design**feature** *By James Sutherland, California Micro Devices*

EEs MUST LEARN WHEN INTERCONNECTIONS AMONG CIRCUITS BEHAVE AS TRANSMISSION LINES, WHICH CAN UNEXPECTEDLY ALTER SIGNALS. SYSTEMS WHOSE DESIGNS IGNORE THESE EFFECTS PERFORM POORLY OR FAIL ALTOGETHER.

As edge speeds increase, wires become transmission lines

ELECTRONIC ENGINEERS, both digital and analog, must now understand how to live with transmission lines. Behind the need for improved understanding are today's ICs, which produce signals that rise and fall so rapidly that they cause interconnections to behave as transmission lines. This behavior can alter waveforms and timing and even damage components.

Rise and fall times are now as short as 100 psec. When presented with such fast signals, many interconnections behave as transmission lines, even though they don't do so when subjected to signals that change more slowly. Transmission-line effects can arise even if a system's clock frequency remains moderate. By understanding transmission lines' analog characteristics, designers can prevent interconnections from limiting the speed and reliability of their systems.

This article explains what transmission lines are and how they function. It includes the basics on how proper terminations allow you to live with transmission lines.

You should care about the speed of system interconnections because electronic signals propagate at a significant fraction of the speed of light. It takes about 2 nsec for a signal to traverse 12 in. (30 cm) of interconnection on a typical pc board. A few years ago, when gate delays were 5 to 10 nsec, the propagation time of a signal through an IC might have been tens to hundreds of nanoseconds. So, a fraction of a nanosecond or even 2 nsec was not a significant portion of the total signal-propagation time.

In the last decade, however, logic-gate delays have decreased to only a fraction of a nanosecond (50 psec in some 0.25-μm, 3.3V ASICs). Rise and fall times are less than 500 psec. Moreover, although many more functions now fit within the typical IC, the typical pc board now performs so many additional functions that the board size hasn't changed appreciably. Therefore, the time that signals spend getting from place to place on pc boards has become a significant factor in system speed.

How can interconnections limit the speed and integrity of high-speed systems?

Improperly terminated transmission lines can cause the following problems:

Ringing delays cause lower system speeds

A high-speed signal transition on an interconnection that behaves as an improperly terminated transmission line can generate reflections. Reflections cause temporary ringing (voltage oscillations above and below the eventual steady-state level). One way to avoid the consequences of ringing is to wait for the reflections to subside before allowing the system to process new data. To achieve the extra delay, you can add one or more clock cycles to each operation, or you can reduce the system's clock frequency.

Overshoot exceeds maximum-IC-voltage ratings

A reflection can cause a voltage to rise above or, because of negative overshoot, fall below the minimum rating an IC's maximum rating . When the excessive voltage persists for more than a trivial amount of time, the IC is overstressed and may latch up. Latch-up can cause temporary or permanent

Figure 1
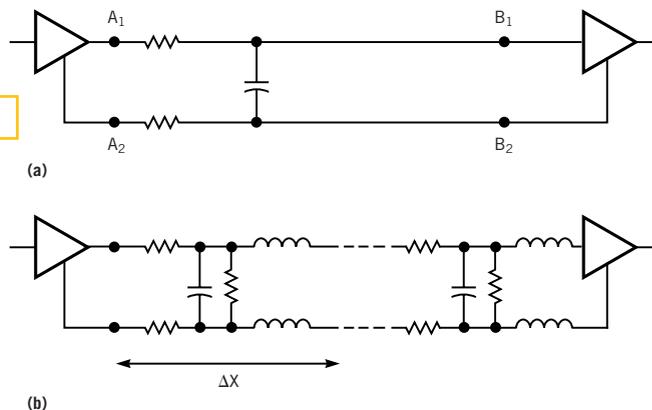
(a)

(b)

A short interconnection or one that operates at a low frequency behaves as a group of lumped circuit elements, such as one capacitor and several resistors (a). If the interconnection is long or the signal frequency is high, the interconnection behaves as a transmission line, which has decidedly different characteristics (b).

damage. (Latch-up is the result of a parasitic npn transistor that causes a CMOS circuit to cease functioning or even to destroy itself.)

Crosstalk increases bit-error rate

Reflections on an improperly terminated transmission line cause larger voltages and currents, which radiate larger electric and magnetic fields and transfer more crosstalk energy into neighboring wires.

Undershoot increases bit-error rate

A signal that exhibits ringing retreats (undershoots or rings back) from its initial maximum-high or minimum-low level. If the signal retreats too far, the receiving IC can read it as the wrong value. If the signal is on a clock line, false triggering can result.

Reduced noise margins increases bit-error rate

In certain configurations, improperly terminated lines cause half- or less-than-full-amplitude signal levels to exist during transition periods. A small noise pulse can carry a signal that is at less than the final amplitude level across the threshold, where it can appear as false data or can cause false triggering.

Driver overload decreases signal integrity, speed, and component integrity

Sometimes, a designer can think that a wire, which is really a 50Ω transmission line, is merely a low-frequency connection to a high-impedance load (say, 10 kΩ). The low line impedance can overload the driver, and, as a result, the load voltage may never reach the proper value.
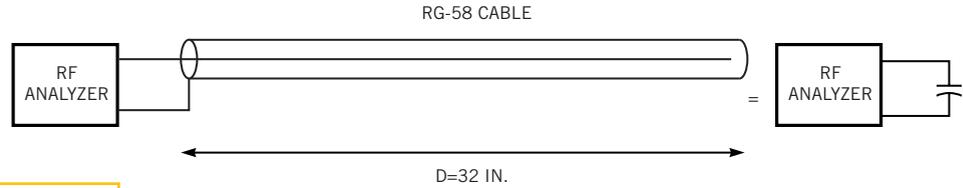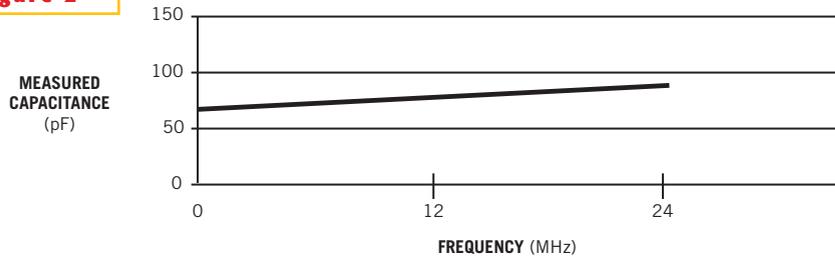
RG-58 CABLE

RF ANALYZER

=

RF ANALYZER

D=32 IN.

If you use a network analyzer to measure the impedance of a 32-in.-long cable, the cable appears to be a simple lumped 70-pF capacitor. (A short length is one whose interconnection length is a small fraction of the signal's wavelength; that is, less than 0.1λ).

Why is signal integrity poor if an interconnection is unterminated?

Under certain conditions, an interconnection ceases to act as a simple pair of wires and behaves as a transmission line, which has different characteristics. The term "wires" includes all types of conductors: traces on pc boards, twisted pairs, ribbon cables, and coaxial cables.

What differentiates a pair of wires from a transmission line?

The length of the interconnection and the highest frequency signal component are the determining factors.

A short interconnection or one that operates at a low frequency behaves as a group of lumped circuit elements, such as one capacitor and several resistors (**Figure 1a**). A typical pc board's interconnect resistance is usually insignificant, so, when you calculate the interconnection delay, you need to consider only the capacitance.

