width approximately equal to the edge rate of the incident wave. The pulse amplitude is positive for low-to-high transitions and negative for high-to-low.

Exact Solution for Step Input

Eliminating v_1^- between (2.44) and (2.45) with $i = \frac{1}{Z_o} v_2^+$ yields

$$v_1^+ = v_2^+ + \tau \frac{dv_2^+}{dt}, \tag{2.48}$$

Assuming that the circuit is initially at rest, then after Laplace transformation,

$$V_2^+ = V_1^+ \frac{1/\tau}{s + 1/\tau}. \tag{2.49}$$

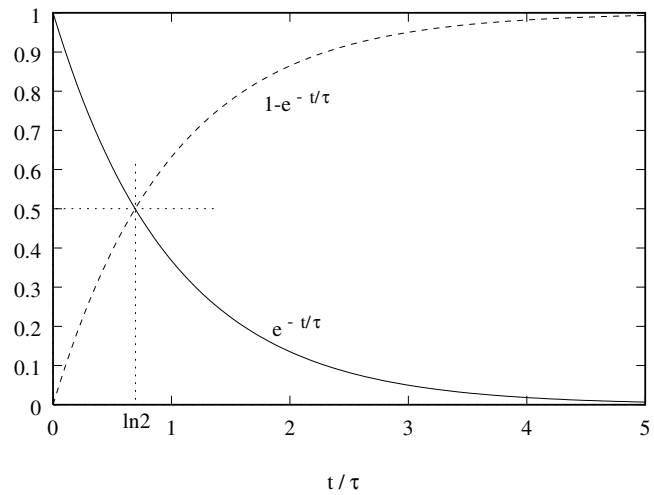


Figure 2.11. Exponential functions describe the reflected and transmitted waves for a step function incident onto series inductance and shunt capacitance.

If v_1^+ is a unit step, then

$$V_2^+ = \frac{1}{s} - \frac{1}{s + 1/\tau}.$$

Inverse Laplace transformation then provides the voltage propagating past the inductor as

$$v_2^+ = (1 - e^{-t/\tau})u(t),$$

and from (2.45), the reflected wave is

$$v_1^- = e^{-t/\tau}u(t).$$

These two functions are plotted in Figure 2.11. The reflected pulse is an exponential spike falling to half its original value by $t = \tau \ln 2$. The transmitted pulse is smoothed, rising to half its final value by time $t = \tau \ln 2$. A SPICE simulation showing step reflection from a series inductor is shown in Figure 2.12.

Considering the frequency domain, at low frequencies the inductor is a short

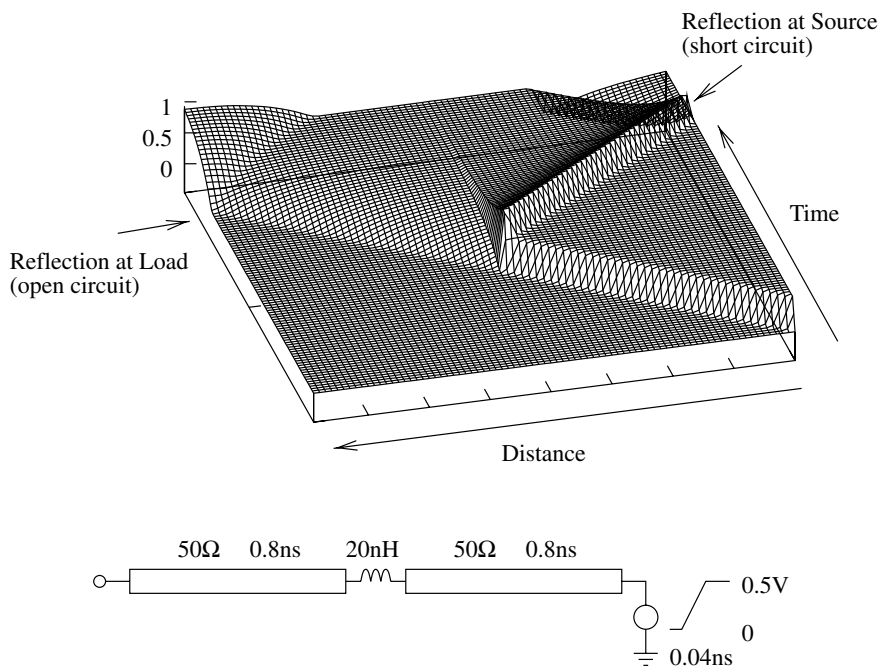


Figure 2.12. A series inductor reflects an exponential pulse back towards the source and low-pass filters the wave that continues down the line.

circuit, so these components readily pass through. At high frequencies, the inductor is an open circuit and reflects these components with a reflection coefficient of +1.

With the high-frequency components reflected back towards the source, the edge rate of the transmitted pulse is rounded off. At the receiver, the slower edge rate causes the wave to be detected later than it would be with a faster edge. If the receiver trip point is $1/2V_{\mathrm{DD}}$, then the delay adder due to the inductor is $\tau \ln 2$.

In total, the series inductor causes noise by creating a reflected pulse and adds delay to the signal due to edge rate degradation. For the step input, the reflected pulse voltage peak equals the amplitude of the step, while the delay adder is $\tau \ln 2$.

Exact Solution for Exponential Input

Let the incident wave v_1^+ be a unit exponential ramp, substitute (2.42) into (2.49), and inverse Laplace transform to get

$$v_2^+ = 1 - \frac{e^{-t/\tau}}{1 - t_o/\tau} - \frac{e^{-t/t_o}}{1 - \tau/t_o}, \tag{2.50}$$

assuming that the circuit is initially at rest.

The delay of the transmitted wave from time $t = 0$ can be found by setting $v_2^+ = 1/2$, then

$$\frac{1}{2} = \frac{e^{-t/\tau}}{1 - t_o/\tau} + \frac{e^{-t/t_o}}{1 - \tau/t_o}, \tag{2.51}$$

but a simple solution for t is not possible. The delay adder is the delay of the transmitted wave minus the delay of the incident wave. For very fast edges with $t_o \ll \tau$, the delay is $t = \tau \ln 2$, so the delay adder is $t_d = \tau \ln 2 - t_o \ln 2 \approx \tau \ln 2$, the same as for the step input. For very slow edges with $t_o \gg \tau$, the delay is $t = t_o \ln 2$ for a delay adder of zero. Just as is seen with the capacitive load, the behavior of the discontinuity is bounded between the fast edge step response and the slow edge ideal case (short circuit, for the series inductor).

An approximate solution to (2.51) is

$$t = \tau \ln 2 + t_o \ln 2,$$

which is asymptotically exact at the extremes of t_o and τ, but reaches -17.4% error near $t_o = \tau$. Using this formula, the delay adder is $t_d = \tau \ln 2$, so to a good approximation, the delay adder for any series inductive discontinuity for any edge rate is just the delay adder for the step function. The approximate solution can be fitted to improve the accuracy near $t_o = \tau$. One solution with maximum error of 1.45% is

$$t = (\tau \ln 2 + t_o \ln 2)\left(1 + 0.194\, e^{-0.174(t_o/\tau + \tau/t_o - 2)}\right).$$

The reflected pulse is found from (2.45), using (2.50), to be

$$v_1^- = \frac{1}{1 - t_o/\tau}\left(e^{-t/\tau} - e^{-t/t_o}\right).$$

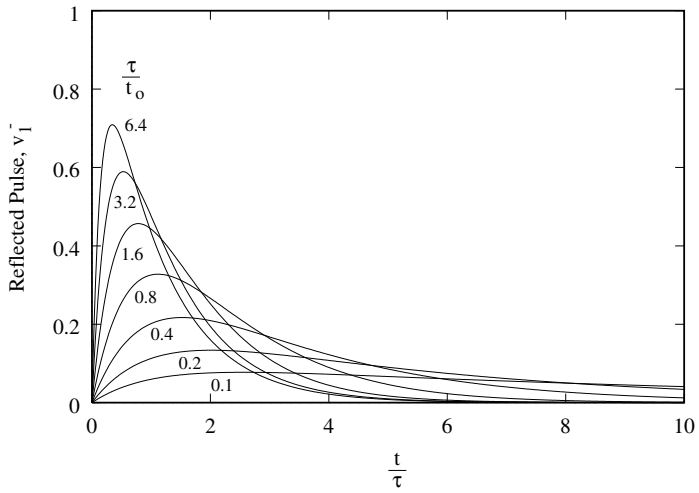


Figure 2.13. Pulses reflected from a series inductive discontinuity for an incident exponential step depend on the rise time of the step.

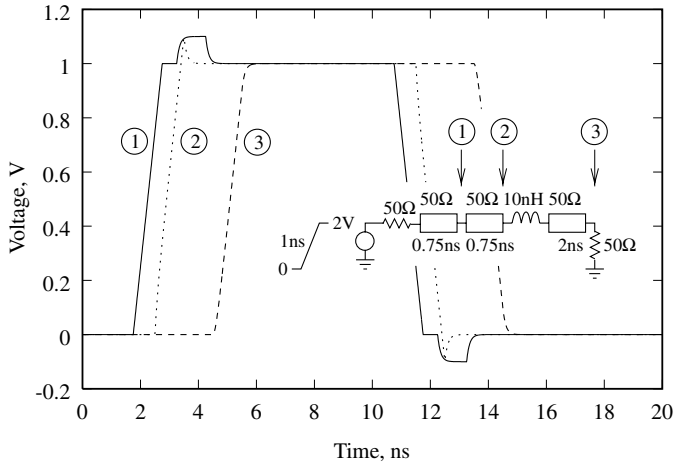
Several pulses are plotted in Figure 2.13, where it can be seen that larger discontinuities (increasing τ) and/or faster edges (decreasing t_o) lead to sharper pulses with higher peak amplitudes but with faster decay rates. The peak occurs at time

$$t_{\text{peak}} = \frac{\ln \frac{t_o}{\tau}}{1/\tau - 1/t_o}.$$

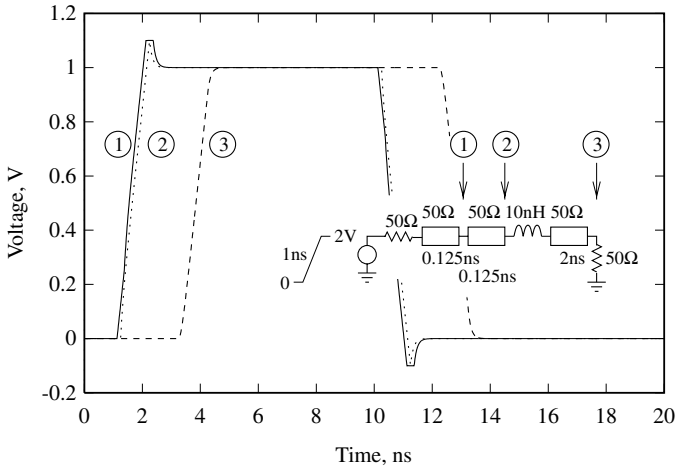
Trapezoidal Ramp Example

Using a linear ramp for the source excitation, a sample SPICE simulation is shown in Figure 2.14, where the schematics are shown in inset. For this problem $\tau = 0.1$ns, so the 1ns rise time appears somewhat slow, and the effect of the discontinuity should be small. In Figure 2.14a, the source is located sufficiently far from the inductor so that it has finished switching by the time the pulse returns. The pulse at point 1 is roughly square since the incident edge is triangular. The peak amplitude of the reflected pulse can be estimated using (2.47) to be

$$v_1^- \approx \frac{1}{2}\frac{10 \times 10^{-9}}{50}\frac{1}{1 \times 10^{-9}} = 0.1\text{V},$$



