

RF Basics

Dr Francesco Fornetti

1 Transmission Lines

1.1 Introduction to Transmission Lines

One of the most common transmission lines is the *coaxial cable*.

A coaxial cable is a two-conductor cable made of a single conductor surrounded by a braided wire jacket, with a plastic insulating material separating the two. As such, the outer (braided) conductor completely surrounds the inner (single wire) conductor and the two conductors are insulated from each other for the entire length of the cable.

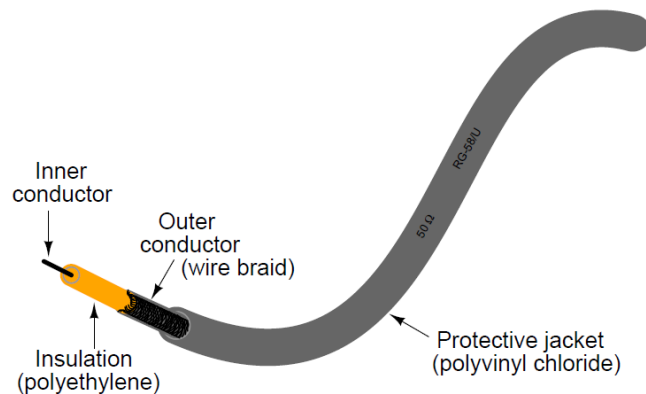


Figure 1.1-1 Coaxial cable construction

Continuous direct current (DC) – such as that used by an ohmmeter to check the cable’s resistance – shows the two conductors to be completely insulated from each other, with nearly infinite resistance between the two. However, the cable’s response, due to the effects of capacitance and inductance distributed along the length of the cable, to short-duration voltage “pulses” and high-frequency AC signals is quite different! When rapidly-changing the applied voltage, the cable is such that it acts as a *finite* impedance (typically 50 or 75 Ω), drawing current proportional to the applied voltage. What we would normally dismiss as being just a pair of wires becomes an important circuit element in the presence of transient and high-frequency AC signals, with characteristic properties all its own. When expressing such properties, we refer to the wire pair as a *transmission line*.

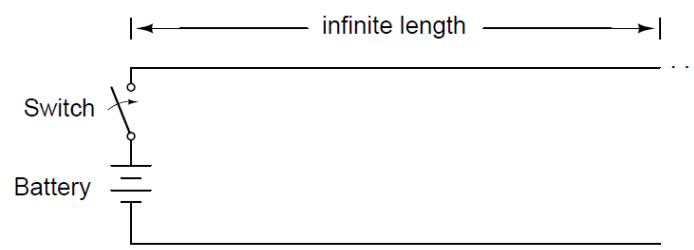


Figure 1.1-2 Driving an Infinite transmission line

Suppose, though, that we had a set of parallel wires of *infinite* length (Figure 1.1-2), with nothing connected at the end. What would happen when we close the switch? Being that there is no longer a load at the end of the wires, this circuit is open. Would there be no current at all?

Despite being able to avoid wire resistance through the use of superconductors in this “thought experiment,” we cannot eliminate capacitance along the wires’ lengths. *Any* pair of conductors separated by an insulating medium creates capacitance between those conductors (Figure 1.1-3)

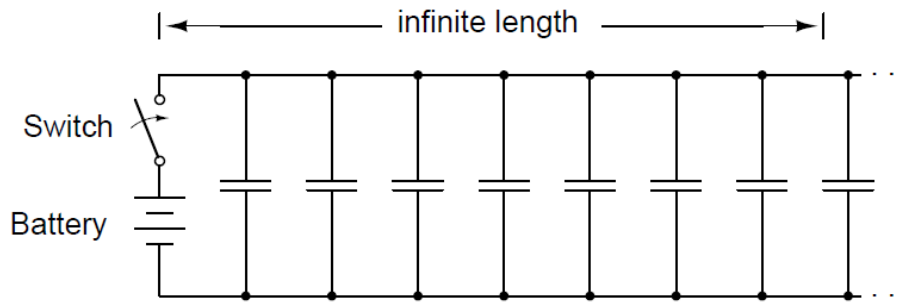


Figure 1.1-3 Equivalent circuit showing stray capacitance between conductors.

Voltage applied between two conductors creates an electric field between those conductors. Energy is stored in this electric field, and this storage of energy results in an opposition to change in voltage. The reaction of a capacitance against changes in voltage is described by the equation $i = C(de/dt)$, which tells us that current will be drawn proportional to the voltage's rate of change over time. Thus, when the switch is closed, the capacitance between conductors will react against the sudden voltage increase by charging up and drawing current from the source. According to the equation, an instant rise in applied voltage (as produced by perfect switch closure) gives rise to an infinite charging current.

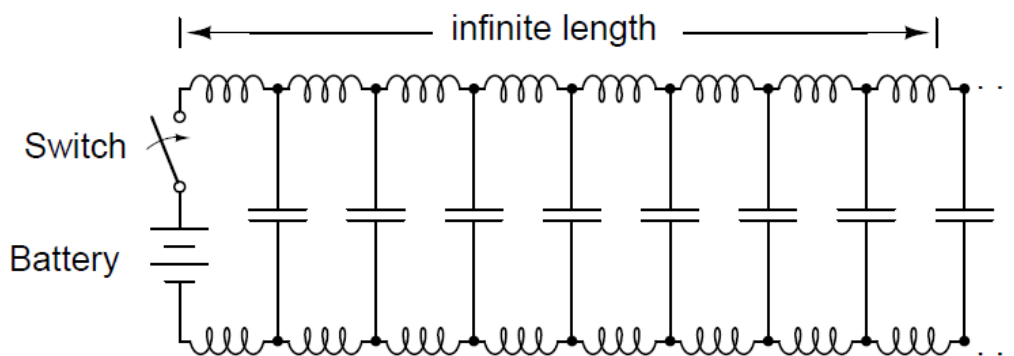


Figure 1.1-4 Equivalent circuit showing stray capacitance and inductance

However, the current drawn by a pair of parallel wires will not be infinite, because there exists series impedance along the wires due to inductance. (Figure 1.1-4) Remember that current through *any* conductor develops a magnetic field of proportional magnitude. Energy is stored in this magnetic field, (Figure 1.1-5) and this storage of energy results in an opposition to change in current. Each wire develops a magnetic field as it carries charging current for the capacitance between the wires, and in so doing drops voltage according to the inductance equation $e = L(di/dt)$. This voltage drop limits the voltage rate-of-change across the distributed capacitance, preventing the current from ever reaching an infinite magnitude.

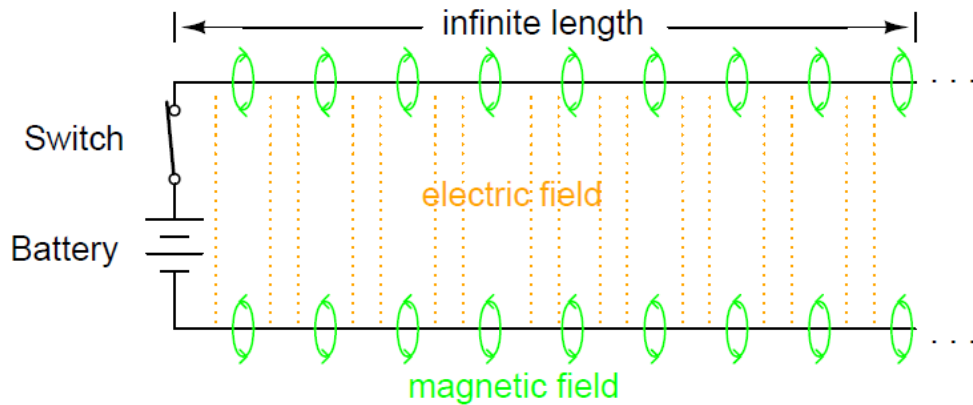


Figure 1.1-5 Voltage charges capacitance, current charges inductance

Because the electrons in the two wires transfer motion to and from each other at nearly the speed of light, the “wave front” of voltage and current change will propagate down the length of the wires at that same velocity, resulting in the distributed capacitance and inductance progressively charging to full voltage and current, respectively, as shown in Figure 1.1-6.