If the interconnection is long or the signal frequency is high, the interconnection behaves as a transmission line, which has decidedly different characteristics (**Figure 1b**). For example, at 1 Hz, a circuit that drives a 100-kΩ load through a long cable sees the cable merely as a 1Ω resistor in series with the 100-kΩ load. However, at 300 MHz, the driver sees only the cable's characteristic impedance, which consists of a continuous uniform resistance, inductance, and capacitance. The driver does not see the 100-kΩ load during the 300-MHz signal's transitions.

This article attempts to answer the following questions:
- What are short and long interconnects?
- What are low- and high-frequency signals?
- What is a transmission line?

TABLE 1—RELATIVE PERMITTIVITY AND PERMEABILITY OF SOME COMMON MATERIALS

Material	Dielectric constant	Permeability	Comments
Air	$\epsilon_R = 1$	$\mu_R = 1$	a great transmission material
Cable (RG-58)	$\epsilon_R = 2.3$	$\mu_R = 1$	
PCB (FR-4)	$\epsilon_R = 4$	$\mu_R = 1$	
Glass	$\epsilon_R = 6$	$\mu_R = 1$	
Ceramic	$\epsilon_R = 10$	$\mu_R = 1$	
Barium titanate	$\epsilon_R = 1200$	$\mu_R = 1$	a great capacitor material; $C = \epsilon A/d$, where A = area, and d = thickness
Steel	$\epsilon_R =$ Not applicable	$\mu_R = 2000$	
Permalloy	$\epsilon_R =$ Not applicable	$\mu_R = 80,000$	(a great magnetic material)

Notes: $\epsilon_R = e/e_0$; $\mu_R = \mu/\mu_0$; and velocity $v = 1/\sqrt{e\mu} = v_0/\sqrt{e_R\mu_R}$.

- How does a transmission line differ from a simple interconnection?
- How can a transmission line hurt the signal's integrity?
- How can proper terminations fix signal-integrity problems?

Short interconnections

A short interconnection is one whose length is a small fraction of the signal's sinusoidal wavelength. A conservative rule of thumb for a small fraction is one-tenth. For example, you can treat a 32-in.-long (81-cm) interconnect as a lumped capacitor at signal frequencies as high as 24 MHz. You arrive at that frequency because 320 in. (8.2m) is the wavelength for which 32 in. is one-tenth. A signal with a wavelength of 8.2m has a frequency of 24 MHz. You can derive this frequency from:

distance=velocity·time
$\qquad$ =velocity·(1/frequency);
velocity of the signal in a cable
$\qquad$ =$(300 \cdot 10^6 \text{m/sec})/(\sqrt{2.3}$
$\qquad$ =$198 \cdot 10^6$m/sec.

So, frequency=$(198 \cdot 10^6$m/sec)/8.2m =24 MHz. If you use a network analyzer to measure the impedance of a 32-in.-long cable, the cable appears to be a simple lumped 70-pF capacitor (**Figure 2**).

Wavelength references

The following are wavelengths for some reference frequencies:

$\lambda = 300 \times 10^6$m/sec/60 Hz
$\quad$ =5000 km (3100 miles).
$\lambda = 300 \times 10^6$m/sec/6 kHz=50 km.
$\lambda = 300 \times 10^6$m/sec/300 MHz=1m.
$\lambda = 300 \times 10^6$m/sec/3 GHz=10 cm.

Wavelength=λ=the distance an electromagnetic wave travels to change phase by 360°.

This relationship is the simple distance=velocity×time formula $\lambda = c_0 \cdot T$, where c_0=the speed of light, T=the period of the signal, and f=the frequency of the signal. $\lambda = c_0 \cdot 1/f$.

These values are for electromagnetic waves in free space. A subsequent discussion covers the effects of a dielectric material adjacent to the conductor. The effect of the dielectric is the reason for the $\sqrt{2.3}$ factor that appears in the earlier discussion of short interconnections.
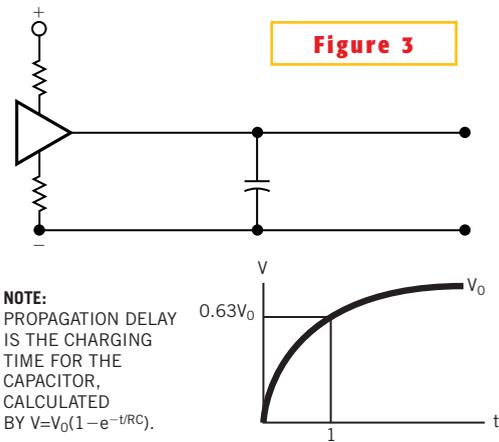
A short-length interconnect looks like a pair of wires and a simple capacitor.

The RC time-constant formula for the charging of a capacitor through a resistor determines the signal delay through a short or low-frequency interconnection (**Figure 3**). The resistance of a conductor on a pc board is usually too small to consider, but you should include the driver's output resistance in calculating the time constant. For example, with a driver that has $R_O = 50\Omega$, a pc-board-trace capacitance of C=50 pF, and a fast 5V input pulse, the signal would be at 3.15V in one time constant (t=RC). Also, at t=2.5 nsec, V=3.15V (63% of 5V); at t=5.0 nsec, V=4.30V (86% of 5V); and at t=7.5 nsec, V=4.75V (95% of 5V).

The interconnection's propagation delay would be the time from the application of the new signal until the signal reaches the receiver's switching threshold.

Does the copper conductor determine the maximum signal speed on a pc-board interconnect?

A discussion such as this one, which focuses on low-frequency phenomena, should consider the speed of the current flow through the bulk-copper conductor. This current flow represents the flow of

NOTE: PROPAGATION DELAY IS THE CHARGING TIME FOR THE CAPACITOR, CALCULATED BY $V = V_0(1 - e^{-t/RC})$.

A short or low-frequency interconnect looks like a pair of conductors and a lumped capacitance. The resistance of a conductor on a pc board is usually too small to consider, but you should include the driver's output resistance in calculating the time constant.

electron (and "hole") charges that occurs when you apply an electric potential to a conductor. The formula in **Figure 4** shows the speed of this charge flow. At extremely low frequencies, there is time for the charge-based current to flow through the bulk of the conductor. However, the speed of approximately 6 cm/sec is wholly inadequate for a high-speed electronic system. You should now be able to appreciate why engineering textbooks call this charge flow "drift current."

So, how do the high-speed signals travel on a pc board?

High-speed signals travel on a pc board as electromagnetic waves at about half the speed of light (15 cm/nsec, or 6 in./nsec) on the surface of the conductors and in the pc-board dielectric. Light, which is also an electromagnetic wave, travels in free space at 30 cm/nsec ($C_0 = 300 \cdot 10^6$m/sec, or 186 miles/msec).

What is a long interconnection?

As you have probably guessed, whereas a short interconnection is less than 0.1 of a signal's wavelength, a long interconnection is longer than 0.1 of the signal's wavelength. The basic guideline is that you must accord transmission-line status to any interconnection whose length is a significant portion of the signal's wavelength. You should recognize, however, that with pulse waveforms, the rise time is easier to use than the wavelength is for determining whether an interconnection is long or short.