(a)



(b)

Figure 2.14. Reflected and transmitted waveforms at a series inductance discontinuity: (a) source far from the discontinuity, (b) source near to the discontinuity.

which closely matches the simulation. The voltage swing is 1V because the 2V source evenly divides between the source and transmission line impedances. The pulse width of 1ns is approximately equal to the edge rate. For the transmitted waveform observed at point 3, there is no significant change in the edge rate, but rounding of the corners is observable.

In Figure 2.14b, the source is close to the inductor so that the initial edge is still rising when the reflected pulse returns. The reflected pulse width and amplitude are similar to the far case in Figure 2.14a, but the reflected pulse voltage superimposes onto the rising edge and is not as easily observed. In this case, some of the pulse is "hidden" within the edge itself, where it helps the signal to rise more quickly.

These two cases show how signal integrity is improved when the edge rate lengthens compared to the delays in the circuit. Slower edges produce less noise and allow the noise to be hidden in the rise/fall times if the interconnects are short enough.

2.4.4 Shunt Capacitance

For a wave incident from the left of the shunt capacitance discontinuity in Figure 2.10b, the voltage and current are

$$v_1 = v_1^+ + v_1^-$$

and

$$i_1 = \frac{1}{Z_o} v_1^+ - \frac{1}{Z_o} v_1^- .$$

On the right of the capacitor, assume that the line is long so there is no reflected wave, then the voltage and current there are

$$v_2 = v_2^+$$

and

$$i_2 = \frac{1}{Z_o} v_2^+ .$$

Applying Kirchhoff's current law yields

$$\frac{1}{Z_o}v_1^+ - \frac{1}{Z_o}v_1^- = C\frac{dv_2^+}{dt} + \frac{1}{Z_o}v_2^+, \tag{2.52}$$

while continuity of the voltage requires

$$v_1^+ + v_1^- = v_2^+. \tag{2.53}$$

Eliminating v_1^+ yields

$$v_1^- = -\tau\frac{dv_2^+}{dt}, \tag{2.54}$$

where $\tau = \frac{CZ_o}{2}$ is the time constant of the discontinuity.

Approximate Solution for Weak Reflections

For small capacitances causing weak reflections, the incident wave passes through almost unchanged, then $\frac{dv_2^+}{dt} \approx \frac{dv_1^+}{dt}$ and

$$v_1^- \approx -\tau\frac{dv_1^+}{dt}. \tag{2.55}$$

The reflected wave is a pulse with width approximately equal to the edge rate of the incident wave. The pulse amplitude is negative for low-to-high transitions and positive for high-to-low.

Exact Solutions and Duality

Eliminating v_1^- between (2.52) and (2.53) yields

$$v_1^+ = v_2^+ + \tau\frac{dv_2^+}{dt}.$$

This equation is exactly the same as that for the series inductance in (2.48), so the shunt capacitor behaves exactly the same as the series inductance but with a different time constant. Similarly, the reflected wave in (2.54) is the same as that in (2.46) except for a sign change. The shunt capacitor reflects the same pulse as the series inductor but with opposite sign and different time constant. The series

inductance and shunt capacitance are duals, a reasonable result since inductance encapsulates the effects of the magnetic field while those for the electric field are captured by capacitance.

Cascades of discontinuities are not covered by the mathematical framework above, but it is interesting to note that since the pulses reflected from shunt capacitive and series inductive discontinuities are of opposite sign, they cancel when $\tau_C = \tau_L$, which occurs when $Z_o = \sqrt{L/C}$.

Trapezoidal Ramp Example

A sample SPICE simulation is shown in Figure 2.15, where the edge rate of 1ns is slow compared to the 0.1ns time constant of the shunt capacitor. In Figure 2.15a, the source is located sufficiently far from the capacitor so that it has finished switching by the time the pulse returns. The pulse at point 1 is roughly square since the incident edge is triangular. The peak amplitude of the reflected pulse can be estimated using (2.55) to be

$$v_1^- \approx -\frac{1}{2}\, 4 \times 10^{-12}\, 50\, \frac{1}{1 \times 10^{-9}} = -0.1\text{V},$$

which closely matches the simulation. The pulse width is approximately equal to the edge rate. For the transmitted waveform observed at point 3, there is no significant change in the edge rate, but rounding of the corners is observable.

In Figure 2.15b, the source is close to the capacitor so that the initial edge is still rising when the reflected pulse returns. The reflected pulse width and amplitude are similar to the fast case in Figure 2.15a, but the reflected pulse voltage superimposes onto the rising edge and is not as easily observed. In this case, some of the pulse is "hidden" within the edge itself, where it inhibits the rise of the signal.

2.4.5 Impedance Step

The general rule of thumb is to keep all transmission lines at the same characteristic impedance; otherwise, reflections are generated with amplitudes given by the reflection coefficient in (2.40). For series inductance and shunt capacitance, the frequency-dependent impedance of these discontinuities causes frequency-dependent

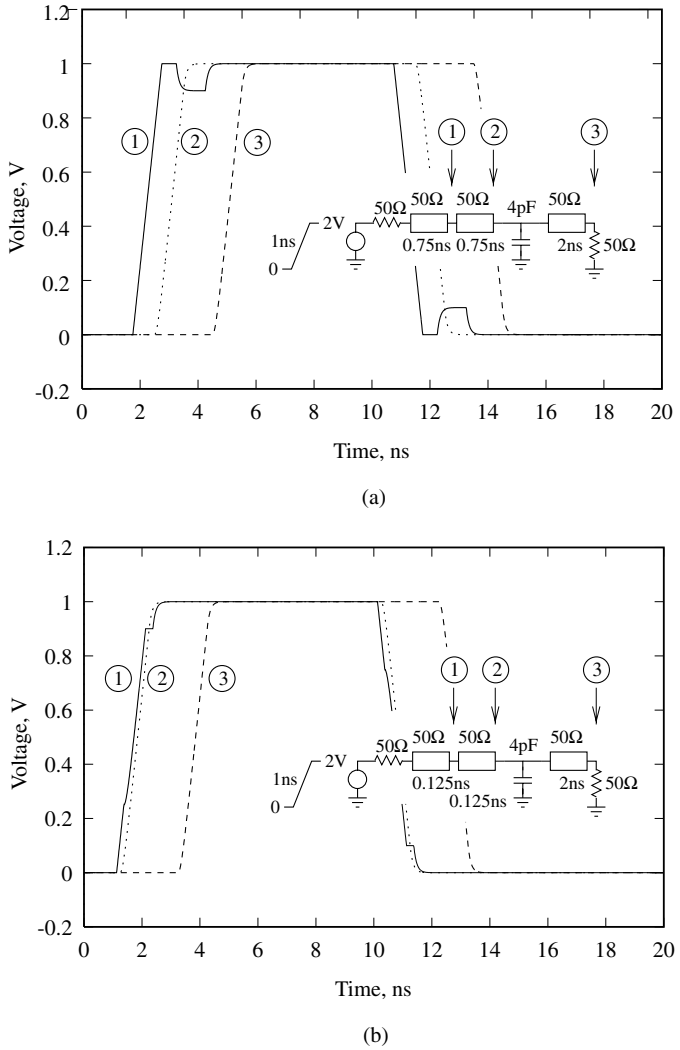


Figure 2.15. Reflected and transmitted waveforms at a shunt capacitance discontinuity. (a) Fast edge produces a reflected pulse with width approximately equal to the edge rate. (b) Slow edge at the source produces a truncated pulse with modified rise time.

reflection coefficients and the generation of narrow spikes with widths similar to the edge rate. In contrast, the characteristic impedance of transmission lines can be essentially constant over very broad bandwidths, so the reflection coefficient can be strongly frequency-independent. In these cases, a fraction of the entire waveform can be peeled off and reflected back down the line.

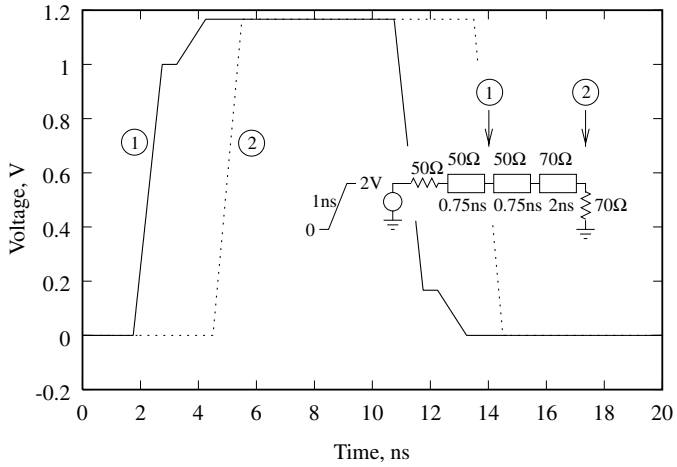
Assume a signal propagates down a transmission line with characteristic impedance Z_1 and meets a discontinuity where the line impedance changes to Z_2 thereafter. When the lines are long or terminated, then (2.40) holds and two cases must be considered. When $Z_1 > Z_2$, the reflection coefficient is negative and the reflected wave is a negative copy of the incident wave with reduced amplitude. When $Z_1 < Z_2$, the reflected wave is a positive copy. The incident and reflected waves superimpose.