The end result of these interactions is a constant current of limited magnitude through the battery source. Since the wires are infinitely long, their distributed capacitance will never fully charge to the source voltage, and their distributed inductance will never allow unlimited charging current. In other words, this pair of wires will draw current from the source so long as the switch is closed, behaving as a constant load. No longer are the wires merely conductors of electrical current and carriers of voltage, but now constitute a circuit component in themselves, with unique characteristics. No longer are the two wires merely *a pair of conductors*, but rather a *transmission line*.

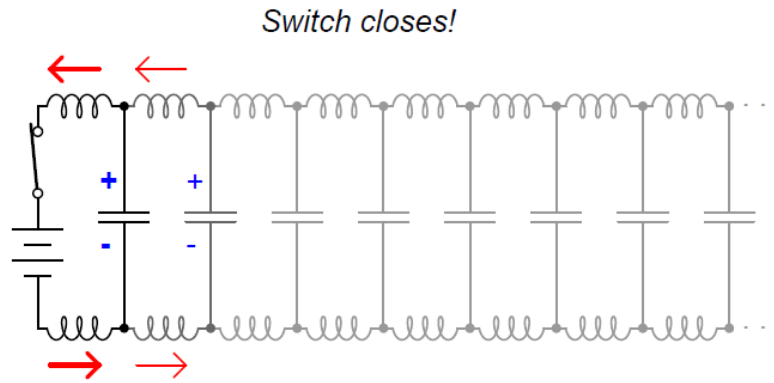


Figure 14.11: *Begin wave propagation.*

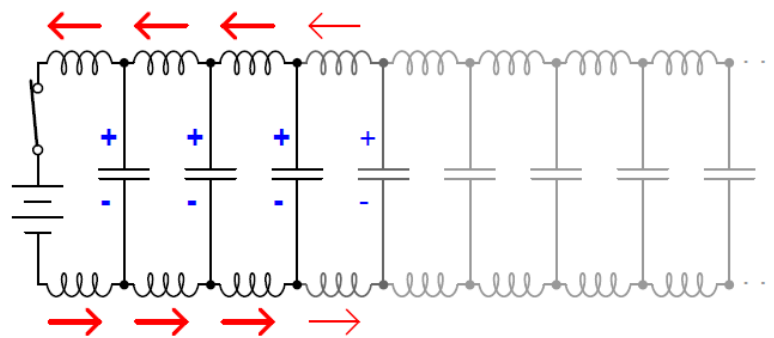


Figure 14.12: *Continue wave propagation.*

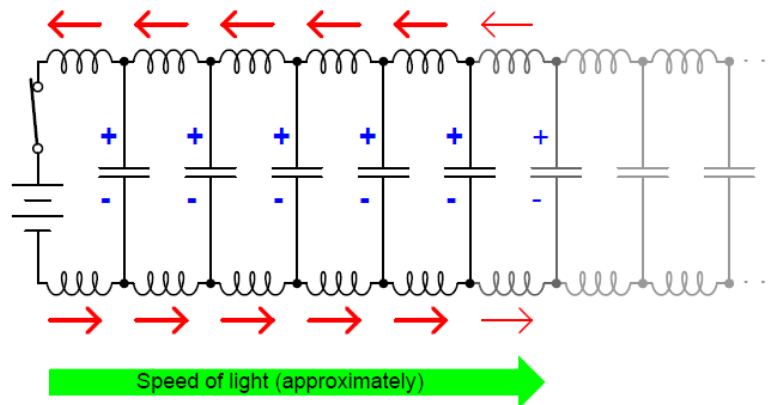


Figure 1.1-6 Propagate at speed of light

As a constant load, the transmission line's response to applied voltage is resistive rather than reactive, despite being comprised purely of inductance and capacitance (assuming superconducting wires with zero resistance). We can say this because there is no difference from the battery's perspective between a resistor eternally dissipating energy and an infinite transmission line eternally absorbing energy. The impedance (resistance) of this line in ohms is called the characteristic impedance, and it is fixed by the geometry of the two conductors. For a parallel-wire line with air insulation, the characteristic impedance may be calculated as such:



$$Z_0 = \frac{276}{\sqrt{k}} \log \frac{d}{r}$$

Where,

Z_0 = Characteristic impedance of line

d = Distance between conductor centers

r = Conductor radius

k = Relative permittivity of insulation bwtween conductors

If the transmission line is coaxial in construction, the characteristic impedance follows a different equation:



$$Z_0 = \frac{138}{\sqrt{k}} \log \frac{d_1}{d_2}$$

Where,

Z_0 = Characteristic impedance of line

d_1 = Inside diameter of outer conductor

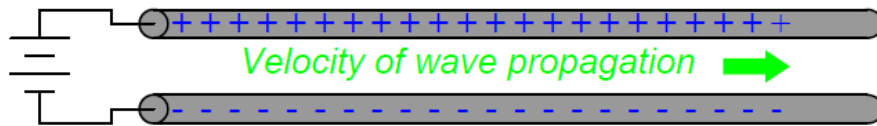
d_2 = Outside diameter of inner conductor

k = Relative permittivity of insulation
between conductors

In both equations, identical units of measurement must be used in both terms of the fraction.

If the insulating material is other than air (or a vacuum), both the characteristic impedance and the propagation velocity will be affected. The ratio of a transmission line's true propagation velocity and the speed of light in a vacuum is called the *velocity factor* of that line.

Velocity factor is purely a factor of the insulating material's relative permittivity (otherwise known as its *dielectric constant*), defined as the ratio of a material's electric field permittivity to that of a pure vacuum. The velocity factor of any cable type – coaxial or otherwise – may be calculated quite simply by the following formula:



$$\text{Velocity factor} = \frac{v}{c} = \frac{1}{\sqrt{k}}$$

Where,

k = Relative permittivity of insulation between conductors

v = Velocity of wave propagation

c = Velocity of light in a vacuum

Characteristic impedance is also known as *natural impedance*, and it refers to the equivalent resistance of a transmission line if it were infinitely long, owing to distributed capacitance and inductance as the voltage and current “waves” propagate along its length at a propagation velocity equal to some large fraction of light speed.

It can be seen in either of the first two equations that a transmission line’s characteristic impedance (Z_0) increases as the conductor spacing increases. If the conductors are moved away from each other, the distributed capacitance will decrease (greater spacing between capacitor “plates”), and the distributed inductance will increase (less cancellation of the two opposing magnetic fields). Less parallel capacitance and more series inductance results in a smaller current drawn by the line for any given amount of applied voltage, which by definition is a greater impedance. Conversely, bringing the two conductors closer together increases the parallel capacitance and decreases the series inductance. Both changes result in a larger current drawn for a given applied voltage, equating to a lesser impedance.

Barring any dissipative effects such as dielectric “leakage” and conductor resistance, the characteristic impedance of a transmission line is equal to the square root of the ratio of the line’s inductance per unit length divided by the line’s capacitance per unit length:

$$Z_0 = \sqrt{\frac{L}{C}}$$

Where,

Z_0 = Characteristic impedance of line

L = Inductance per unit length of line

C = Capacitance per unit length of line

1.2 Finite Length transmission line

A transmission line of infinite length is an interesting abstraction, but physically impossible.