TABLE 2–ELECTROMAGNETIC-WAVE SPEEDS

Material	ϵ_R at 10 kHz	Square root of ϵ_R	Speed (in./nsec)	Speed (cm/nsec)
Air	1	1	12	30.5
Teflon	2	1.41	8.5	21.6
Silica glass	3.8	1.95	6.2	15.7
Epoxy glass	5	2.2	5.4	13.7
Ceramic	10	3.2	3.75	9.5

$$\text{SPEED} = \frac{C_0}{\sqrt{\epsilon_R}} = \frac{\text{SPEED OF LIGHT} = 300 \times 10^6 \text{ m/SEC (OR 1 FT/nSEC)}}{\sqrt{\text{RELATIVE DIELECTRIC CONSTANT}}}.$$

Interconnection resonance

When the interconnect length approaches one-fourth of the signal's wavelength, the interconnect starts to resonate and no longer behaves as a lumped capacitor. For example, the 32-in.-long cable in **Figure 5** is one-fourth of a 60-MHz signal's wavelength.

$d=vt=v/f$, and $f=v/d=(3\cdot10^8\text{m/sec}/\sqrt{2.3})/(4\times32\text{ in.}\times1\text{m}/39\text{ in.})=60$ MHz. So, this resonance phenomenon is a reason for not having an interconnection length close to one-fourth the wavelength.

Component size versus wavelength

Figure 6 shows 100-MHz, 1-GHz, and 10- GHz current-signal frequencies. At 100 MHz, the current in the 2-cm-long resistor is the same throughout the resistor. At 1 GHz, the current in the resistor differs slightly at each end. You can't even define the current in the lumped-component resistor at 10 GHz, however: At one end, the current is positive; at the other end, it is negative. Conventional components' distributed nature is one reason to use small thin-film resistors at such high frequencies.

What is a high-frequency interconnection?

A high-frequency interconnection is one in which the signal's rise time (t_R) is less than twice the propagation time (t_{PD}), or time of flight, for the signal's electromagnetic wave to reach the end of the interconnect.

Figure 4

I (CURRENT)

AREA=XY.
LENGTH=Z=VELOCITY×TIME.
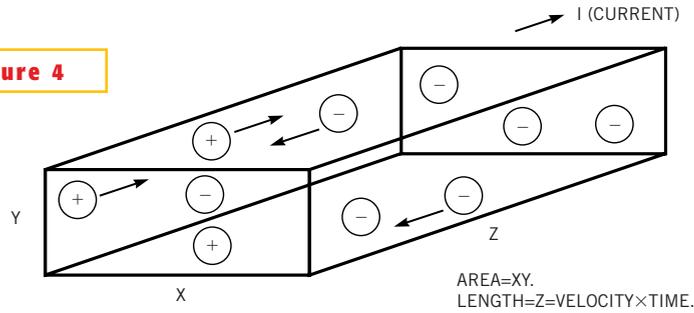
CHARGE IN WIRE=Q=(VOLUME×NO. OF ELECTRON CHARGES(n)/VOLUME)×(CHARGE/ELECTRON(q)).
Q=(AREA×VELOCITY×TIME)×n×q.
I=Q/TIME.
I=AREA×VELOCITY×n×q.

VELOCITY=I/Anq=
J/nq=((10⁵COULOMBS/SEC)/CM²)/((10²³CHARGES/CM³)×1.6×10⁻¹⁹C).

VELOCITY OF CHARGES IN COPPER=6.25 CM/SEC=6.25×10⁻⁹ CM/nSEC.
VELOCITY OF ELECTROMAGNETIC WAVE ON COPPER AND DIELECTRIC OF PC BOARD=15 CM/nSEC.

At extremely low frequencies, charge-based current has time to flow through the bulk of a conductor. However, the speed of approximately 6 cm/sec is wholly inadequate for a high-speed electronic system.

Although the wavelength criterion illustrates a basic principle, this rise-time rule is easier to use and should form the basis for deciding whether an interconnection behaves as a transmission line. The reason is simple: All digital signals contain frequency components that are higher than the fundamental. To have sharp corners, square waves contain at least the third and fifth harmonics. Although you could use the frequency of these harmonics, most digital-system designers think in rise and fall times, which are easier to measure with an oscilloscope. If you can't measure the rise and fall times, you can estimate them. Typically, each time is 10 to 20% of the period.

Designers should not fall into the trap of thinking that their ICs' rise times are those on that data sheets. Most data sheets don't even specify rise and fall times, but the data sheets that do generally provide only the maximum (slowest) values. The parts that you receive are likely to be at least slightly faster and may be several times faster. So, you should use a rise-time figure that you base on a manufacturer's guarantee or on your own tests.

For the earlier example of a 32-in.-long cable, you must treat the cable as a transmission line for signals at frequencies greater than 24 MHz. The following formula illustrates this point:

$$t_{PD} = \frac{\text{DISTANCE}}{\text{VELOCITY}} = \frac{D}{C_O/\sqrt{\varepsilon_R}} =$$

$$32\text{ IN.}\times\frac{\sqrt{2.3}}{300\times10^6\,\text{m}/\text{SEC}\times39\,\text{IN./m}} =$$

$$\frac{48.5}{11.7\times10^9}.$$

where t_{PD} equals 4.15 nsec, which is the time it takes for the electromagnetic wave to make a one-way trip. The $2t_{PD}$ figure equals 8.3 nsec, which is the time it takes for a two-way trip. So, if the rise
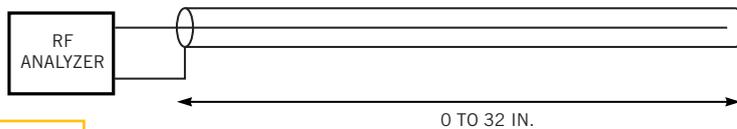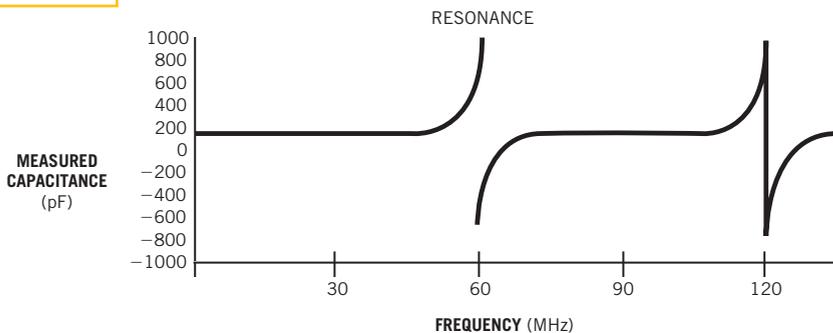
Figure 5

RF ANALYZER

0 TO 32 IN.

RESONANCE

MEASURED CAPACITANCE (pF)

FREQUENCY (MHz)

A long interconnect is one whose length is a significant percentage of the signal's wavelength (that is, greater than 0.1λ). When the interconnect length approaches one-fourth of the signal's wavelength, the interconnect starts to resonate and no longer behaves as a lumped capacitor.

time is less than 8.3 nsec, this cable is a transmission line. If you assume the rise time to be 20% of the signal's period (0.2T=8.3 nsec), then T=41.5 nsec, and frequency=24 MHz. So, for signals with frequencies higher than 24 MHz, this cable is a transmission line, in which ϵ is the relative dielectric constant (2.3 for RG-58 cable) and C_O is the speed of light (300×10^6 m/sec).

To save calculation time, you can use a graph that shows the boundary between wire pairs, which don't need terminations, and transmission lines, which do (**Figure 7**). The data is based on interconnection lengths on a pc board and the signal's rise time.