Voltage is continuous at the discontinuity, so the signal continues onto the second transmission line with peak amplitude based on the total voltage on the first line. When the incident and reflected waves have the same sign, they add, and the voltage signal on the second transmission line is large. Conversely, a negative reflection coefficient produces a smaller transmitted voltage signal.

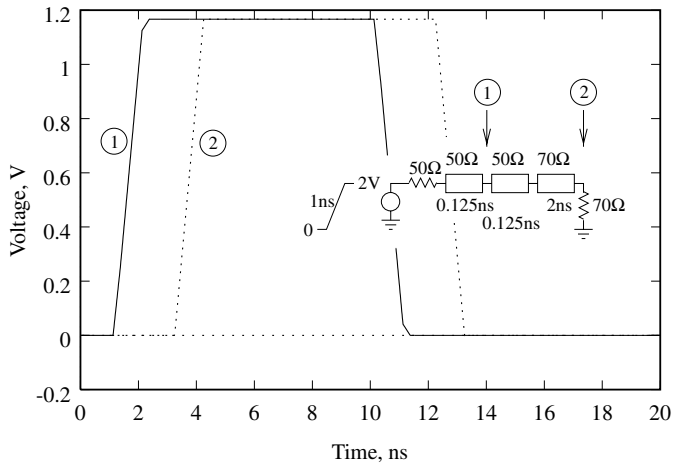
Example

Sample SPICE simulations for both cases are shown in Figures 2.16 and 2.17, where the schematics are shown in inset. See section 2.2.1 for a discussion on fast and slow edges. In Figure 2.16, the reflection coefficient is +0.1667, so the reflected wave has the peak amplitude 1V*0.1667=0.1667V. The incident and reflected waves add up to 1.1667V when they happen to be coincident; otherwise, just the 1V incident wave or 0.1667V reflected wave is observable. The transmitted wave is launched with the peak voltage of 1.1667V, so its slew rate increases because it transitions over a larger voltage in the same time as the incident wave.

In Figure 2.17, the reflection coefficient is -0.1667, so the reflected wave has the peak amplitude -0.1667V, and the transmitted wave has the peak amplitude -0.8333V. The slew rate of the transmitted wave decreases.

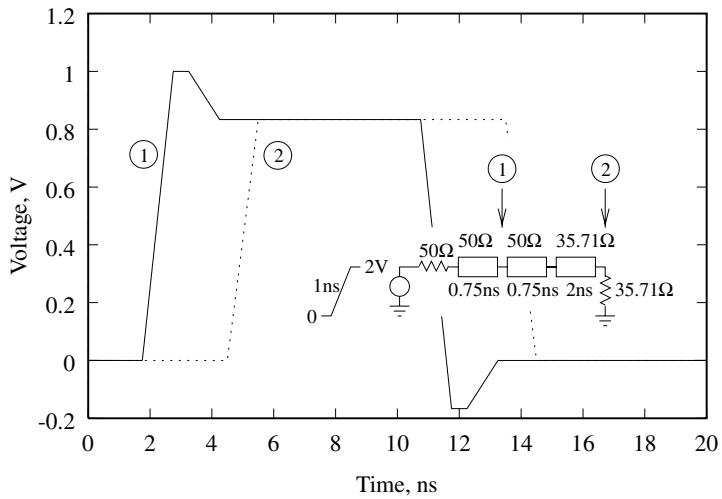


(a)

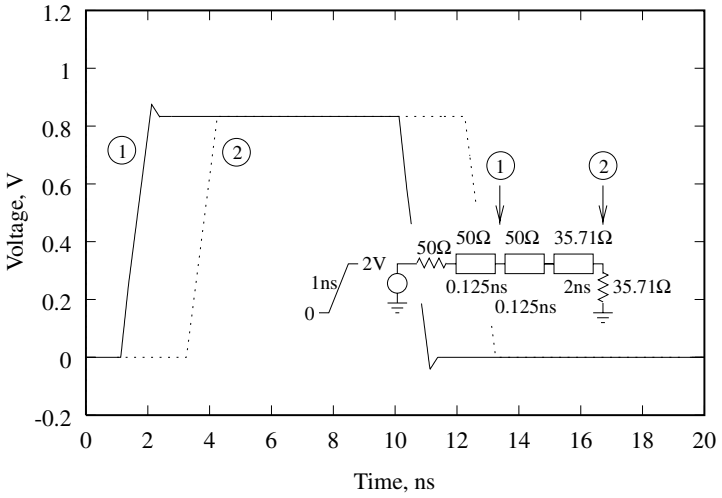


(b)

Figure 2.16. Reflected and transmitted waveforms for an impedance discontinuity to a larger impedance: (a) fast edge, (b) slow edge.



(a)



(b)

Figure 2.17. Reflected and transmitted waveforms for an impedance discontinuity to a smaller impedance: (a) fast edge, (b) slow edge.

2.5 Crosstalk

Leakage of a signal from one conductor to another is called *crosstalk*, and it can be induced through three coupling mechanisms: capacitive, inductive, and radiative. Radiative coupling is essentially a self-induced EMI disturbance and would be treated within an EMI design framework. This section is concerned with capacitive and inductive coupling.

2.5.1 Capacitive Crosstalk

All conductors have some capacitance between them, and when sufficiently close, the capacitance can become large enough to couple significant energy from one line (the aggressor, or active line) to another (the victim, or passive line). The capacitance allows displacement current to cross the gap and inject into the victim line. Since the impedance is equal looking both ways up and down the line, the current splits equally and sends waves propagating in each direction. The coupling is sketched in Figure 2.18, where capacitance is distributed along the length of two transmission lines. The mutual capacitance per-unit-length is c_m. The problem in general requires

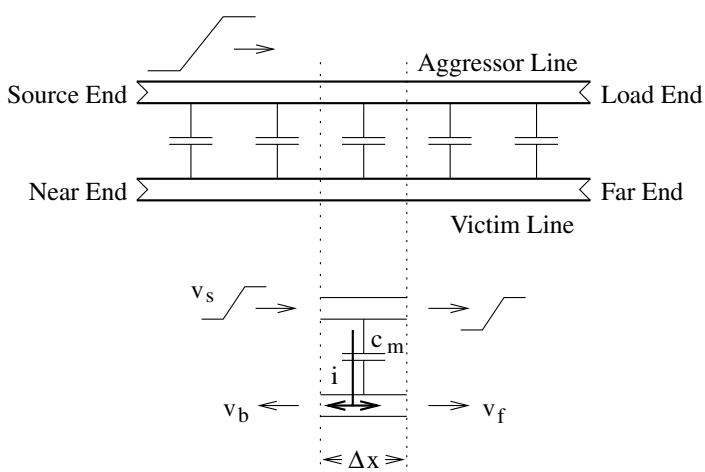


Figure 2.18. Sketch of capacitive coupling producing crosstalk.

the simultaneous solution of the coupled differential equations representing the two lines. Approximations can be applied to show the basic behavior and to derive useful formulas suitable for weak coupling.

A very short length of the coupled transmission lines is shown in cutout in Figure 2.18. As the voltage wave passes on the aggressor line, it injects current onto the victim line through displacement current, then this current splits evenly in each direction, due to symmetry. Kirchhoff's current law applied to the victim line yields

$$\frac{v_b}{Z_o} + \frac{v_f}{Z_o} = c_m \Delta x \frac{dv_s}{dt},$$

where it is assumed that, due to weak coupling, the voltage across the mutual capacitor is v_s, the amplitude of the voltage wave on the aggressor line. The voltage is continuous, so $v_b = v_f$, and

$$v_f = v_b = \frac{1}{2} Z_o c_m \Delta x \frac{dv_s}{dt}. \tag{2.56}$$

The aggressor wave creates pulses that have widths about equal to the edge rate and that propagate in opposite directions on the victim line. A low-to-high transition on the aggressor line produces positive pulses on the victim line, while a high-to-low transition produces negative pulses.

Once a crosstalk signal has been launched on the victim line, it too can create crosstalk back onto the aggressor, where it can upset the waveform there and complicate further computations of crosstalk. When this secondary crosstalk is negligible, then the coupling is said to be weak; otherwise, the coupling is strong. Simple formulas for crosstalk rely on weak coupling.

The aggressor and forward victim waveforms travel together towards the load and the far end, respectively. Therefore, at each increment on the lines, the aggressor edge can add to the victim pulse, and the victim pulse grows in magnitude the farther down the lines the waves travel. At the far end, for line length d, the total noise for weak coupling from (2.56) is

$$v_{FE} = \frac{1}{2} Z_o c_m d \frac{dv_s}{dt}, \tag{2.57}$$

which matches the capacitive portion of the same result in (10.36), derived using coupled mode theory. The noise at the far end—far-end noise (FEN)—is a single pulse with width approximately equal to the edge rate of the signal on the aggressor.

The aggressor and backward victim waveforms travel in opposite directions, so the overlap where the aggressor can inject current is only one-half the rise time, Δt. After this time period, the pulse travels unchanged to the near end. However, these pulses are generated continuously, so the near end receives them for at least 2×TOF, when the last pulse generated at the far end propagates back to the near end.

The interaction distance in (2.56) is $\Delta x = v_p \frac{1}{2} \Delta t$, where v_p is the phase velocity. To simplify the result, assume that the aggressor edge is triangular, then $dv_s/dt = v_o/\Delta t$, where v_o is the peak voltage. Substituting these into (2.56) yields the near-end noise (NEN) for weak coupling as

$$v_{NE} = \frac{1}{4} Z_o c_m v_p v_o.$$

However, $Z_o v_p = \sqrt{\ell/c}/\sqrt{\ell c} = 1/c$, where c is the capacitance per-unit-length of either line, so

$$v_{NE} = \frac{1}{4} \frac{c_m}{c} v_o.$$

Expressing as a coupling coefficient yields

$$K_{NE} = \frac{1}{4} \frac{c_m}{c}, \tag{2.58}$$

where $K_{NE} = v_{NE}/v_o$. This result matches the capacitive portion of the same result in (10.31), derived using coupled mode theory.