All transmission lines have some finite length, and as such do not behave precisely the same as an infinite line. If an infinitely long piece of 50 Ω cable was measured with an ohmmeter one would be able to measure 50 Ω worth of resistance between the inner and outer conductors. But since it cannot be infinite in length, it will always measure as “open” (infinite resistance).

Nonetheless, the characteristic impedance rating of a transmission line is important even when dealing with limited lengths. An older term for characteristic impedance is *surge impedance*. If a transient voltage (a “surge”) is applied to the end of a transmission line, the line will draw a current proportional to the surge voltage magnitude divided by the line’s surge impedance ($I=E/Z$). This simple, Ohm’s Law relationship between current and voltage will hold true for a limited period of time, but not indefinitely.

If the end of a transmission line is open-circuited – that is, left unconnected – the current “wave” propagating down the line’s length will have to stop at the end, since electrons cannot flow where there is no continuing path. This abrupt cessation of current at the line’s end causes a “pile-up” to occur along the length of the transmission line, as the electrons successively find no place to go.

When this electron “pile-up” propagates back to the battery, current at the battery ceases, and the line acts as a simple open circuit. All this happens very quickly for transmission lines of reasonable length, and so an ohmmeter measurement of the line never reveals the brief time period where the line actually behaves as a resistor. For a mile-long cable with a velocity factor of 0.66 (signal propagation velocity is 66% of light speed, or 122,760 miles per second), it takes only 1/122,760 of a second (8.146 microseconds) for a signal to travel from one end to the other.

For the current signal to reach the line’s end and “reflect” back to the source, the round-trip time is twice this figure, or 16.292 μ s.

A signal propagating from the source-end of a transmission line to the load-end is called an *incident wave*. The propagation of a signal from load-end to source-end (such as what happened in this example with current encountering the end of an open-circuited transmission line) is called a *reflected wave*.

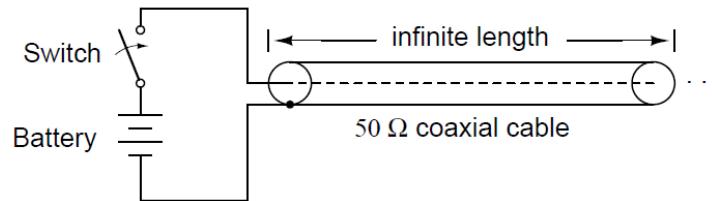
High-speed measurement instruments are able to detect this transit time from source to line-end and back to source again, and may be used for the purpose of determining a cable’s length. This technique may also be used for determining the presence *and* location of a break in one or both of the cable’s conductors, since a current will “reflect” off the wire break just as it will off the end of an open-circuited cable. Instruments designed for such purposes are called *time-domain reflectometers* (TDRs). The basic principle is identical to that of sonar range finding: generating a sound pulse and measuring the time it takes for the echo to return.

A similar phenomenon takes place if the end of a transmission line is short-circuited: when the voltage wave-front reaches the end of the line, it is reflected back to the source, because voltage cannot exist between two electrically common points. When this reflected wave reaches the source, the source sees the entire transmission line as a short-circuit. Again, this happens as quickly as the signal can propagate round-trip down and up the transmission line at whatever velocity allowed by the dielectric material between the line’s conductors.

Reflections may be eliminated from the transmission line if the load’s impedance exactly equals the characteristic (“surge”) impedance of the line. For example, a 50 Ω coaxial cable that is either open-circuited or short-circuited will reflect all of the incident energy back to the source. However, if a 50 Ω resistor is connected at the end of the cable, there will be no reflected energy, all signal energy being dissipated by the resistor.

This makes perfect sense if we return to our hypothetical, infinite-length transmission line example. A transmission line of 50 Ω characteristic impedance and infinite length behaves exactly like a 50 Ω resistance as measured from one end. (Figure 1.2-1) If we cut this line to some finite length, it will behave as a 50 Ω resistor to a constant source of DC voltage for a brief time, but then behave like an open- or a short-circuit, depending on what condition we leave the cut end of the line: open (Figure 1.2-2) or shorted.(Figure 1.2-3)

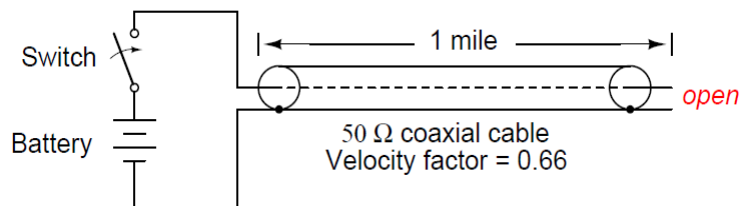
However, if we *terminate* the line with a 50Ω resistor, the line will once again behave as a 50Ω resistor, indefinitely: the same as if it were of infinite length again: (Figure 1.2-4)



Cable's behavior from perspective of battery:

Exactly like a 50Ω resistor

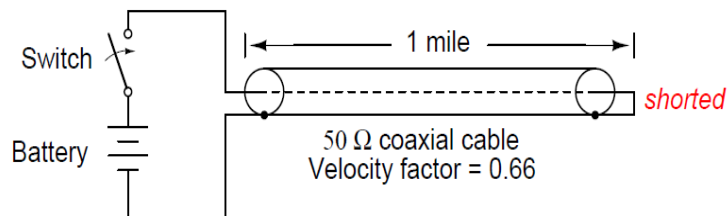
Figure 1.2-1 Infinite transmission line looks like a resistor



Cable's behavior from perspective of battery:

Like a 50Ω resistor for $16.292\mu\text{s}$,
then like an open (infinite resistance)

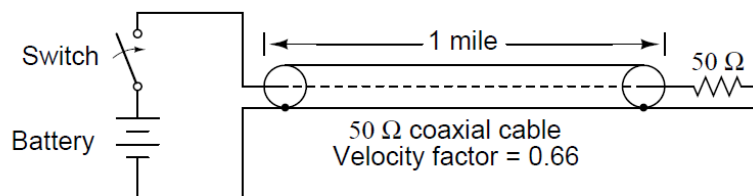
Figure 1.2-2 One mile transmission



Cable's behavior from perspective of battery:

Like a 50Ω resistor for $16.292\mu\text{s}$,
then like a short (zero resistance)

Figure 1.2-3 Shorted Transmission Line



Cable's behavior from perspective of battery:

Exactly like a 50Ω resistor

Figure 1.2-4 Line terminated in characteristic impedance

In essence, a terminating resistor matching the natural impedance of the transmission line makes the line “appear” infinitely long from the perspective of the source, because a resistor has the ability to eternally dissipate energy in the same way a transmission line of infinite length is able to eternally absorb energy.

Reflected waves will also manifest if the terminating resistance isn’t precisely equal to the characteristic impedance of the transmission line, not just if the line is left unconnected (open) or jumpered (shorted). Though the energy reflection will not be total with a terminating impedance of slight mismatch, it will be partial. This happens whether or not the terminating resistance is *greater* or *less* than the line’s characteristic impedance.

Re-reflections of a reflected wave may also occur at the *source end* of a transmission line, if the source’s internal impedance (Thevenin equivalent impedance) is not exactly equal to the line’s characteristic impedance. A reflected wave returning back to the source will be dissipated entirely if the source impedance matches the line’s, but will be reflected back toward the line end like another incident wave, at least partially, if the source impedance does not match the line. This type of reflection may be particularly troublesome, as it makes it appear that the source has transmitted another pulse.