WAVEFORMS ON INTERCONNECTIONS

Figure 8 shows signal propagation along simple wires and along transmission lines. The **figure** shows signal amplitude versus distance along the line at specific times. With short interconnections or low-frequency signals, the applied signal voltage appears instantaneously at every point on the interconnection (**Figure 8a**). (Propagation is instantaneous if, as it usually is on pc boards and short cables, the cable's RC time constant is smaller than the signal's period.)

Actually, the signal's electromagnetic wave travels at a finite speed along the interconnect from the source to the load (and back if there is a reflection). However, this movement is unnoticeable. The wave travels so often from the source to the load and back after every amplitude change that the transients are too small to observe.

Figure 6
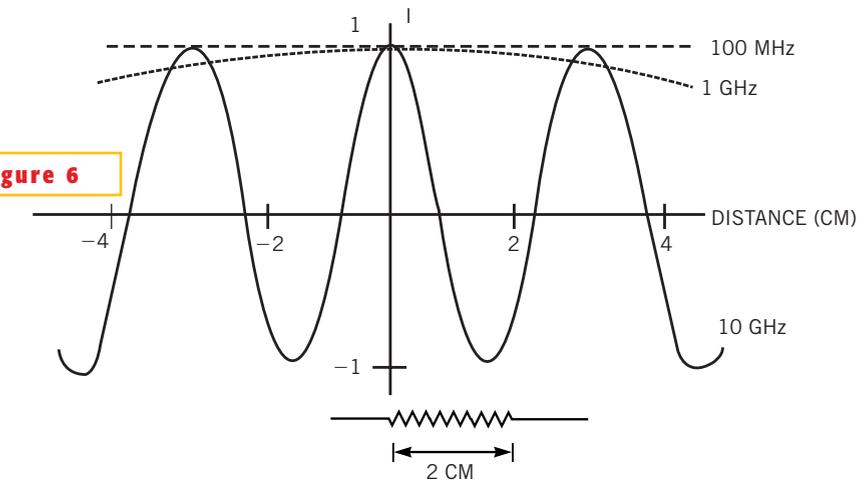
You can't even treat resistors as lumped components at high frequencies because the current density varies widely within their bulk. At 100 MHz, a 2-cm-long resistor has the same current throughout the resistor. At 1 GHz, the resistor has slightly different current at each end. You can't even define the current in the lumped-component resistor at 10 GHz, however: At one end, the current is positive; at the other end, it is negative.

However, with long interconnections or high-frequency signals, the electromagnetic-wave propagation time is comparable with the rise time (**Figure 8b**). Hence, the transient incident and reflected-wave amplitudes are comparable in size to the applied signal. So, by t_1 in **Figure 8**, the applied signal voltage has traveled only a short distance from the source end. Only after the passage of time, t_2, does the voltage appear at a point closer to the load.

A signal propagates down a transmission line as an electromagnetic wave at a velocity close to the speed of light. In free space (in an antenna, for example), where there is no dielectric (insulating) material next to the conductor, the sig-

nal's electromagnetic wave travels at exactly the speed of light. On a pc board, the electromagnetic wave's speed is lower because of the dielectric material adjacent to the interconnect conductors.

The electromagnetic wave's speed on the pc-board interconnections equals the speed of light divided by the square root of the relative dielectric constant of the material adjacent to the conductors. For example, on a board made of FR-4 material, the relative dielectric constant (ϵ_R)=4, and $\sqrt{4}$=2. So, the signal speed on the board is half the speed of light ($1.5 \cdot 10^8$ m/sec or 6 in./nsec).

WAVE PROPAGATION

The concept of waves is one of the great unifying concepts of physics. The physical environment has many types of waves: light waves, sound waves, waves on water, heat waves, radio and TV waves, seismic waves, and even traffic waves. Moving waves carry energy. Most waves travel through media, such as earth, air, water, and steel, without actually carrying the substance with them, though they need the substance to propagate.

Water waves exemplify wave propagation that requires an intermediate material. When you drop a pebble into still water, the water particles near the pebble immediately move from their equilibrium positions. The motion of these particles disturbs adjacent particles,

Figure 7
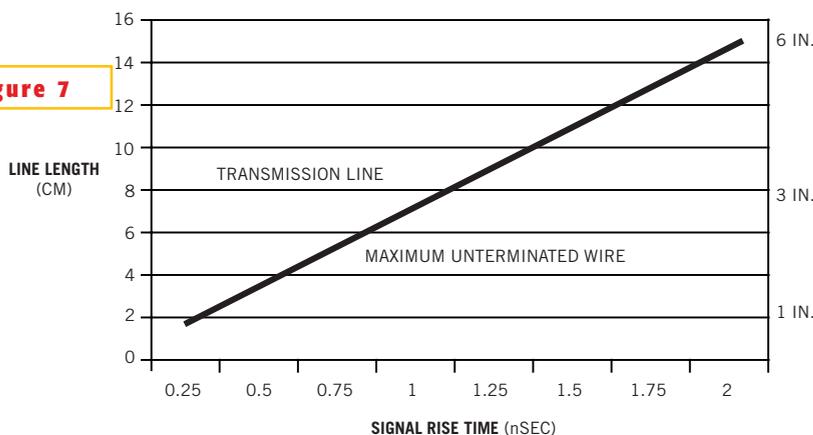
This graph shows the boundary between wire pairs, which don't need terminations, and transmission lines, which do.

causing them to move, and the process continues, creating a wave. The water wave consists of ripples that move along the surface away from the initial disturbance. Individual water particles move mainly up and down with a slight side motion. However, the cumulative effect of all the particles produces a wave that moves radially outward from the initial disturbance point.

Electromagnetic waves, which have a property uniquely different from those of other waves, are of interest in transmission lines. Electromagnetic waves can propagate in a vacuum without any matter present. So neither the copper on the pc board nor the electrons in it needs to move to get a transient signal to its destination. Similarly, the ability of electromagnetic waves to propagate in a vacuum allows communication with satellites in the vacuum of outer space.

A compact set of principles known as Maxwell's equations describes electromagnetic-wave phenomena. These equations are based upon experimental observations and provide the most accurate model of the basic relationships of electric fields, charges, currents, and magnetic fields. The equations cover the en-
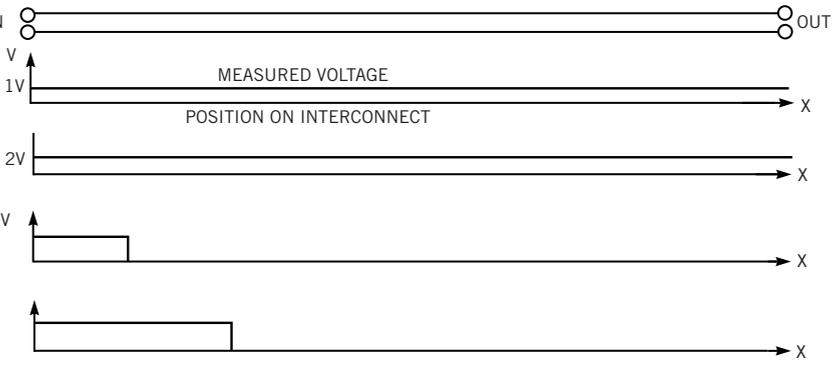

Figure 8

NOTES:
PC-BOARD INTERCONNECT BEHAVING AS A TRANSMISSION LINE.
SIGNAL PROPAGATES AS ELECTROMAGNETIC WAVE (SPEED OF LIGHT $\div \sqrt{\text{RELATIVE DIELECTRIC CONSTANT}}$).

$$\text{PROPAGATION TIME} = \frac{\text{DISTANCE}}{\text{VELOCITY}} = \frac{D}{C/\sqrt{\epsilon_R}} \ .$$

With short interconnections or low-frequency signals, the applied signal voltage appears instantaneously at every point on the interconnection (a). In (b), by t_1, the applied signal voltage has traveled only a short distance from the source end. Only after the passage of time t_2 does the voltage appear at a point closer to the load.

tire electromagnetic spectrum of radio waves, infrared rays, visible-light rays, X-rays, and gamma rays.