In summary, capacitive crosstalk results in a short pulse at the far end and a long signal at the near end. The far-end signal grows in amplitude with longer lines, while near-end noise grows in width with longer lines. The crosstalk noise is positive for low-to-high transitions and negative for high-to-low. The results are summarized in Figure 2.19.

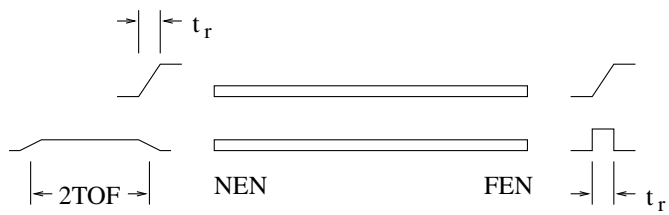


Figure 2.19. Summary of noise waveforms for weak capacitive coupling.

2.5.2 Inductive Crosstalk

The closed loops formed by two signal lines are coupled by mutual inductance, which causes a crosstalk voltage to be generated on the victim line due to changes in the current on the aggressor line according to $v = m\frac{di}{dt}$. In contrast to the capacitive case, where current is injected into the victim line, the net change in current is zero; the aggressor line can only drive current along the victim line. As a result, the forward and backward crosstalk have opposite polarities. The coupling is sketched in Figure 2.20, where mutual inductance, m, is distributed along the length of the line. Similar to the case for capacitive crosstalk, the problem in general requires

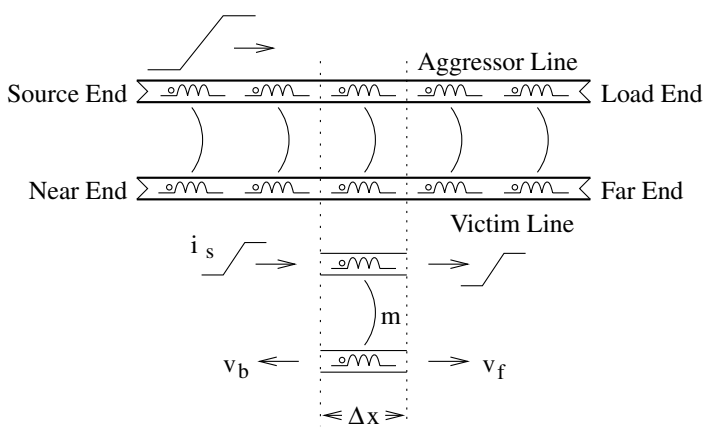


Figure 2.20. Sketch of inductive coupling producing crosstalk.

the simultaneous solution of the coupled differential equation representing the two lines. Approximations can be applied to show the basic behavior and to derive useful formulas suitable for weak coupling.

A very short length of the coupled transmission lines is shown in cutout in Figure 2.20. As the current wave passes on the aggressor line, it induces a series voltage on the victim line through the mutual inductance. Kirchhoff's voltage law applied to the victim line yields

$$v_b = m\Delta x \frac{di_s}{dt} + v_f.$$

The currents are continuous, so $v_b/Z_o = -v_f/Z_o$. Eliminating v_f and substituting $i_s = v_s/Z_o$ yields the backward wave amplitude as

$$v_b = \frac{1}{2}\frac{m}{Z_o}\Delta x \frac{dv_s}{dt},$$

and similarly for the forward wave

$$v_f = -\frac{1}{2}\frac{m}{Z_o}\Delta x \frac{dv_s}{dt}.$$

The aggressor wave creates pulses that have widths about equal to the edge rate and that propagate in opposite directions on the victim line. A low-to-high transition on the aggressor line produces a positive backward pulse and a negative forward pulse on the victim line, while a high-to-low transition reverses the signs.

At this point, the inductive crosstalk derivation follows that of capacitive crosstalk. The forward noise travels with the aggressor wave and picks up amplitude continuously. The backward noise only picks up noise for half the edge rate since it and the aggressor travel in opposite directions. The results for inductive crosstalk are

$$v_{FE} = -\frac{1}{2}\frac{m}{Z_o} d \frac{dv_s}{dt} \tag{2.59}$$

and, assuming a triangular edge on the aggressor waveform,

$$v_{NE} = \frac{1}{4}\frac{m}{\ell} v_o.$$

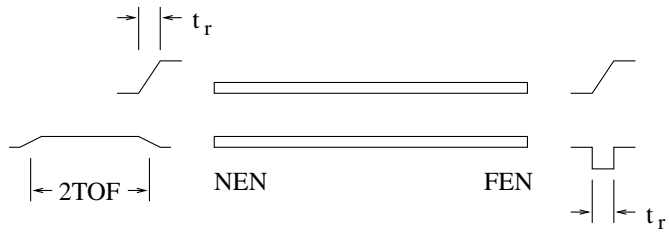


Figure 2.21. Summary of noise waveforms for weak inductive coupling.

Expressing the near-end noise as a coupling coefficient yields

$$K_{NE} = \frac{1}{4}\frac{m}{\ell}. \tag{2.60}$$

Note that these results match (10.36) and (10.31), derived using coupled mode theory.

In summary, inductive crosstalk is very similar to the capacitive crosstalk: a short pulse at the far end and a long signal at the near end. The difference is in the signs, where forward inductive and capacitive crosstalk have opposite signs. The results are summarized in Figure 2.21.

2.5.3 Total Crosstalk

In the general case, capacitive and inductive crosstalk are simultaneously present. Summing (2.57) and (2.59) yields

$$v_{FE} = \frac{1}{2}d\left(Z_o c_m - \frac{m}{Z_o}\right)\frac{dv_s}{dt}, \tag{2.61}$$

which applies for weak coupling. An implicit assumption is that the phase velocities on the two coupled lines are equal. Because of the opposite polarities for capacitive and inductive crosstalk, far-end noise enjoys some cancellation. Far-end noise is eliminated when $Z_o c_m - \frac{m}{Z_o} = 0$, or

$$\frac{c_m}{c} = \frac{m}{\ell},$$

a balanced condition where both coupling terms represent the same fraction of the aggressor line's per-unit-length parameters.

For near-end noise, summing the capacitive crosstalk from (2.58) and the inductive crosstalk from (2.60) yields

$$K_{NE} = \frac{1}{4}\left(\frac{c_m}{c} + \frac{m}{\ell}\right). \tag{2.62}$$

Cancellation is not possible, so near-end crosstalk always exists.

Note that the cancellation of far-end crosstalk enables the creation of directional couplers. With far-end cancellation, a sample of a signal moving left-to-right only appears at the left port. Similarly, a signal moving right-to-left produces a signal only at the right port. Such a directional coupler can be used to create network analyzers, which operate on the principal of directional traveling waves. From the theory, good performance requires equal phase velocities on two lines, plus careful balance of capacitive and inductive crosstalk.

In digital systems, many closely spaced lines exist with a mix of crosstalk components. For capacitive crosstalk, lines shield each other so c_m decreases rapidly for victim lines farther from the aggressor. Therefore, capacitive crosstalk is typically strong to a line's neighbors but weak to other lines. For inductive crosstalk, magnetic fields are not shielded by normal metals, so m decreases slowly and inductive crosstalk is strong for many of a line's neighbors. Summing these two results, the typical crosstalk behavior is that coupling to neighbors is both inductive and capacitive, with capacitive crosstalk often dominating, while coupling to farther lines is mostly inductive.

2.6 Topology

A point-to-point connection between two components yields the highest signal integrity, but most circuits require several components to be connected with a single net. A net connecting three or more pins is a multidrop net. For low-speed operation, just about any topology that completes the connection is acceptable. To

achieve high-speed operation, the layout of the net becomes important. Several topologies for interconnecting components are sketched in Figure 2.22.

The daisy chain simply stitches from one circuit to the next and is suitable for low-speed nets. With no termination (other than any driver source impedance), each drop introduces a shunt capacitance to ground, and the stubs between drops support complex multiple reflections. Daisy chains work well when the delay of the entire length of the net is short compared to the rise time; then the net acts as a capacitive load and the driver can be scaled accordingly.

The near-end cluster or star cluster can be very effective but requires delicate balance to achieve good performance for fast edges. Consider the example in Figure 2.23, where a single stub is connected in the middle of a 10ns line for a short rise-time pulse. As the stub length increases, the ringing period increases and the ringing amplitude takes longer to settle. The voltage dividers, due to the source impedance and the impedance step at the stub, can be seen within the first 12ns on the near-end plot. When the stub reaches the same length as the main line length beyond the branch, the waveform quality improves dramatically, although multiple reflections are still required to reach the full line value. Star clusters require equal stub length and loading.

Just as for the daisy chain, if the edge rate is sufficiently long, then the line lengths in the star cluster are not important. In Figure 2.24, the star cluster of Figure 2.23 is resimulated with a slow 30ns edge, which is long enough for the source edge to propagate to the far-end load and back plus 10ns, which is one round trip TOF on the 5ns far-end section. The deviations from the perfect ramp are primarily isolated to the edge's ramp time with almost no ringing later.