1.3 “Long” and “Short” transmission Lines

In DC and low-frequency AC circuits, the characteristic impedance of parallel wires is usually ignored. This includes the use of coaxial cables in instrument circuits, often employed to protect weak voltage signals from being corrupted by induced “noise” caused by stray electric and magnetic fields. This is due to the relatively short time spans in which reflections take place in the line, as compared to the period of the waveforms or pulses of the significant signals in the circuit. As we saw in the last section, if a transmission line is connected to a DC voltage source, it will behave as a resistor equal in value to the line’s characteristic impedance only for as long as it takes the incident pulse to reach the end of the line and return as a reflected pulse, back to the source. After that time (a brief 16.292 μs for the mile-long coaxial cable of the last example), the source “sees” only the terminating impedance, whatever that may be.

If the circuit in question handles low-frequency AC power, such short time delays introduced by a transmission line between when the AC source outputs a voltage peak and when the source “sees” that peak loaded by the terminating impedance (round-trip time for the incident wave to reach the line’s end and reflect back to the source) are of little consequence. Even though we know that signal magnitudes along the line’s length are not equal at any given time due to signal propagation at (nearly) the speed of light, the actual phase difference between start-of-line and end-of-line signals is negligible, because line-length propagations occur within a very small fraction of the AC waveform’s period. For all practical purposes, we can say that voltage along all respective points on a low-frequency, two-conductor line are equal and in-phase with each other at any given point in time. In these cases, we can say that the transmission lines in question are electrically short, because their propagation effects are much quicker than the periods of the conducted signals.

By contrast, an electrically long line is one where the propagation time is a large fraction or even a multiple of the signal period. A “long” line is generally considered to be one where the source’s signal waveform completes at least a quarter-cycle (90° of “rotation”) before the incident signal reaches line’s end. To put this into perspective, we need to express the distance traveled by a voltage or current signal along a transmission line in relation to its source frequency. Whatever distance we calculate for a given frequency is called the wavelength of the signal.

A simple formula for calculating wavelength is as follows:

$$\lambda = \frac{v}{f}$$

Where,

λ = Wavelength

v = Velocity of propagation

f = Frequency of signal

The lower-case Greek letter “lambda” (λ) represents wavelength, in whatever unit of length used in the velocity figure (if miles per second, then wavelength in miles; if meters per second, then wavelength in

meters). Velocity of propagation is usually the speed of light when calculating signal wavelength in open air or in a vacuum, but will be less if the transmission line has a velocity factor less than 1.

If a “long” line is considered to be one at least 1/4 wavelength in length, for a 60 Hz AC power system, power lines would have to exceed 775 miles in length before the effects of propagation time became significant. Cables connecting an audio amplifier to speakers would have to be over 4.65 miles in length before line reflections would significantly impact a 10 kHz audio signal!

When dealing with radio-frequency systems, though, transmission line length is far from trivial. Consider a 100 MHz radio signal: its wavelength is a mere 9.8202 feet, even at the full propagation velocity of light (186,000 m/s). A transmission line carrying this signal would not have to be more than about 2-1/2 feet in length to be considered “long!” With a cable velocity factor of 0.66, this critical length shrinks to 1.62 feet.

When an electrical source is connected to a load via a “short” transmission line, the load’s impedance dominates the circuit. This is to say, when the line is short, its own characteristic impedance is of little consequence to the circuit’s behaviour. We see this when testing a coaxial cable with an ohmmeter: the cable reads “open” from centre conductor to outer conductor if the cable end is left unterminated. Though the line acts as a resistor for a very brief period of time after the meter is connected, it immediately thereafter behaves as a simple “open circuit”. Since the combined response time of an ohmmeter and the human being using it *greatly exceeds* the round-trip propagation time up and down the cable, it is “electrically short” for this application, and we only register the terminating (load) impedance. It is the extreme speed of the propagated signal that makes us unable to detect the cable’s 50Ω transient impedance with an ohmmeter.

If we use a coaxial cable to conduct a DC voltage or current to a load, and no component in the circuit is capable of measuring or responding quickly enough to “notice” a reflected wave, the cable is considered “electrically short” and its impedance is irrelevant to circuit function.

Taking the same length of cable, though, and using it to conduct a high-frequency AC signal could result in a vastly different assessment of that cable’s “shortness!”

When a source is connected to a load via a “long” transmission line, the line’s own characteristic impedance dominates over load impedance in determining circuit behaviour. In other words, an electrically “long” line acts as the principal component in the circuit, its own characteristics overshadowing the load’s. With a source connected to one end of the cable and a load to the other, current drawn from the source is a function primarily of the line and not the load.

This is increasingly true the longer the transmission line is. Consider our hypothetical 50Ω cable of infinite length, surely the ultimate example of a “long” transmission line: no matter what kind of load we connect to one end of this line, the source (connected to the other end) will only see 50Ω of impedance, because the line’s infinite length prevents the signal from *ever reaching* the end where the load is connected. In this scenario, line impedance exclusively defines circuit behaviour, rendering the load completely irrelevant.

The most effective way to minimize the impact of transmission line length on circuit behaviour is to match the line’s characteristic impedance to the load impedance. If the load impedance is equal to the line impedance, then *any* signal source connected to the other end of the line will “see” the exact same impedance, and will have the exact same amount of current drawn from it, regardless of line length. In this condition of perfect impedance matching, line length only affects the amount of time delay from signal departure at the source to signal arrival at the load. However, perfect matching of line and load impedances is not always practical or possible.

The next section discusses the effects of “long” transmission lines, especially when line length happens to match specific fractions or multiples of signal wavelength.

1.4 Standing waves and resonance

Whenever there is a mismatch of impedance between transmission line and load, reflections will occur. If the incident signal is a continuous AC waveform, these reflections will mix with more of the oncoming incident waveform to produce stationary waveforms called *standing waves*.

The following illustration shows how a triangle-shaped incident waveform turns into a mirror-image reflection upon reaching the line's unterminated end. The transmission line in this illustrative sequence is shown as a single, thick line rather than a pair of wires, for simplicity's sake. The incident wave is shown travelling from left to right, while the reflected wave travels from right to left. (Figure 1.4-1) If we add the two waveforms together, we find that a third, stationary waveform is created along the line's length (Figure 1.4-2)

This third, "standing" wave, in fact, represents the only voltage along the line, being the representative sum of incident and reflected voltage waves. It oscillates in instantaneous magnitude, but does not propagate down the cable's length like the incident or reflected waveforms causing it. Note the dots along the line length marking the "zero" points of the standing wave (where the incident and reflected waves cancel each other), and how those points never change position: (Figure 1.4-3)