ELECTROMAGNETIC-WAVE PROPAGATION

Electromagnetic waves consist of time-varying electric and magnetic fields. When the electric charge on a pair of conductors changes, the electric field that the charges create changes. This time-varying electric field generates a magnetic field. The time-varying magnetic field, in turn, creates an electric field. (Two charged electrodes or a changing magnetic field can create an electric field.) These time-varying fields continue to generate one another in an ever-expanding region, and the resulting waves propagate away from the location of the initial charge change.

Figure 9 shows both the electric and magnetic waves traveling on a transmission line. These waves are the result of a sinusoidal signal at the source. These waves are in a transverse-electromagnetic (TEM) mode because the fields are perpendicular (transverse) to the direction of travel. The electric (E) and magnetic (H) fields are also perpendicular to each other.

Wave propagation in free space

The permittivity and permeability of the medium in which it travels determine

an electromagnetic wave's velocity of propagation.

Permittivity, ϵ, is the ability of a dielectric to store electrical potential energy under the influence of an electric field. The permittivity of free space is $\epsilon_0 = (1/36\pi) \times 10^{-9}$ F/m.

Permittivity is an important parameter of a capacitor. The presence of farads in the measurement unit should help you to remember the definition and to distinguish permittivity from permeability.

Permeability, μ, is the property of a magnetic substance that determines the degree to which the substance modifies the magnetic flux in the region of a magnetic field that the substance occupies. The permeability of free space is $\mu_0 = 4\pi \times 10^{-7}$ H/m. Velocity, v, in free space is $v_0 = 1/\sqrt{\epsilon_0 \mu_0} = 300 \times 10^6$ m/sec.

Wave propagation in materials other than free space

Although you could express permittivity and permeability as absolute numbers, you almost always express each quantity as a relative value. The relative value is the ratio of the permittivity or permeability of the substance you are characterizing to the corresponding free-space value. Another name for relative permittivity is relative dielectric constant or, more commonly, dielectric constant. **Table 1** shows the dielectric constants
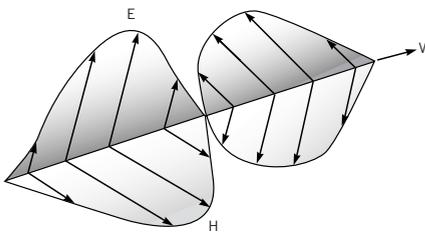

Figure 9

Both electric and magnetic waves travel on transmission lines. These waves, which are the result of a sinusoidal signal at the source, are in a transverse-electromagnetic mode. The fields are perpendicular (transverse) to the direction of travel. The electric (E) and magnetic (H) fields are also perpendicular to each other.

and relative permeabilities of various materials.

Skin effect

To an electric field outside a good conductor, such as copper, the conductor appears to exhibit infinite conductivity. However, inside a perfect conductor, there can be no electric or magnetic field. Inside a good conductor, a varying electric field can penetrate only a thin region, or skin. As the signal frequency increases, the thickness of the skin decreases.

The formula for calculating the skin depth in meters is $\sqrt{\rho/\pi\mu f}$, where ρ is the conductor resistivity in ohm-meters, $\mu=4\pi10^{-7}$H/m, and f is the frequency in hertz. (The skin depth in copper, for example, might be only 10 μm.)

Speed of wave on a transmission line

Table 2 shows electromagnetic-wave speeds on transmission lines using various dielectrics.

LUMPED-COMPONENT CIRCUIT MODELS

The above electromagnetic-wave models are the most accurate that physicists have devised to explain their experimental observations. However, these models and their accompanying equations are inconvenient for electronics engineers who are accustomed to circuit models that contain lumped components, such as resistors, capacitors, and inductors. The designers of the lumped-component circuit models intended them to handle only dc and very-low-frequency currents, but engineers have adapted these models to represent transmission lines.

Many small, lumped R, L, and C components represent the transmission line's continuous uniform resistance, inductance, and capacitance. Conventional currents and voltages represent the electromagnetic waves. **Figure 10** shows two versions of this lumped-component model. The second-order transmission-
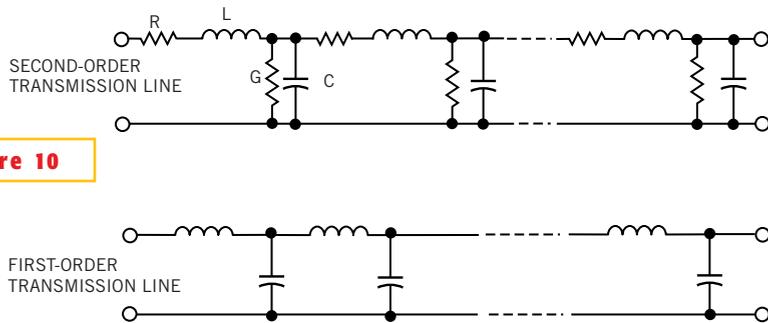

The second-order transmission-line model is more accurate, because it incorporates all three component types. However, the first-order model sees more use because most transmission-line cables and pc-board conductors have L and C values that dwarf the R values.

line model is more accurate, because it incorporates all three component types. However, engineers more often use the first-order model because most transmission-line cables and pc-board conductors have L and C values that dwarf the R values. Therefore, it is often safe to ignore series resistance and shunt conductance.

Characteristic impedance of circuit model

The most important parameter of the transmission line's lumped-component circuit model is the characteristic impedance, Z_0. Z_0 is the effective transmission-line impedance that the source signal driver sees during the signal's high-speed transition. After the transition ends, the impedance of the cable or pc-board conductors returns to that of simple wires, which each have a near-zero resistance. The most surprising aspect of a transmission line's Z_0, however, is that even though the circuit model consists of inductors and capacitors, Z_0 is effectively a pure resistor with no reactive component. Moreover, unlike the conductors' series resistance, Z_0 is not near-zero in value.

During the signal transition, the driver sees only Z_0 and does not see the load

impedance, Z_L, which is at the end of the transmission line.

You can derive Z_0 from the transmission-line circuit model (**Figure 11**).

In the spirit of differential calculus, consider a short length, Δx, of the transmission line. The quantity ΔV is the drop across the series inductor, L. The voltage per unit length, $dV/dx=z\cdot I$, where z is the impedance and I is the current. The quantity ΔI is the current that flows through the capacitor, C, so $dI/dx=y\cdot V$, where y is the admittance. Admittance, whose units are siemens, is the reciprocal of impedance. So, you derive Z_0 with the following equations:

$$\Delta V=V_S-V_L, \text{ and}$$

$$dV/dx=zI=j\omega LI, \qquad (1)$$

where $\omega=2\pi f$, and z=the impedance per unit length of series elements.