In general, stubs create two principal opportunities for loss of signal quality. First, the branch in the line creates an impedance discontinuity by placing the two lines in parallel. The impedance discontinuity will appear as a simple voltage divider for the shortest round trip TOF to the closest discontinuity. Second, reflections from the unmatched terminations will return to the branching point and reflect again,

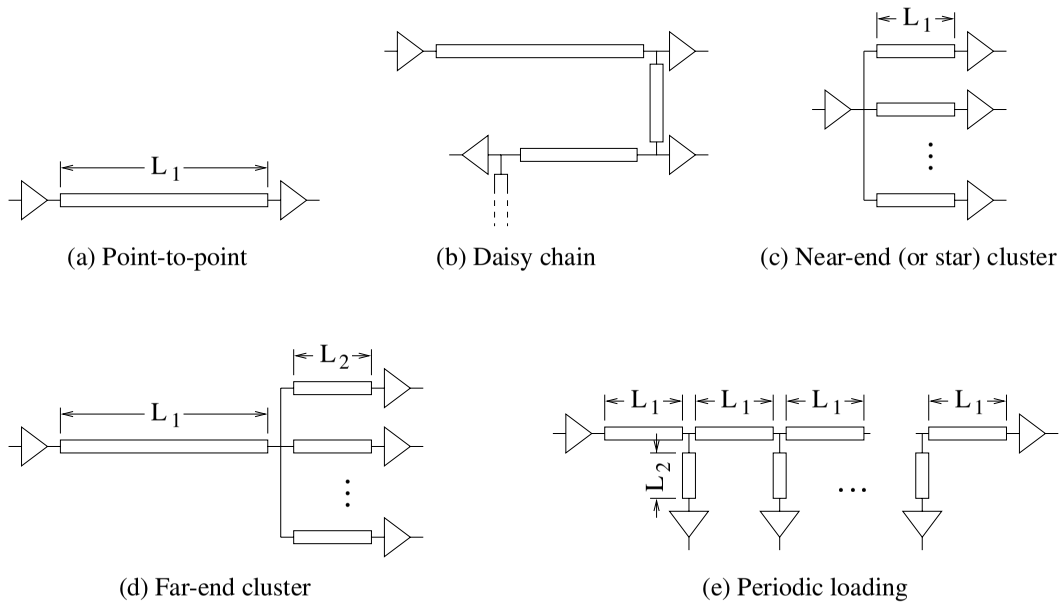


Figure 2.22. A few topologies for connecting components.

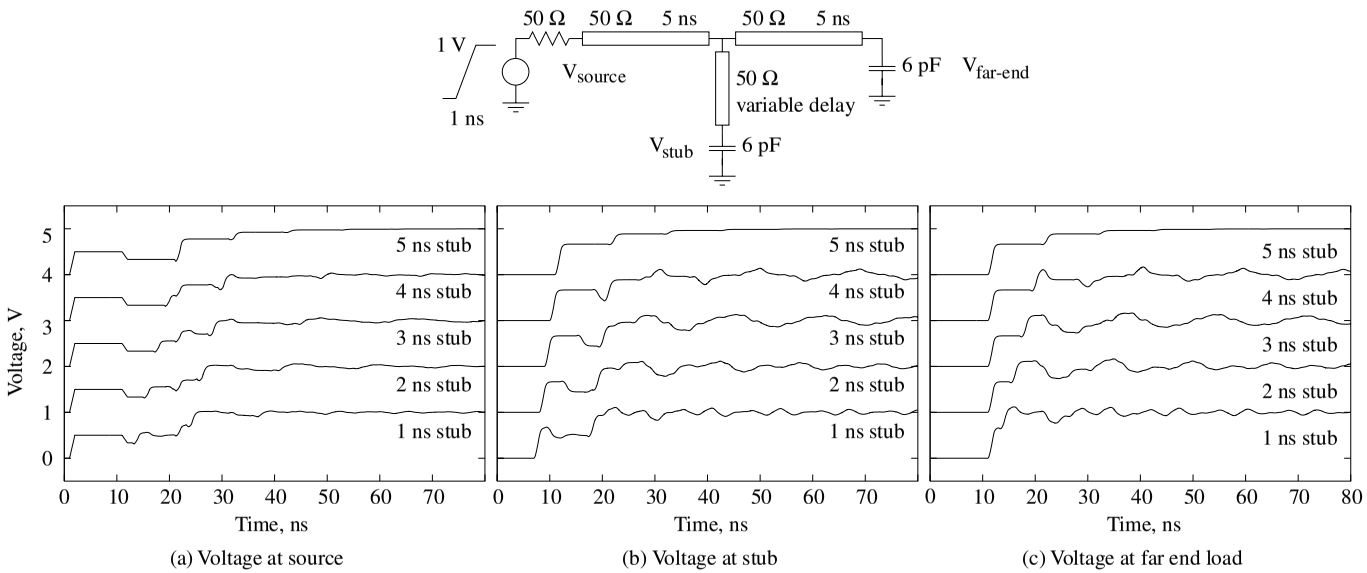
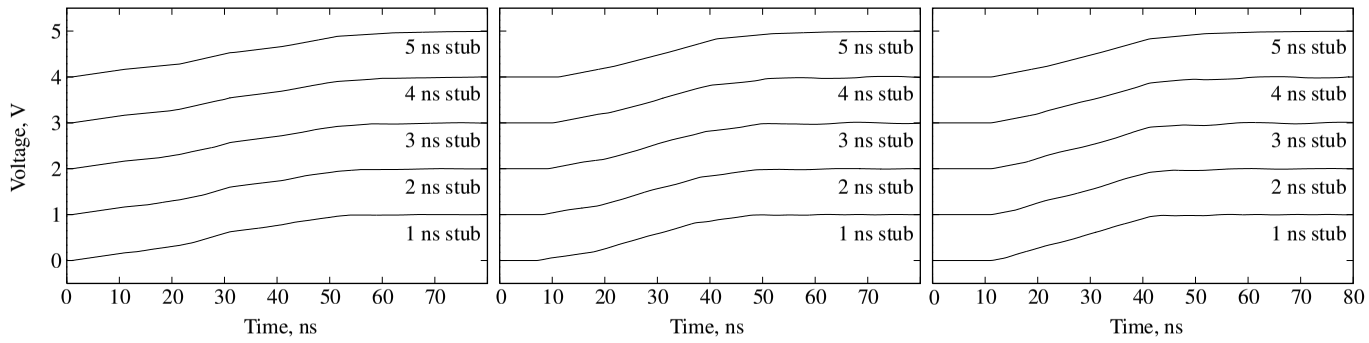


Figure 2.23. Example SPICE simulations of the effects of stubs with a fast 1ns edge. Note the 1V offsets added to separate the waveforms.



(a) Voltage at source. (b) Voltage at stub. (c) Voltage at far end.

Figure 2.24. Example SPICE simulations of the effects of stubs with a slow 30ns edge. Note the 1V offsets added to separate the waveforms.

causing multiple reflections. The sum of all of these interactions can be quite complex and will in general require detailed simulation.

The far-end cluster improves signal integrity by limiting the length of the far-end stubs. By keeping the lengths of stubs short compared to the edge rate, the far-end cluster behaves like a single capacitive load.

Finally, good signal integrity can be achieved by periodically loading a transmission line with additional capacitance. A straight transmission line can be tapped along its length, as shown in Figure 2.22e, to create a multidrop net well suited to creating an expansion bus. Assuming that the stubs and the distance between them are short compared to the signal rise time, then the composite structure acts as an artificial transmission line with characteristic impedance lower than the unloaded trunk transmission line. The inductance is unaffected, but the capacitive loading from the stubs and associated receivers adds to the capacitance of the trunk to pull down the characteristic impedance. If the total capacitive load is C_L, then the distributed loading is C_L/d, where d is the length of the trunk line. The loaded characteristic impedance and phase velocity are then

$$Z_L = \sqrt{\frac{\ell}{c + C_L/d}} = \frac{Z_o}{\sqrt{1 + C_L/(cd)}} = \frac{Z_o}{\sqrt{1 + C_L/C}},$$

and

$$v_L = \frac{1}{\sqrt{\ell(c + C_L/d)}} = \frac{v_p}{\sqrt{1 + C_L/(cd)}} = \frac{v_p}{\sqrt{1 + C_L/C}},$$

where $C = cd$ is the total capacitance of the trunk line, and Z_o and v_p are the unloaded values for the trunk line. Note how the loading capacitance only lowers impedance and phase velocity. To avoid reflections from an impedance step, the loaded impedance must be matched to other system impedances. High-performance expansion bus specifications include limits on stub lengths and capacitive loading per slot to ensure impedance match and signal integrity.

Wiring Rules

In a large system, detailed simulation of every net is too time consuming. Precharacterization of several topologies can provide a library of acceptable layouts, called *wiring rules*, that are guaranteed to produce acceptable signal quality and delay. A system designed using such a library is then correct by construction. Any generic topology, such as one from Figure 2.22, can be considered for characterization for wiring rules.

Detailed simulations over process corners with each variable combination are required, so the effort to construct a usable set of wiring rules can be considerable. Metrics from the simulation may include delay, rise and fall time, overshoot, undershoot, crosstalk, settling time, and SSN. After appropriate metrics and their bounds have been selected, a large number of simulations can be performed to find the range of driver strengths and loadings, characteristic impedances, and line lengths that fit within the bounds. The set of valid values are fitted to equations, and these equations become the wiring rules.

To demonstrate, the simple circuit shown in inset in Figure 2.25 has fixed driver impedance and rise time, load capacitance, and transmission line delay. What range of characteristic impedance is allowable so that the voltage at the load rises to 0.75V within 3ns? Running several cases, it can be seen that impedances above 40Ω are allowed. If a limit for overshoot is also specified, then an upper limit on characteristic impedance can be established.

2.7 Simultaneous Switching Noise

Ideally, every signal would be provided with separate and isolated power and ground connections. For packaged digital components, such dedicated pins drive up packaging costs by increasing pin counts. Common practice is to share several ground and power pins among all of the signal pins, and the performance of the system is strongly related to the ratio of the number of signal pins to the number of ground and/or power pins. For example, an 8:1 signal-to-ground ratio is common, where the

number of power pins equals the number of ground pins. Very high-speed interfaces require ratios as small as 2:1, while very low-performance interfaces may need only one power pin and one ground pin.

Shared ground and power pins enable coupling from one signal line to another. Because the power distribution inductance is small, the coupled noise is small if few drivers are switching at any given time. However, the level of coupled noise increases with the number of switching outputs, and when the whole bus is switching, the noise level can become excessive. For this reason, this noise mechanism is called *simultaneous switching noise* (SSN). It is also known as *delta-I noise* since it results from the rate-of-change of current across the package inductance. SSN is treated in detail in chapter 3.