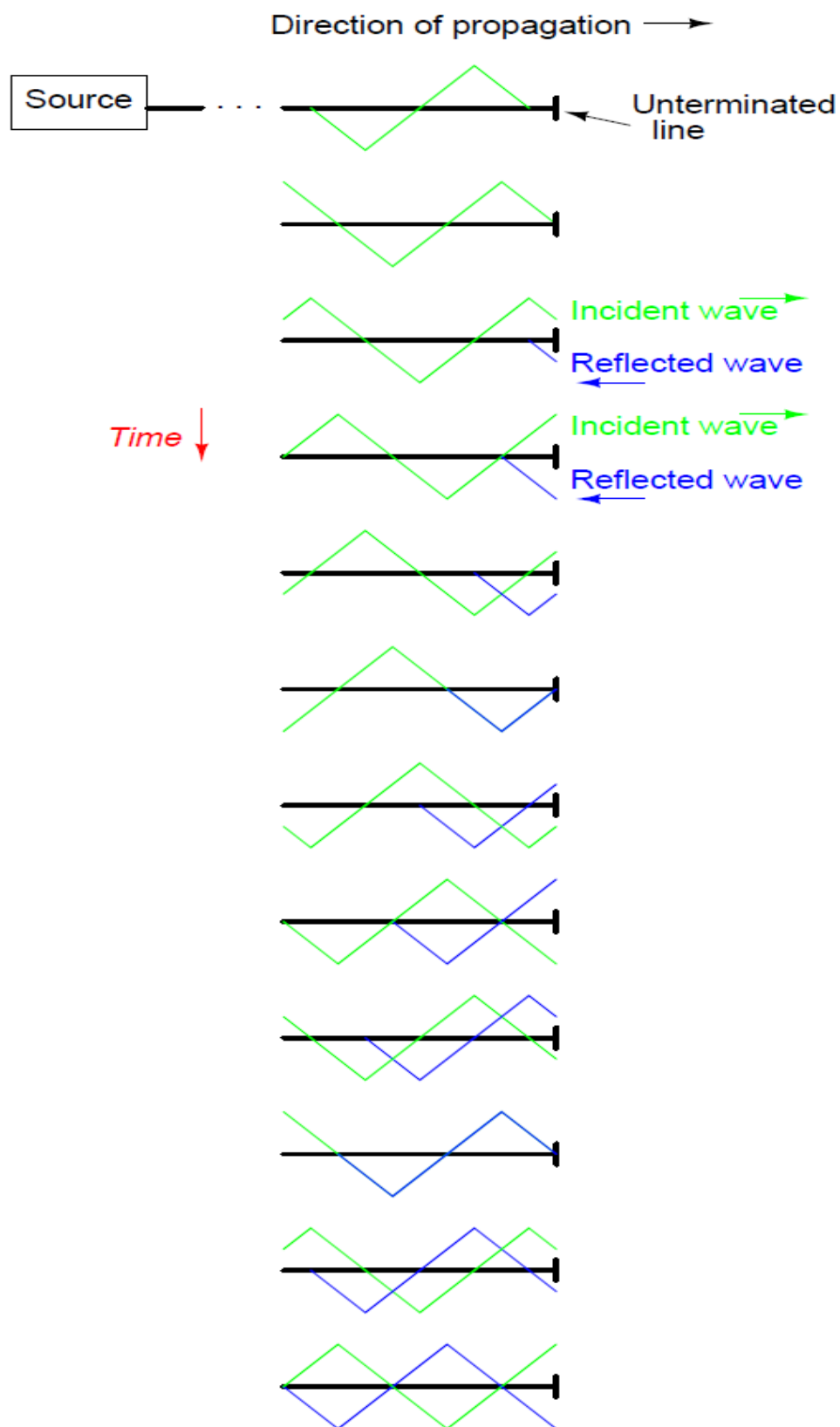


Figure 1.4-1 Incident wave reflects off end of unterminated transmission line

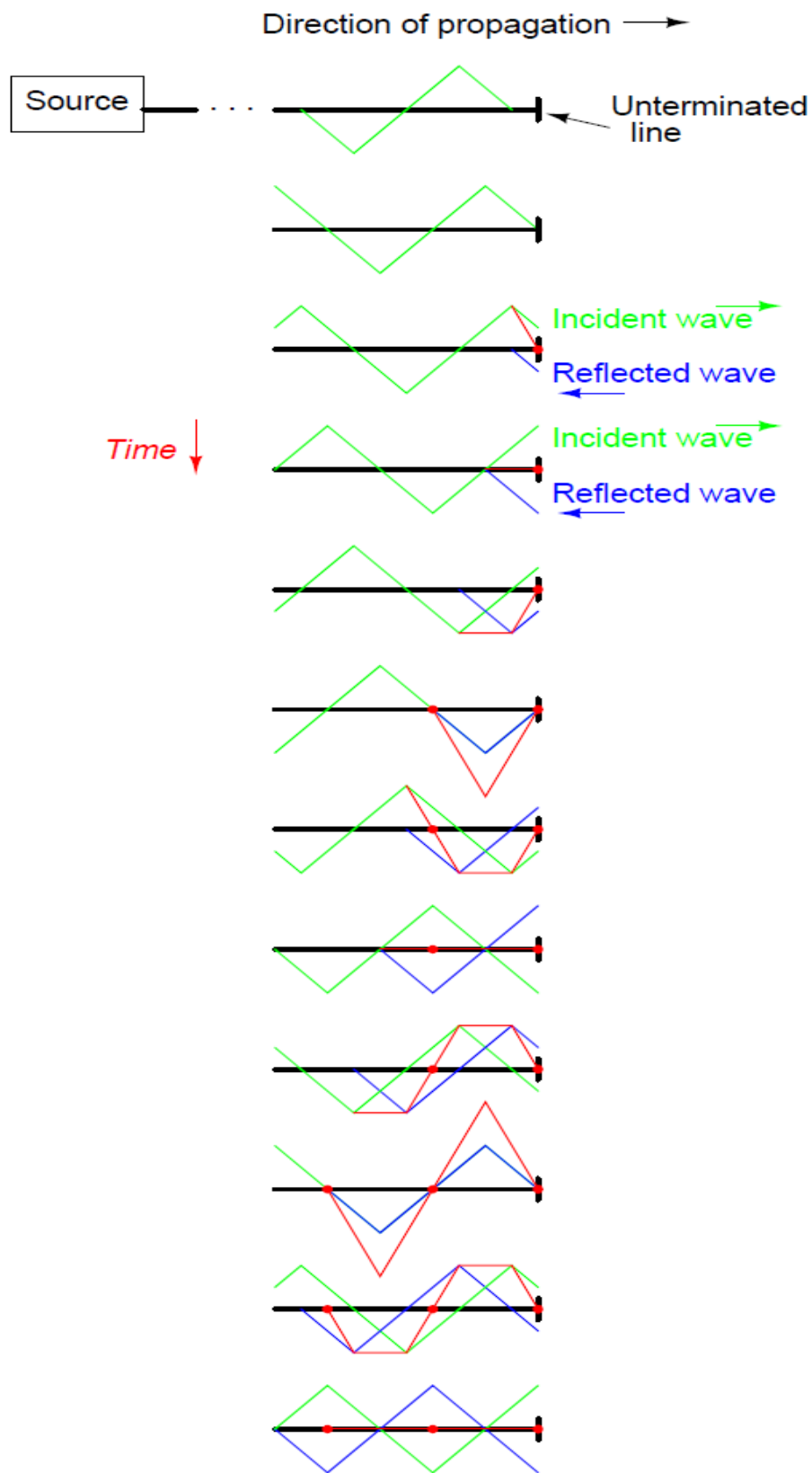


Figure 1.4-2 The sum of the incident and reflected waves is a stationary wave

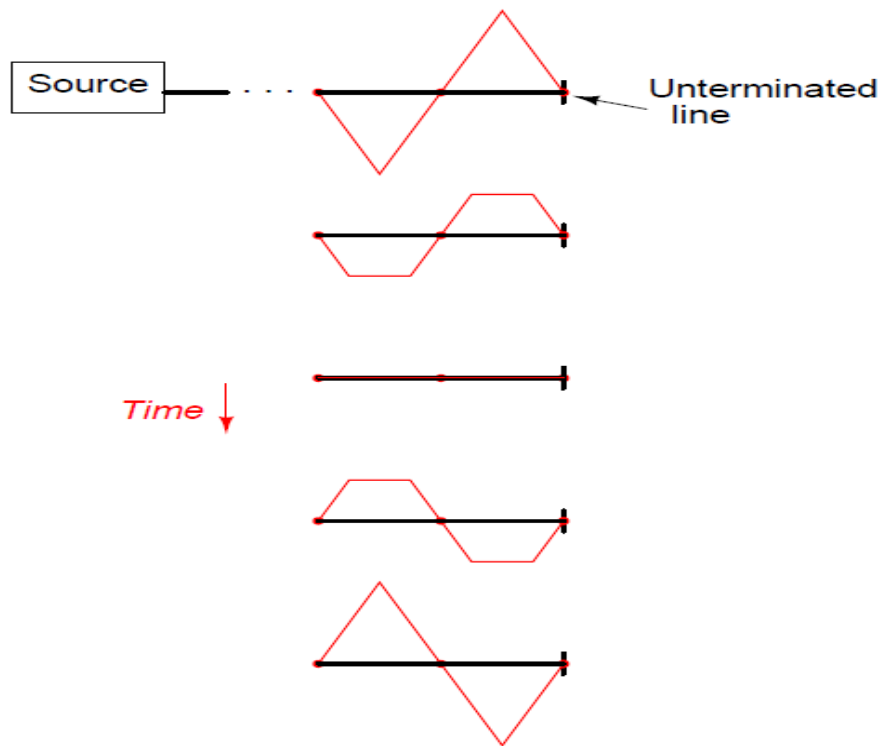
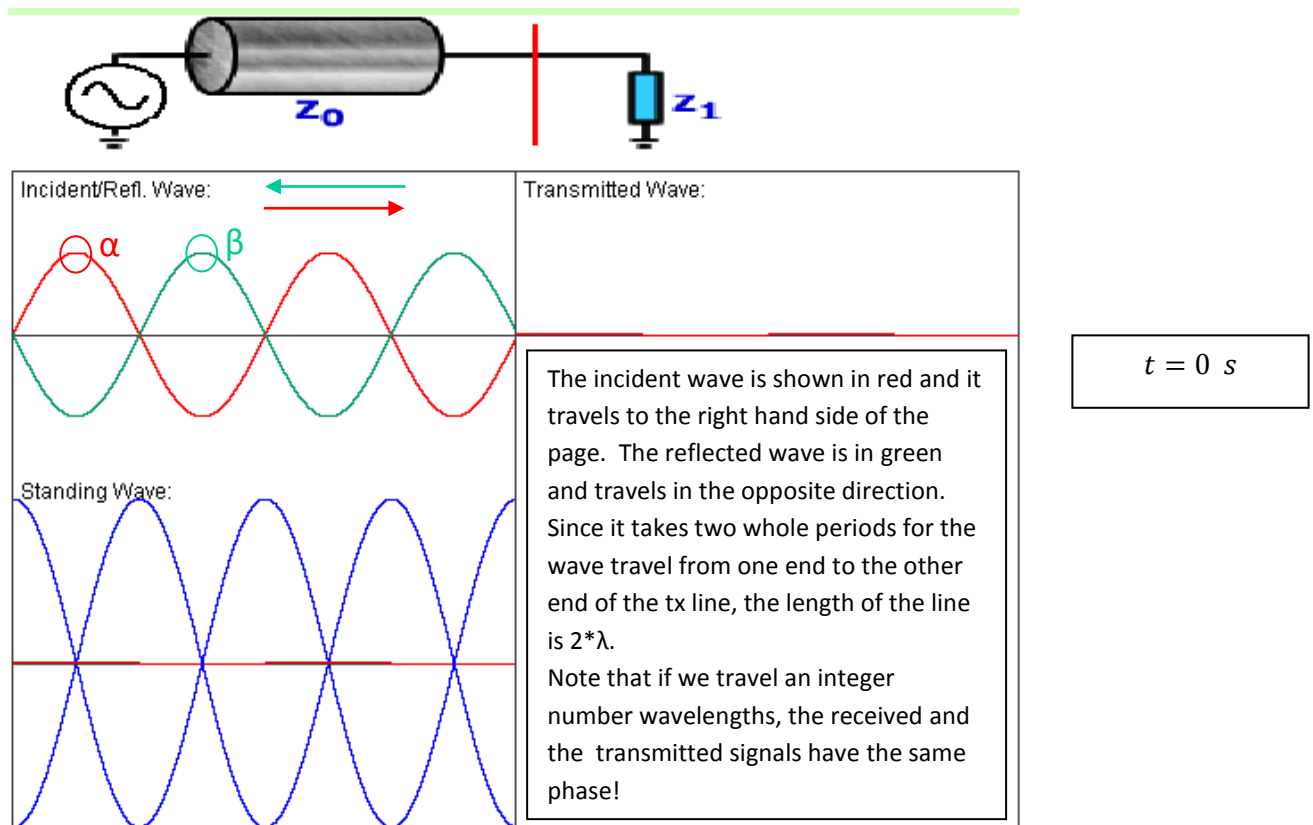


Figure 1.4-3 The standing wave does not propagate along the transmission line

The material presented in this section was extracted from chapter 14 of “volume II – AC” on [allaboutcircuits.com](http://www.allaboutcircuits.com). The document is an excellent introductory tutorial for transmission line theory and is available for free on the web http://www.allaboutcircuits.com/vol_2/chpt_14/index.html