Also, $\Delta I=I_S-I_L$, and

$$dI/dx=yV=j\omega CV, \qquad (2)$$

where y=the admittance per unit length of shunt elements.

By taking the second derivative of **Equation 1** and substituting in **Equation 2**, **Equation 2** becomes $d^2V/dx^2=z\times dI/dx=zyV$, whose solution is

$$V = V_1 e^{\sqrt{zy}\,x} + V_2 e^{-\sqrt{zy}\,x}, \qquad (3)$$

and

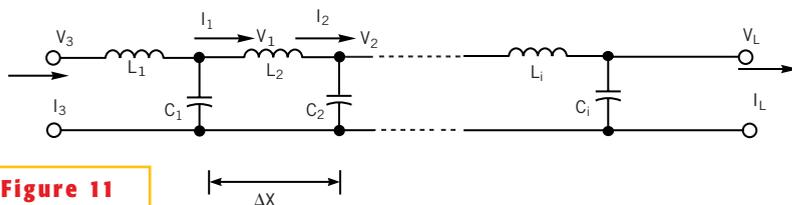$$I = I_1 e^{\sqrt{zy}\,x} + I_2 e^{-\sqrt{zy}\,x}. \qquad (4)$$

Figure 11

You can derive Z_0 from the transmission-line circuit model.

Equation 3 is the solution to the equation $d^2V/dx^2 = z \cdot dI/dx = zyV$, where V is the voltage at any point on the transmission line as x approaches zero and $V_L = V_1 + V_2$.

When you substitute **equations 3** and **4** into **equations 1** and **2** and combine them, the equations become:

$$\text{INCIDENT } V_1 = \sqrt{z/y}\ I_1,$$

$$\text{REFLECTED } V_2 = -\sqrt{z/y}\ I_2, \text{ AND}$$

$$Z_0 = \sqrt{z/y} = \sqrt{j\omega L/j\omega C} = \sqrt{L/C},$$

which is the characteristic impedance.

Note that the characteristic impedance is a pure resistance that is completely independent of frequency, even though the transmission-line circuit consists of reactive inductive and capacitive elements.

A derived formula expresses velocity as a function of the transmission line's L and C per unit length. Velocity$=1/\sqrt{LC}$. Usually, however, it is easier to derive the velocity from the speed of light divided by the square root of the relative dielectric constant, ϵ. You can more easily obtain the value of ϵ than you can determine the transmission line's L and C per unit length.

Characteristic impedance from physical dimensions

Because the inductance and capacitance of a transmission line depend solely on the physical characteristics of the implementation, you can also derive a transmission line's characteristic impedance from the line's physical dimensions. **Figure 12** shows three configuration examples. Typical dimensions for microstrip transmission lines are: T (conductor thickness)= 0.0015 in. for 1-oz copper, H (substrate thickness)=0.062 in. for G-10 glass epoxy, and W (conductor width)=0.015 in.

Multiple loads on a transmission line

Most academic textbooks on transmission lines assume that every transmission line has only one load, but many pc boards place multiple loads on one transmission line. So, practicing engineers have developed an empirical formula to modify the characteristic impedance and signal velocity to account for the additional capacitance that multiple loads add. The formula is as follows:

CABLE
$Z_0 = 138 \log_{10}(D/d)$.
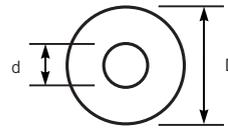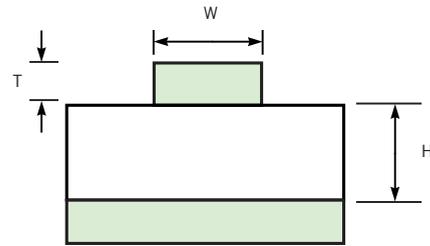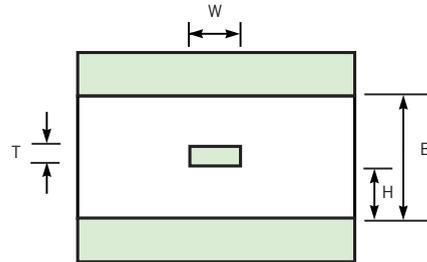$C = 24.1K_e/\log_{10}(D/d)$ pF/m.
$L = 0.46 \log_{10}(D/d)$ µH/m.

Figure 12

MICROSTRIP
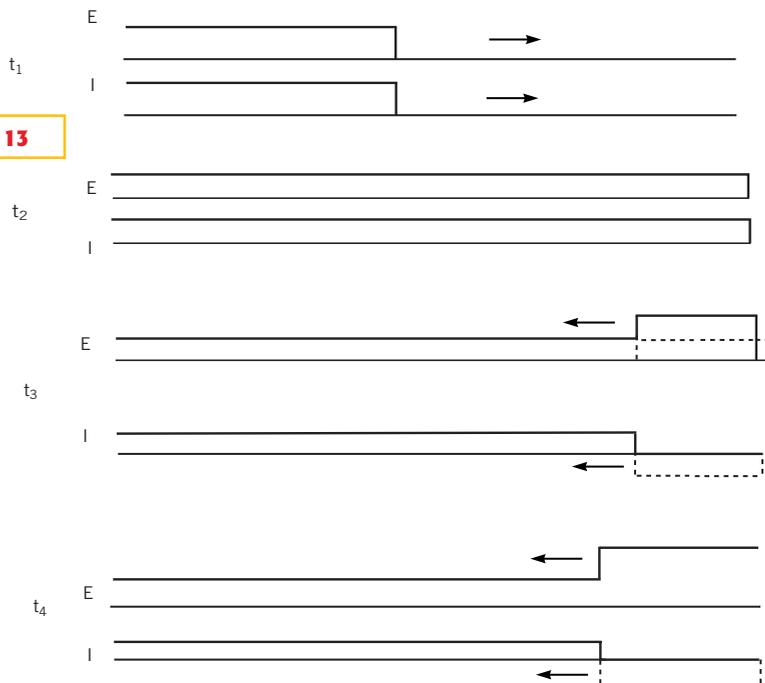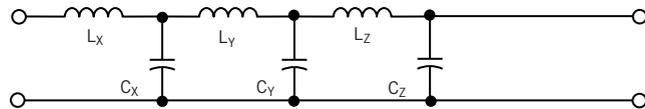$Z_0 = (87/\sqrt{\epsilon_R + 1.41})\ln.(5.98H/(0.8W+T))$.

STRIPLINE
$Z_0 = (60/\sqrt{\epsilon_R})\ln.(4B/0.67\pi W(0.8+(TW)))$.

Because the inductance and capacitance of a transmission line depend solely on the physical characteristics of the implementation, you can derive a transmission line's characteristic impedance from the line's physical dimensions.