2.8 System Timing

Prior sections have discussed in detail many mechanisms that cause degradation in digital signaling. If signal integrity is sufficiently reduced, then the performance of

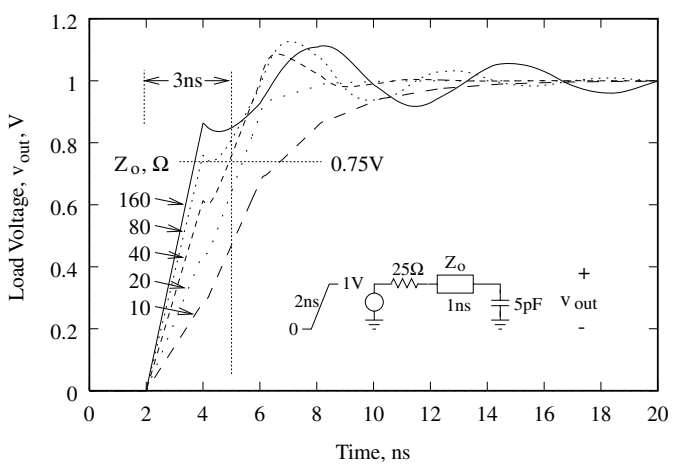


Figure 2.25. Example simulations to generate a wiring rule for allowed transmission line impedance given delay, driver impedance, and load.

the digital system will suffer. The maximum clock rate of a digital system is related to the signaling specifications, driver and receiver performance, and interconnect length. Eye diagrams offer a convenient method to verify that signaling specifications are met while providing estimates of skew.

2.8.1 Maximum Clock Rate

The maximum clock rate of a synchronous system is of primary interest, and the waveforms in Figure 2.26 show the major components that set the upper limit on the clock rate: driver propagation delay (t_p), TOF on the interconnect, and the setup and hold times at the receiver. Note that the driver's rise/fall time is subsumed into the driver's propagation delay; a longer rise time causes the driver's output to reach the receiver's trigger point later, so delay is increased.

Assuming that the data transitions and is sampled on the rising clock edge, then before the rising clock edge, it must be true that

$$P > t_p + \text{TOF} + t_S.$$

Since the clock rate is inversely proportional to frequency, the maximum clock rate is

$$f_{max} < \frac{1}{t_p + \text{TOF} + t_S}. \tag{2.63}$$

After the rising clock edge, it must be true that

$$t_H < t_p + \text{TOF}. \tag{2.64}$$

So the maximum clock rate is given by (2.63) under the condition that (2.64) is true.

Normally, short driver delay is desired to help maximize the system clock frequency. However, for very short and fast interconnects, the TOF is very small and longer t_p may be needed to satisfy the receiver hold time.

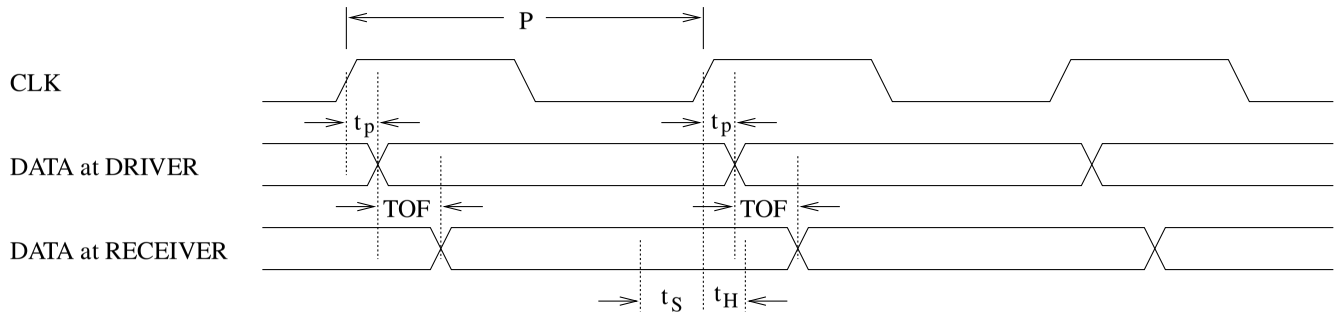


Figure 2.26. Timing diagram for determining maximum clock rate.

2.8.2 Eye Diagrams

Signaling specifications must be met for every signal at every clock cycle. For a long stream of bits, it can be difficult to tell if the signal meets the specifications. To facilitate analyses, the bits can be superimposed to create a single plot from which all of the signaling specifications can be checked.

Consider the clocked data presented in Figure 2.27, where it is assumed that the data is sampled on the rising edge of the clock. The data associated with each rising edge can be cut out and plotted together, as shown for a few of the edges. A good plot results when the waveform is plotted for one clock period before and after the edge, then the full data bit, plus half the one before and half one after, are captured. Superimposing all of the bits then builds an eye diagram. For multiconductor interconnects, the eye diagrams for each line can be superimposed to examine the signal specifications for the whole interconnect.

As an example, consider the schematic in Figure 2.28, where the driver and

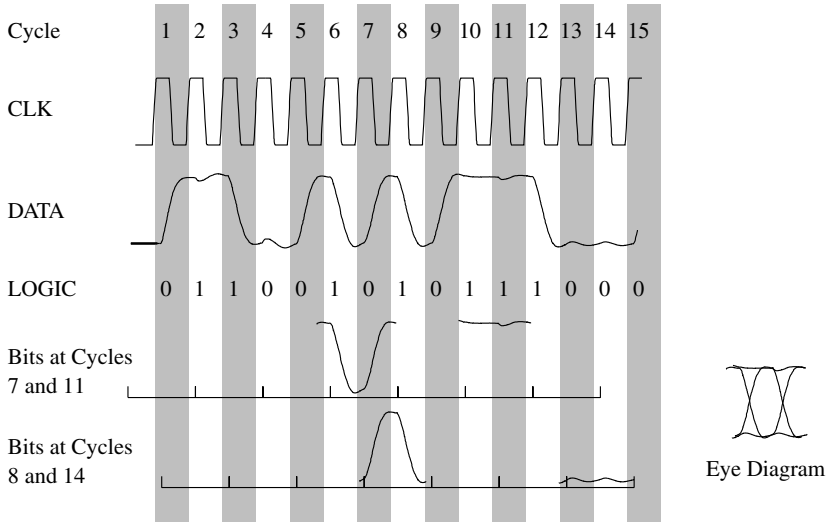


Figure 2.27. Eye diagrams are constructed by overlapping the individual bits in a bit stream.

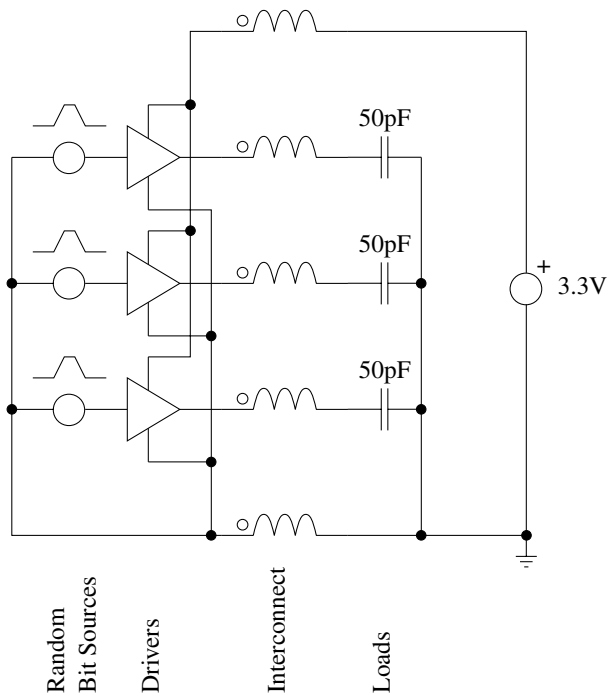


Figure 2.28. Schematic for demonstration of eye diagrams from SPICE simulations.

interconnect models are those appearing in the netlist in Appendix C. Each driver is driven by a random bit stream to simulate a digital signaling sequence. SSN and crosstalk introduced by the package degrade the signal quality at the receiver, modeled here by simple capacitance. In Figure 2.29, the data is driven at a 125MHz clock rate, and the resulting eye diagram for all three signal lines taken together is very clean. As the clock rate is increased to 250MHz in Figure 2.30, the eye diagram begins to show significant impact due to finite edge rate and skew. By 500MHz, shown in Figure 2.31, the eye diagram begins closing down. Note how as the clock frequency rises, the waveforms do not trace each other as well.

One bit of data is often called a *symbol*. In a single stream of data, if the waveform of one bit is not completely settled by the time of the clock transition

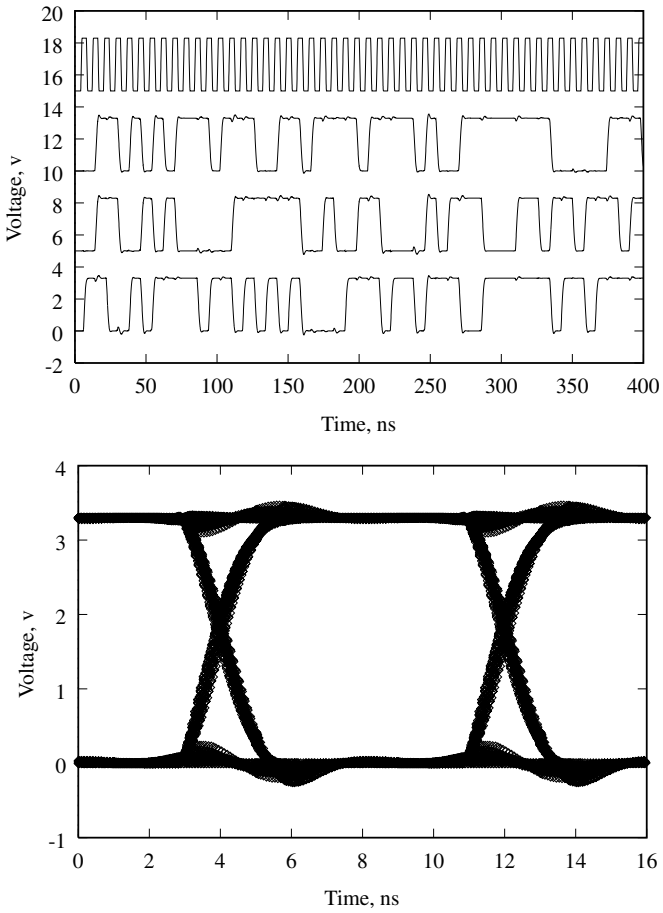


Figure 2.29. Clock and output waveforms (with 5V offsets) and eye diagram of the outputs for the circuit in Figure 2.28 at 125MHz.