1.5 Reflection and Transmission in an unterminated, shorted, matched and mismatched line

Open circuit



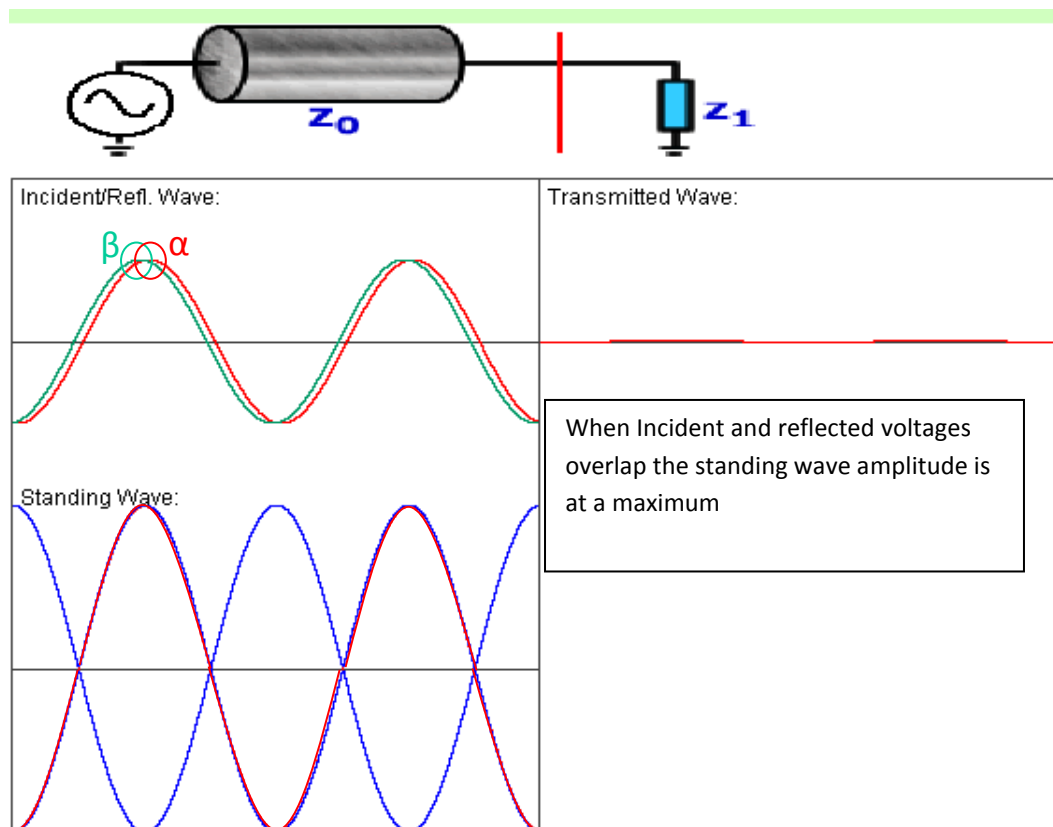
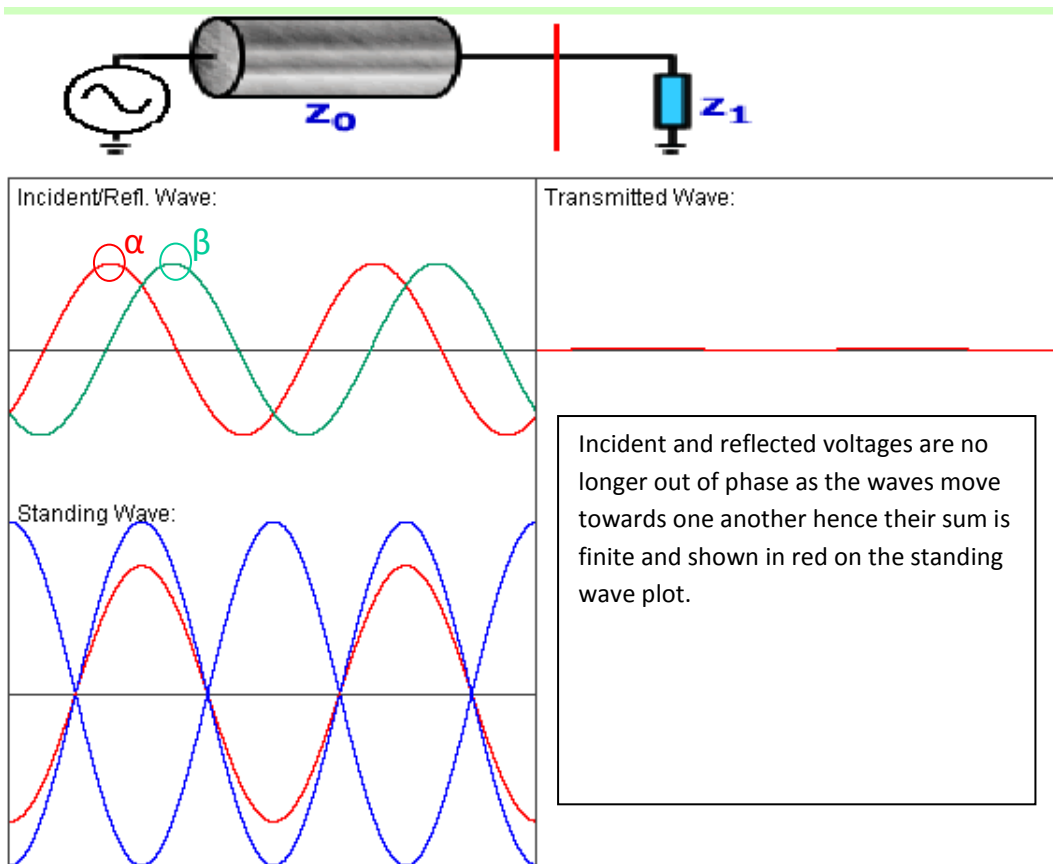
On the Standing Wave plot, the blue curve shows the envelope of the sum of incident and reflected waves, whereas the red curve shows the instantaneous sum of incident and reflected waves. At $t=0\text{s}$, the incident and reflected voltages are 180° out of phase hence their sum is zero!

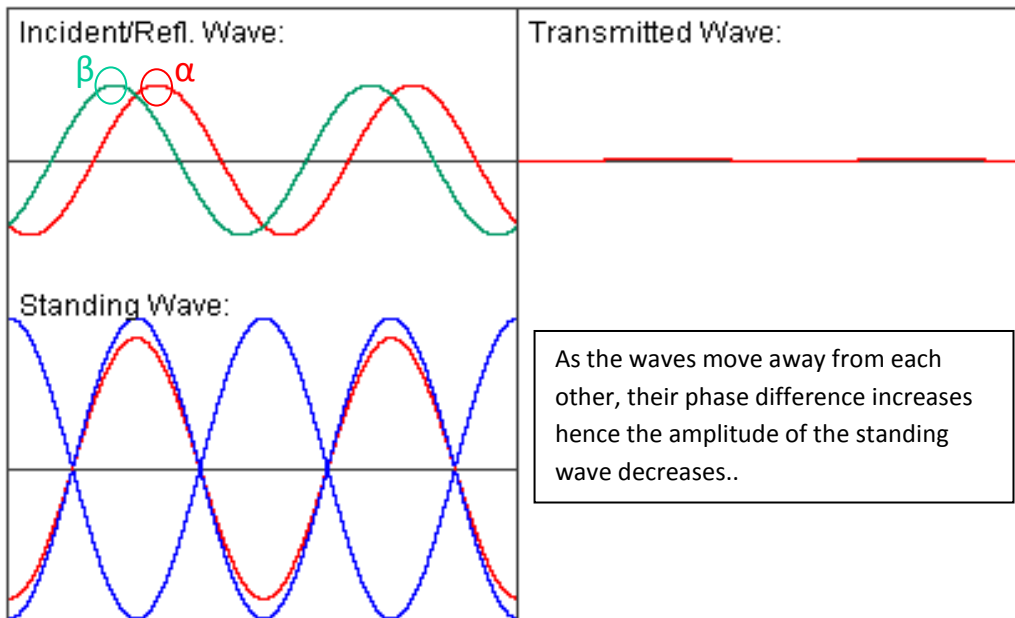
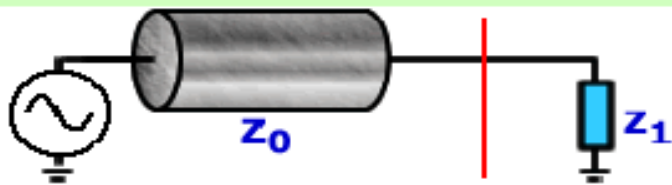
We have picked two points on the transmitted and reflected waves and we will now see how they move along a line terminated in an open circuit.

It will be apparent that, although the location of the maxima and minima of the standing wave do not change along the direction of propagation (hence the "standing" wave definition), the instantaneous amplitude of the standing wave does change!