Figure 13

The lumped-component circuit model uses uniformly distributed inductance and capacitance, which appear as many infinitesimal lumped elements. Voltage and current define the signal.
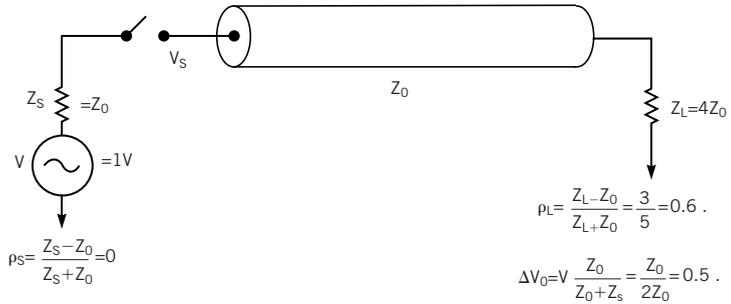
$$Z_{0\,EFFECTIVE} = Z_0' = Z_0 \div$$
$$\sqrt{1+(C_L/C_O)}, \text{ AND}$$

$$t_{PD\,EFFECTIVE} = t_{PD}' =$$
$$t_{PD}\sqrt{1+(C_L/C_O)}.$$

C_0 is unloaded capacitance per unit length of the transmission line, and C_L is the added load capacitance per unit length. The added capacitance decreases the effective Z_0 and increases the propagation delay. Although these formulas aren't in the theoretical textbooks, they are in many application notes, including one classic—the *Motorola* (www. motorola.com) *MECL* (Motorola emitter-coupled logic) *Handbook*. This handbook is also a source of C_0 values for microstrip transmission lines.

Figure 14

Reflections on transmission lines with open-circuit loads

What follows is a heuristic description of a signal wave traveling on a transmission line from the source to an open-cir-

$\rho_S = \dfrac{Z_S - Z_0}{Z_S + Z_0} = 0$

$\rho_L = \dfrac{Z_L - Z_0}{Z_L + Z_0} = \dfrac{3}{5} = 0.6$.

$\Delta V_0 = V\,\dfrac{Z_0}{Z_0 + Z_s} = \dfrac{Z_0}{2Z_0} = 0.5$.

$V = V_1 + V_R.$
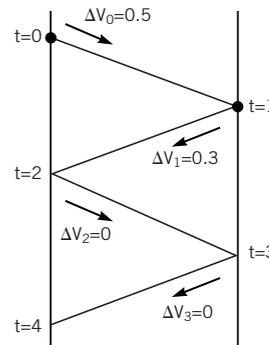$V_1 = \Delta V_0 + \rho_L \Delta V_0.$
$V_1 = 0.5 + (0.6)(0.5) = 0.8V.$

$\Delta V_1 = \rho_L \Delta V_0 = 0.3V.$

$V_3 = 0.8V.$

You can use a lattice diagram to calculate reflections.

cuited load end where the signal reflects back to the source. The lumped-component circuit model in **Figure 13** uses uniformly distributed inductance and capacitance, which appear as many infinitesimal lumped elements. Voltage and current define the signal.

When the signal arrives at the load end of the transmission line, its orderly progress is interrupted. At time t_1, the wave is approaching the load end. The voltage and current of the signal's energy wave are progressing together: L_Y is carrying current, but L_Z is not, and C_Y is charged to E, but C_Z is not. There is voltage across C_Y, so current immediately begins to flow through L_Z. Because L_Z carries as much current as L_Y, the charge on C_Y increases no further than E. Current entering L_Z flows into C_Z, charging it to E. At time t_2, the energy wave reaches the end. All capacitors are now charged to E, and all inductors are now carrying current I.

At this point, the progress of the wave cannot continue. L_Z is carrying current to C_Z, but there is no inductor beyond C_Z to

Figure 15
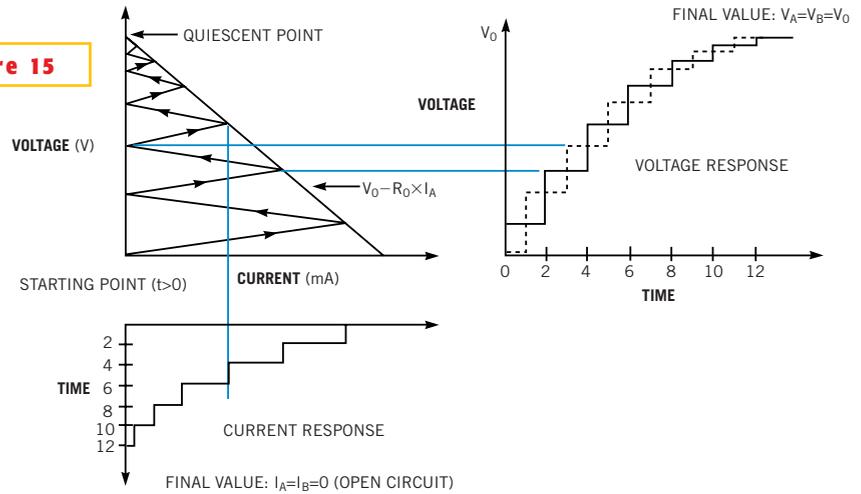
You might calculate reflections using the more sophisticated Bergeron-plot technique.

act as an outlet for current as C_Z becomes charged. You can see the result at time t_3 when C_Z becomes overcharged. L_Z cannot stop carrying current until it exhausts its magnetic energy. L_Z continues to drive current into C_Z until C_Z's voltage is 2E. At this point, the current in L_Z is zero. At time t_4, when L_Z stops carrying current, all current that L_Y carries is driven into C_Y, doubling the voltage of C_Y and forcing the current in L_Y to stop.

At the same time, the voltages at the two ends of L_Z become equal, so the overcharge on C_Z cannot escape. The voltage

across C_Y and C_Z is 2E, and the current in L_Y and L_Z is zero. This process continues progressively along the line, assuming that the energy of each L equals the energy of each C. The result is a voltage wave of magnitude E traveling back to the source and adding to the original wave. The reflected-current wave appears as a negative I wave traveling back to the source and canceling the original wave, so that the net current is zero. No current flows back to the source, but a wave front does propagate. The current that flowed away from the source continues to do so until it meets the reflected wave.

Remember that this explanation is merely an attempt to model experimental data (reality) with lumped components originally intended for low-frequency circuits. You can more accurately model these phenomena with electro-

SUMMARY OF KEY POINTS ON TRANSMISSION LINES

- Interconnections via pc boards or cables can behave differently depending upon the ratio of the interconnection length to the signal's wavelength. Signal-integrity problems occur if the interconnection is a transmission line and the system's design doesn't account for the transmission-line characteristics.
- A high-speed signal on a pc board or in a cable travels at the speed of electromagnetic waves (speed of light) modified by the dielectric constant (permittivity) of the material between the conductors.
- You should treat an interconnection as a transmission line when the signal's rise time is less than twice the electromagnetic wave's propagation time over the length of the interconnect. In other words, treat interconnections as transmission lines if their length is a significant portion of the signal's wavelength. Remember, however, that the criteria of $t_{RISE} < 2$ times time of flight and interconnect length > 0.1 wavelength are not absolute but are rules of thumb. The effects of the impedance mismatches and reflections are analog in nature.
- During a signal's high-frequency transitions, the characteristic impedance, Z_0, of a transmission line is purely resistive and depends upon the uniformly distributed values of capacitance and inductance; that is, Z_0 depends on the geometry of the conductors and dielectric. At low frequencies, the interconnection functions as simple wires between a driver and its load.
- To avoid signal-integrity problems in most situations, the transmission-line load impedance, the source impedance, or both should equal Z_0 or should add a termination resistor to make it equal Z_0.

magnetic-wave differential equations or wave equations and boundary conditions.

Reflections on transmission lines with short-circuit loads

Reflections for a short-circuit load are similar to the above with the voltage and current waveforms interchanged.