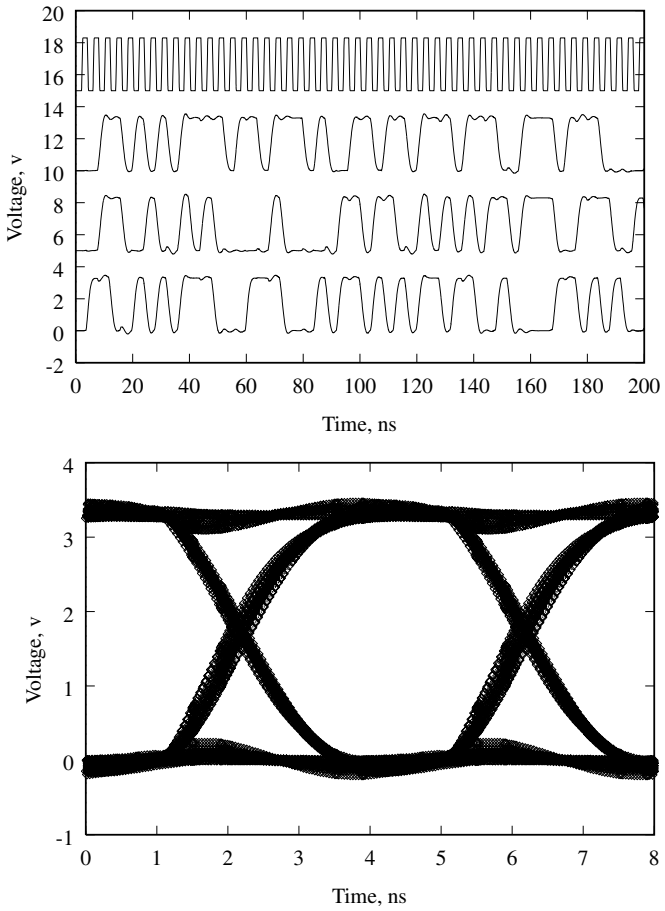


Figure 2.30. Clock and output waveforms (with 5V offsets) and eye diagram of the outputs for the circuit in Figure 2.28 at 250MHz.

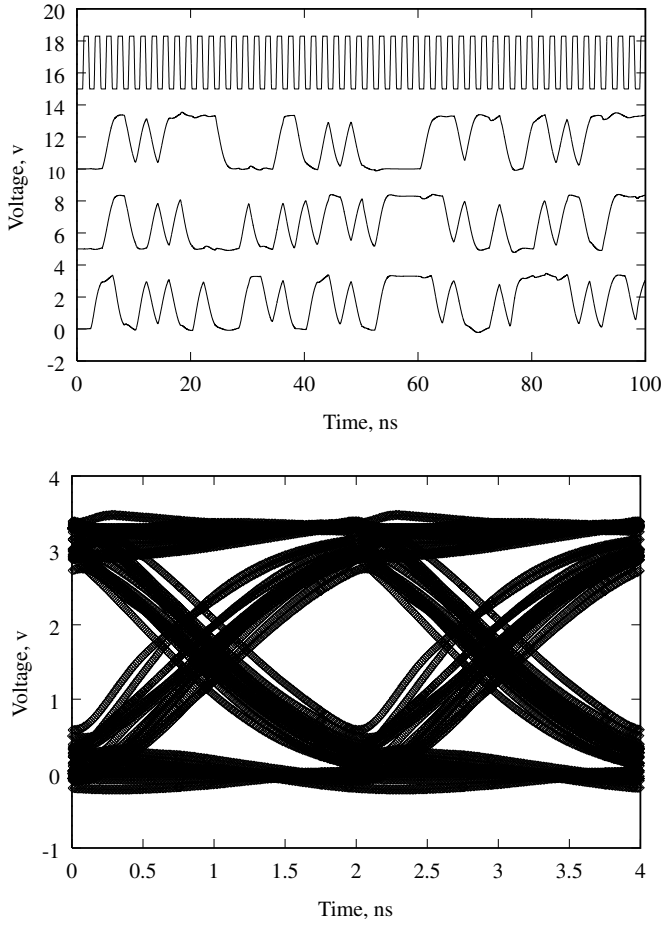


Figure 2.31. Clock and output waveforms (with 5V offsets) and eye diagram of the outputs for the circuit in Figure 2.28 at 500MHz.

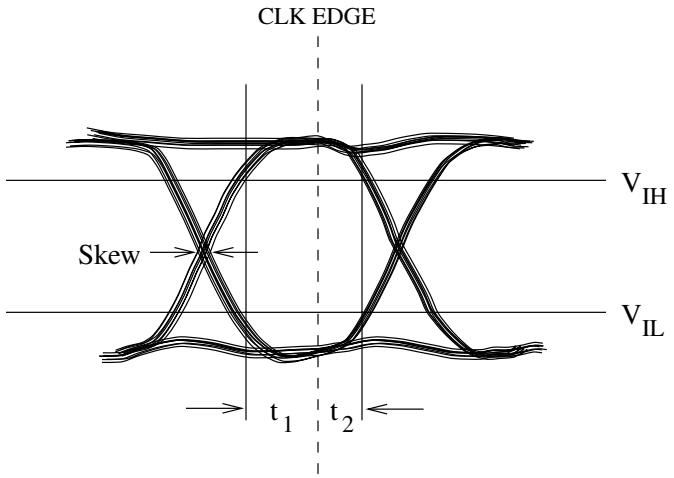


Figure 2.32. Eye diagrams help determine if signals satisfy signaling specifications.

for the next bit, it will affect the shape of the following bit, leading to lower eye quality. The effect is readily apparent in Figure 2.31. When one symbol affects the next, *intersymbol interference* (ISI) is said to occur.

Signal quality only really matters at receivers, where the signal is interpreted according to the logic specifications. Consider the representative eye diagram shown in Figure 2.32. The input specification is superimposed onto the eye diagram. Assuming that the signal is sampled at the clock edge, then the receiver specifications will be satisfied when $t_1 > t_S$ and $t_2 > t_H$, the setup and hold times, respectively.

2.8.3 Skew, Jitter, and Margin

The timing diagram in Figure 2.26 implies that the clock and data edges fall at precise times. As can be seen in the eye diagrams in Figures 2.29 through 2.31, data edges fall along a range of times due to the influences of SSN, crosstalk, TOF variations, and other effects. The uncertainty in the arrival time of a signal edge is the *signal skew*, and all sources of skew must be taken into account to ensure reliable operation of the system. *Clock skew* refers specifically to skew from all sources on

the clock line in a synchronous system.

Clock generators inherently produce some variation in the timing of clock edges at their outputs, and this variation is called *clock jitter*. In a system driven by a single clock generator, all components see the same variation in the timing of the clock edges, so the timing budget is not directly affected by clock jitter. However, jitter does effectively reduce the available clock cycle by occasionally producing one clock edge late and the next one early, and the lower effective clock cycle does affect the timing budget. So clock jitter is included in the timing budget, not because of the uncertainty in the clock edge, but because of the uncertainty in the clock period.

Skews and jitter can be incorporated into the system clock estimates in (2.63) and (2.64), with the precise representation depending on how the skew is broken down. Assuming that the driver propagation delay and the receiver setup and hold times are all worst-case from the manufacturer's data sheet, then with TOF skew, clock skew, clock jitter, and margin, the maximum clock rate is given by

$$f_{max} < \frac{1}{t_p + \mathrm{TOF} + t_S + \Delta t_{\mathrm{TOF}} + \Delta t_{clock} + \Delta t_{jitter} + \Delta t_{margin}}$$

assuming that

$$t_H < t_p + \mathrm{TOF} + \Delta t_{\mathrm{TOF}} + \Delta t_{clock} + \Delta t_{jitter} + \Delta t_{margin}$$

holds true.

If eye diagrams are not available, skew can be estimated from peak noise and the signal slew rate. In-phase noise pushes a waveform higher, so the rising edge arrives at the load earlier. Similarly, out-of-phase noise pushes a waveform lower for later arrival of rising edges. Therefore, additive noise directly causes skew in the timing of the rising and falling edges. The effect of delay caused by noise is often referred to as *pushout*. Assuming that the edges are approximately linear, the slew rate is $\Delta v/\Delta t$. The skew is then approximately the noise voltage times the inverse of the slew rate:

$$\Delta t_{\mathrm{TOF}} = v_{\mathrm{noise}} \frac{\Delta t}{\Delta v}. \tag{2.65}$$

For example, in Figure 2.31, the peak-to-peak noise at the center of the eye is approximately 1V, while the slew rate is 3.3V/2ns. Applying these numbers to (2.65) yields about 0.6ns of skew compared to about 0.7ns from the eye diagram.

Note from (2.65) that a steeper waveform edge induces less skew for a given amount of noise. On the other hand, sharp edges induce noise through additional crosstalk and SSN. Therefore, there is a tradeoff between noise and timing skew. Because digital signaling operates with noise margins, there is no benefit to reducing noise below a certain level. Since lowering the noise through reduced edge rate pays a penalty in timing, an optimal system will utilize the fastest edge rates consistent with an acceptable noise level.

Skew estimates from peak noise voltage can also be used to help set noise budgets. Assuming worst-case additive skew, then the allowed skew from the timing budget can be allocated to noise sources such as crosstalk and SSN to help resolve design options like routing density and package style. Because detailed eye diagram simulations are time-consuming to set up and run, they may be more appropriate in later design confirmation stages.

2.8.4 Dual Data Rate

In some systems, data is sampled on both the rising and falling edges of the clock, and these systems are said to be *dual data rate* (DDR). One of the main advantages of DDR is that the bandwidth of the clock is halved for a given data rate, so DDR is an excellent choice for high-speed interconnects. However, to latch the data with a rising edge, the clock must be delayed by 90° with a delay-locked loop (DLL) or doubled in frequency with a phase-locked loop (PLL), so extra circuitry is required at the receive end. Timing considerations are unchanged for DDR except for the halving of the period plus the addition of jitter for the DLL or PLL.