Reflections on transmission lines with a matched load impedance

When the load impedance equals the transmission line's characteristic impedance, no reflection occurs. The load looks like the uniform characteristic impedance of the transmission line, so the load absorbs all of the wave energy. In general, this configuration is ideal. Also, in general, when the load impedance doesn't match the transmission line's characteristic impedance (Z_0), you should add a Z_0 termination resistor in parallel with the load. The load impedance is usually high, so its value is inconsequential. Because the source impedance is usually low, you can put a termination resistor in series with the source so the total resistance equals Z_0.

Figure 16
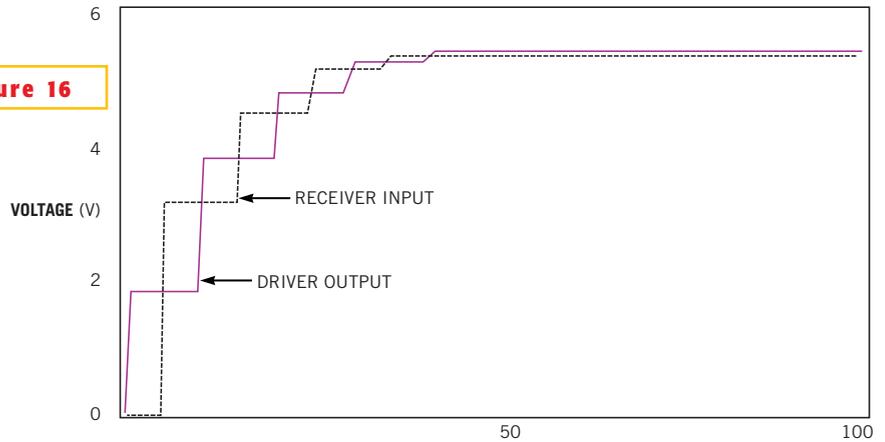
This transmission line's source and load impedances produce reflections that cause the voltage to rise in stair-step fashion. This phenomenon occurs when Z_S and Z_L are only slightly higher than Z_0.

Use caution in applying these guidelines, because exceptions do occur.

Reflections on transmission lines with a discontinuity or intermediate load

A discontinuity (change in the uniform characteristic impedance) or intermediate load on a transmission line causes a reflection just as does an impedance mismatch at the load end. Any intermediate reflection travels back to the source and adds to the other waves traveling on the transmission line.

Reflections on transmission lines with any load impedance

If the load impedance, Z_L, doesn't match the transmission line's Z_0 and if Z_L

is neither zero nor infinity, the signal energy that the load doesn't absorb reflects back toward the source. When the reflection reaches the source, if the source impedance, Z_S, doesn't match Z_0, there is a reflection from the source back toward the load. The reflections continue until the load, the source, and losses along the transmission line fully absorb the wave's energy.

The following formula for ρ determines the size of the reflection at a transmission line's mismatched end (source or load):

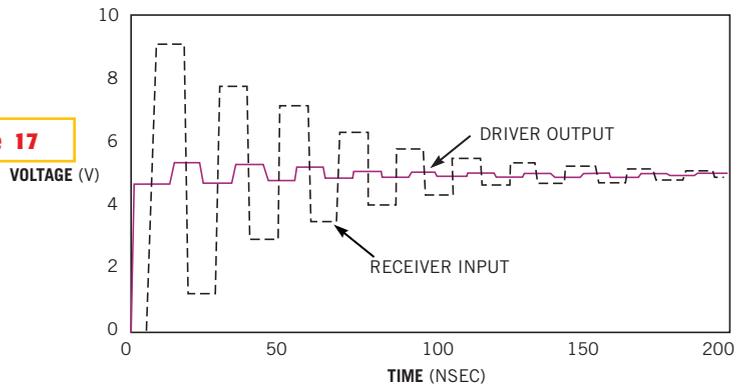$$V_L = V_{INCIDENT} + V_{REFLECTED};$$
$$I_L = I_{INCIDENT} + I_{REFLECTED};$$

$$Z_L = \frac{V_L}{I_L} = \frac{V_I + V_R}{I_I + I_R} =$$

$$\frac{V_I + V_R}{\dfrac{V_I}{Z_0} - \dfrac{V_R}{Z_0}} = Z_0 \frac{V_I + V_R}{V_I - V_R}; \text{ AND}$$

REFLECTION FACTOR $\rho =$

$$\frac{V_R}{V_I} = \frac{Z_L - Z_0}{Z_L + Z_0}.$$

Figure 17

VOLTAGE (V)

DRIVER OUTPUT

RECEIVER INPUT

TIME (NSEC)

A different configuration of source impedance, transmission line, and load produces reflections that appear as ringing on the transmission line. The overshoot can damage the receiver circuit, or, if the receiver reads the signal before the ringing subsides, corrupt the data.

This formula is for the load mismatch. You can use the formula for a source end if you substitute Z_S for Z_L. For the open-circuit calculation, divide the formula by Z_L to avoid infinity over infinity.

Lattice diagram

You can use a lattice diagram to calculate reflections (**Figure 14**). First, note that the initial source voltage, V_0, at V_S is 0.5V, because the voltage, V, of 1V divides between Z_S and Z_0. The lattice diagram shows that at t=0, the V_0=0.5V voltage wave propagates to the load end where the reflection factor is 0.6. So the total

voltage, V_L, at the load at t=1 is 0.8V. The reflected voltage of 0.3V then travels back to the source on the 0.5V incident wave. When the 0.3V wave reaches the source, no reflection occurs because $Z_S = Z_0$, making the reflection factor zero.

Bergeron plot

You might calculate the reflections using the more sophisticated Bergeron-plot technique (**Figure 15**). This technique is especially useful for nonlinear terminations; that is, where the driver or receiver has different impedances depending on whether the signal level is high or low.

Stair-stepping reflections

Figure 16 is an example of a transmission line whose source and load impedances produce reflections that cause the voltage to rise in stair-step fashion. This phenomenon occurs when Z_S and Z_L are only slightly higher than Z_0. The first signal wave is a fraction of the full driver signal because the driver signal divides between Z_S and Z_0. The signal at the receiver needs approximately four round trips of reflections (40 nsec) be-

THE BERGERON-PLOT TECHNIQUE IS ESPECIALLY USEFUL FOR NONLINEAR TERMINATIONS; THAT IS, WHERE THE DRIVER OR RECEIVER HAS DIFFERENT IMPEDANCES DEPENDING ON WHETHER THE SIGNAL LEVEL IS HIGH OR LOW.

fore it reaches its full value.

Ringing reflections

Figure 17 provides an example of a different configuration of source impedance, transmission line, and load that produces reflections, which appear as ringing on the transmission line. The overshoot can damage the receiver circuit or, if the receiver reads the signal before the ringing subsides, corrupt the data.

The waveform in the **figure** has square edges instead of curves only because it is from a simulation. In this simulation, $Z_L >> Z_0$. The ringing decays as the load and cable losses gradually absorb the signal's energy. You can now see that the criteria of $t_{RISE} < 2$ times time of flight and interconnect length > 0.1 wavelength are not absolute demarcations but are rules of thumb because the effects of impedance mismatches and reflections are analog in nature.□

Author's Biography

James Sutherland is a senior applications engineer with California Micro Devices (Milpitas, CA), where he has worked for the past year. His responsibilities include defining new products and assisting customers. In a previous job, he developed ICs for DSPs, FPGAs, and disk drives. He holds a BSEE and an MBA from Stanford University (Palo Alto, CA) and is a member of the IEEE. His activities outside work include worldwide travel, photography, and parenting.