2.9 Exercises

1. Show that

$$v^+ = \frac{v + Z_o i}{2}$$

and

$$v^- = \frac{v - Z_o i}{2}$$

on a lossless transmission line.

2. Show that the inductance per-unit-length of the parallel-plate transmission line in Figure 2.33 is

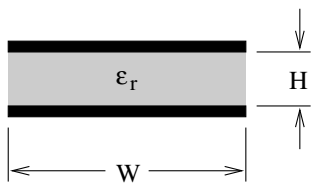


Figure 2.33. Cross section of a parallel plate transmission line.

$$\ell = \mu_o \frac{H}{W}$$

when $W \gg H$. Show that the characteristic impedance is

$$Z_o = \sqrt{\frac{\mu_o}{\epsilon_r \epsilon_o}} \frac{H}{W}.$$

3. (a) Using the notation in Figure 2.34, show that the input impedance of a loaded lossy transmission line is

$$Z_{\text{in}} = Z_o \frac{Z_L + Z_o \tanh(\gamma d)}{Z_o + Z_L \tanh(\gamma d)}. \tag{2.66}$$

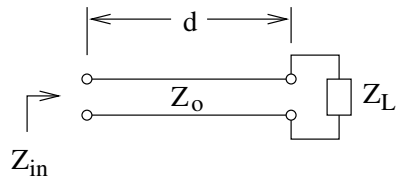


Figure 2.34. Notation for a loaded transmission line.

(b) Show that the input impedance of a loaded lossless transmission line is

$$Z_{\text{in}} = Z_o \frac{Z_L + \jmath Z_o \tan(\beta d)}{Z_o + \jmath Z_L \tan(\beta d)}. \tag{2.67}$$

4. (a) Describe the signal integrity effect of a series capacitor in terms of frequency-dependent impedance and reflection coefficients.

   (b) Set up and run a SPICE simulation to demonstrate the effects of a series 10pF capacitor on the reflected and transmitted signal. Use a 50Ω system, 1ns rise and fall times, 7ns wide pulse, and choose the transmission line lengths to be long compared to the edge rate.

5. Assume that a capacitor is inserted in series into a long transmission line. Set up and solve for the exact reflected, transmitted, and total voltages for an incident unit step, assuming the circuit is initially at rest. Find an approximate solution for small capacitance. Show that the time constant for the discontinuity is $\tau = 2CZ_o$.

6. Repeat problem 4 for a shunt inductance of 25nH. Is this an effective way to block the transmitted signal?

7. Assume that a long transmission line is shorted to ground in the middle with a shunt inductor. Set up and solve for the exact reflected, transmitted, and total voltages for an incident unit step, assuming the circuit is initially at

rest. Find an approximate solution for small inductance. Show that the time constant for the discontinuity is $\tau = 2L/Z_o$.

8. Describe the signal integrity at the load for the circuit in Figure 2.4b as the driver impedance Z_o varies from 1Ω to 2500Ω. Assume a receiver with 6pF of capacitive loading, a 50Ω transmission line with 10ns of delay, and a driver pulse of 1ns rise and fall, 7ns width, and 1V peak.

9. How are forward and backward crosstalk affected when

   (a) The characteristic impedances of the two transmission lines are different, but the phase velocities are the same?

   (b) The characteristic impedances of the two transmission lines are equal, but the phase velocities are different?

10. (a) Find the largest driver impedance that supports first incidence switching on a long 70Ω transmission line with a small capacitive load.

    (b) Repeat for second incidence switching.

    Use a 5V source swing and TTL logic high (2V) receiver threshold with no margin.

11. Discuss the impact of using a driver impedance lower than necessary for first incidence switching on timing, SSN, crosstalk, and power dissipation. (A single SPICE example may be used for the discussion.)

12. Sketch 35ns of the waveform at points 1 and 2 for the schematic in Figure 2.7, modified as follows: source voltage rises from 0 to 3V in 2ns, source impedance is 150Ω, transmission line impedance is 75Ω with 10ns of delay, and the load capacitance is small.

13. In source synchronous systems, the clock is edge-aligned and routed with the data to clock the data into the receivers. This system eliminates clock skew, driver delay, and TOF from the timing budget, enabling higher clock rates.

Derive the setup and hold time requirements for a DDR source synchronous system.

14. Show that far-end skew caused by crosstalk is approximately independent of edge rate. Assume weak coupling and trapezoidal edges.

15. (a) Show that the input impedance of a lossless transmission line shorted at the far end is given by

$$Z_{\text{in}} = \jmath Z_o \tan(\beta d).$$

Draw a sketch of Z_{in} for βd ranging from 0 to 3π and label the frequency bands where the transmission line looks inductive and where it looks capacitive. At what frequencies does this shorted transmission line look like an open circuit?

(b) Show that the input impedance of a lossless transmission line open-circuited at the far end is given by

$$Z_{\text{in}} = -\jmath Z_o \cot(\beta d).$$

Draw a sketch of Z_{in} for βd ranging from 0 to 3π and label the frequency bands where the transmission line looks inductive and where it looks capacitive. At what frequencies does this opened transmission line look like a short circuit?

(c) Repeat for a transmission line terminated in a resistance equal to the characteristic impedance of the line.

16. Find the transmission line equations by applying Kirchhoff's voltage and current laws to the differential transmission line section in Figure 2.35. How does the result compare to (2.3) and (2.4)?

17. Let $\tau = z + \nu t$ and find the direction of travel, velocity, and characteristic impedance of a lossless transmission line for an arbitrary waveform.

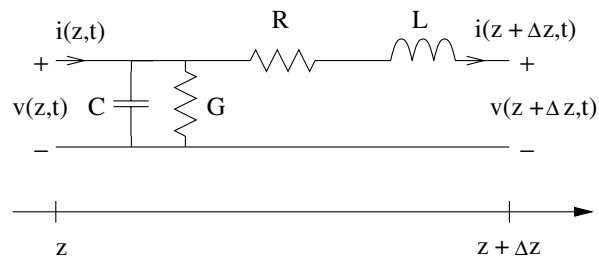


Figure 2.35. Lumped model of a short length of a transmission line. Note the reversal in component order from Figure 2.1.

18. Assume that a voltage source is connected in series at the junction between two lossy transmission lines. Find the reflection and transmission coefficients for the junction. What is the condition for matched impedance ($\Gamma = 0$)?

19. A 1V 50Ω driver with a 1ns edge rate drives a transmission line with $\ell =$ 289nH/m, $c =$ 115pF/m, and $r =$ 20Ω/m loaded at the far end by 5pF. For the lossy and lossless cases, plot the waveforms at the source and load for a low-to-high transition for line lengths of 10cm, 30cm, 90cm, and 270cm. When is the lossless assumption valid?

20. Determine a wiring rule for the variable length line in Figure 2.23 such that the ringing is limited to ±20% overshoot and undershoot.

21. For the star and daisy chain topologies in Figure 2.36,use SPICE simulations to determine which is better in terms of delay, skew, overshoot/undershoot, and power dissipation. Use multicycle simulation and eye diagrams, if available.

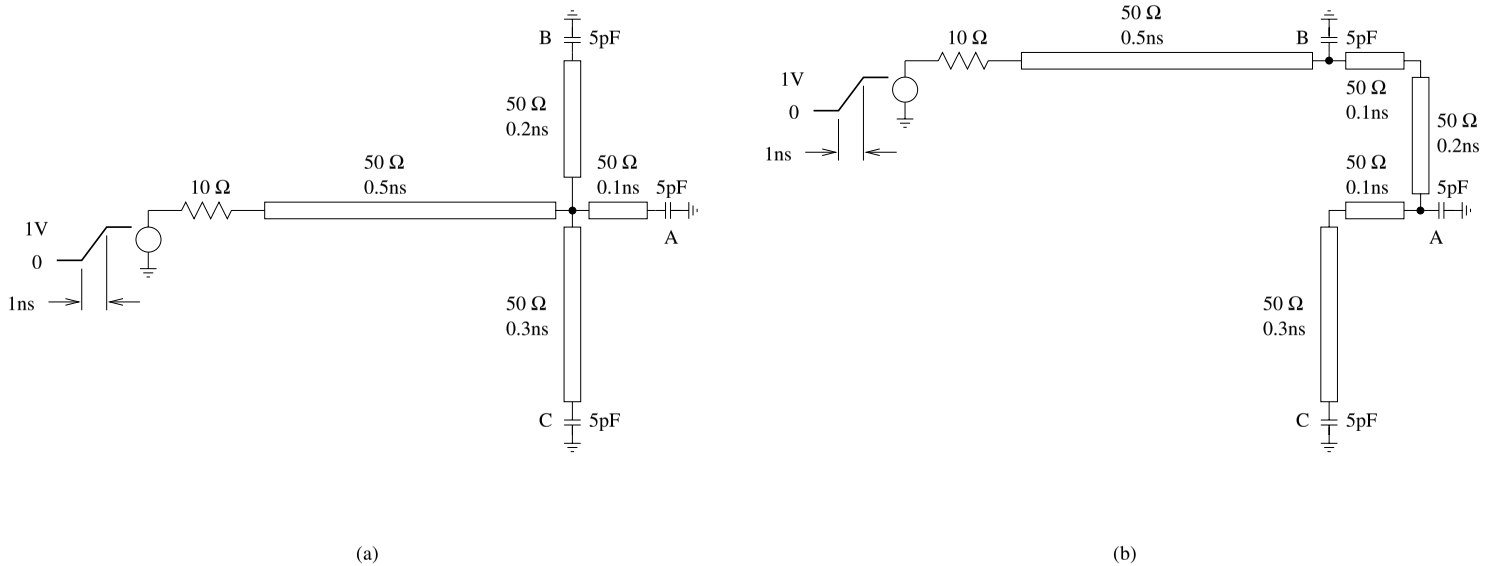


Figure 2.36. Two routing schemes to connect three loads to a single driver: (a) star, (b) daisy chain.