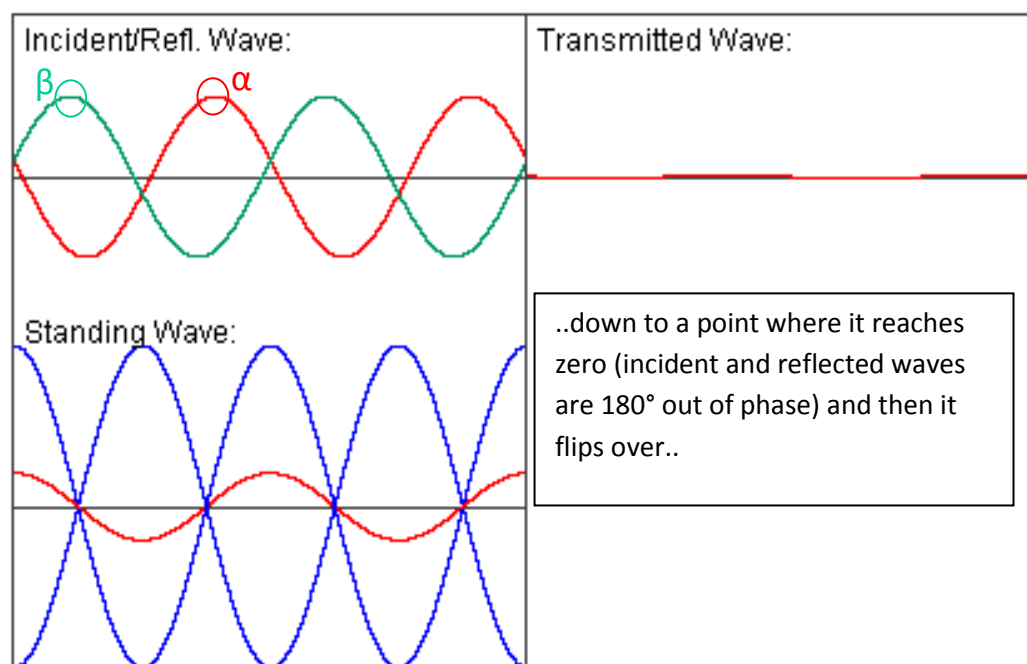
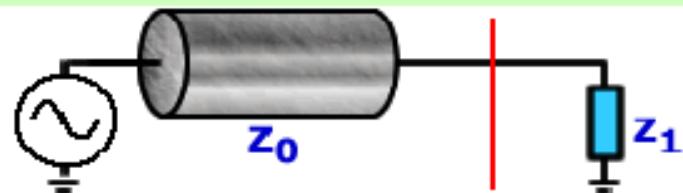
The nulls are spaced at $\lambda/2$ hence, if a line is terminated in an open circuit, one could just measure the distance between two nulls (which is possible with a slotted line) and work out the wavelength (and hence the frequency) of the wave travelling down the transmission line.

As will be seen in the next section, a short circuit will also do the trick. It is important to notice that, although the distance between two peaks could also be measured, it would not be as reliable a measurement, since the standing wave pulsates and hence the peaks are not stable!

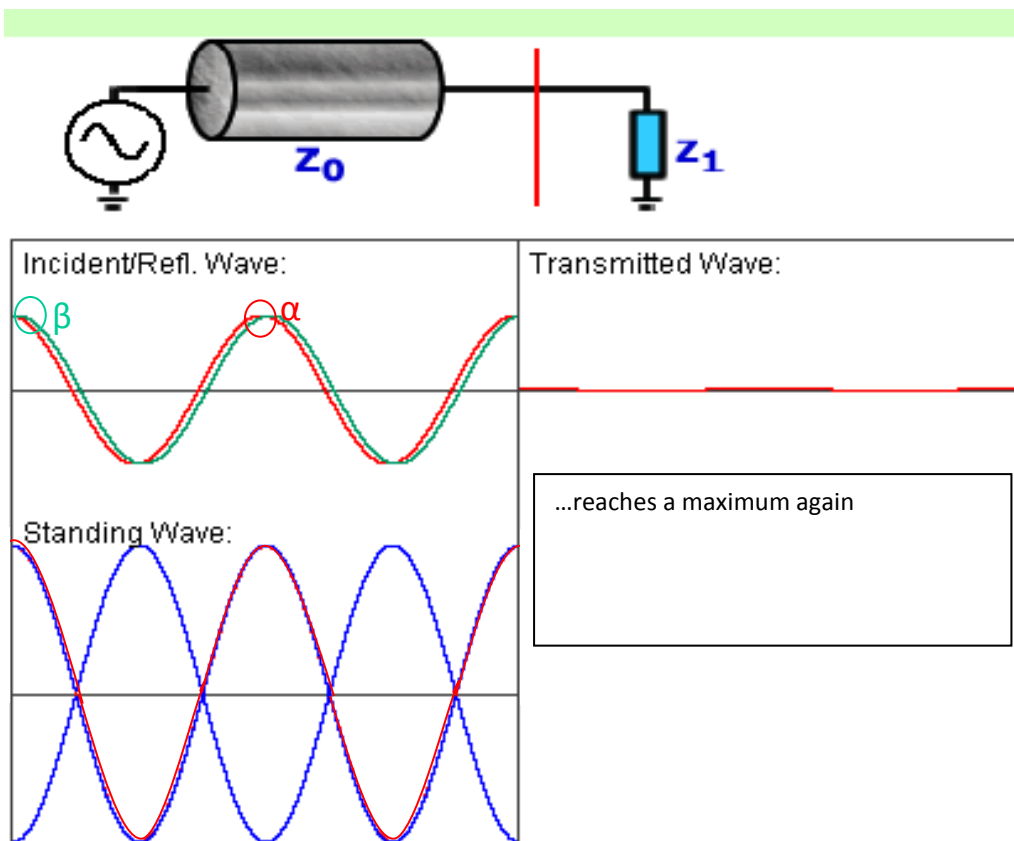
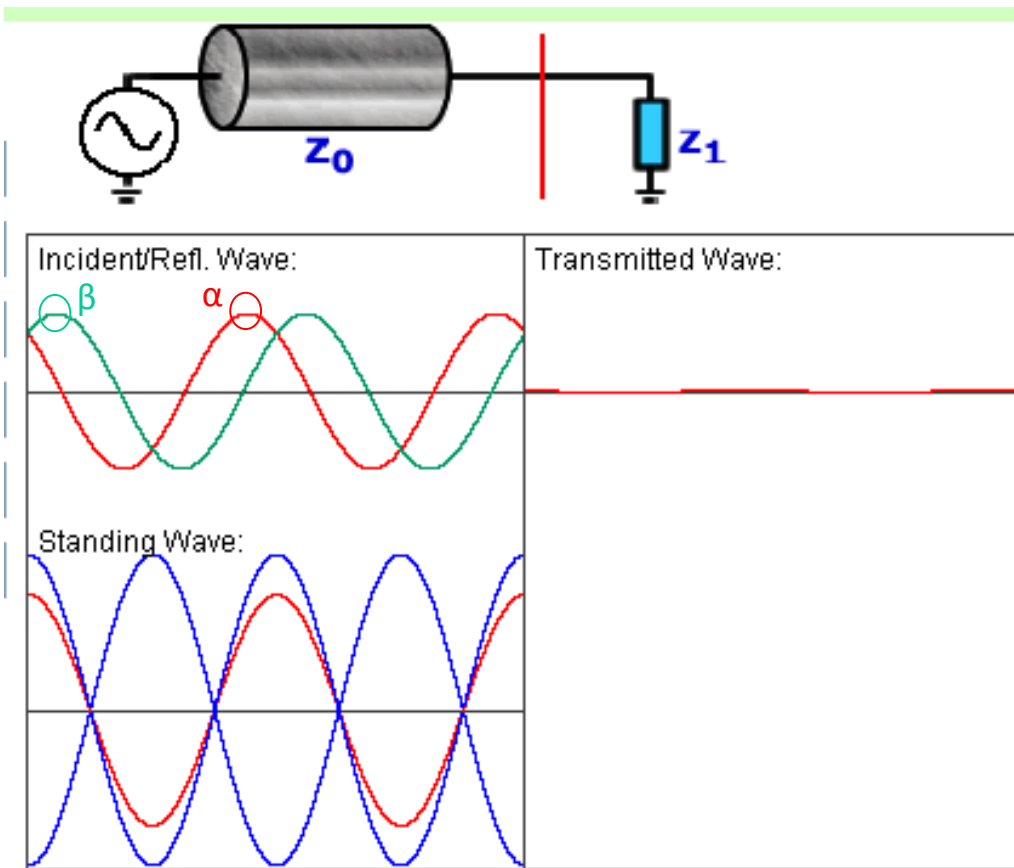


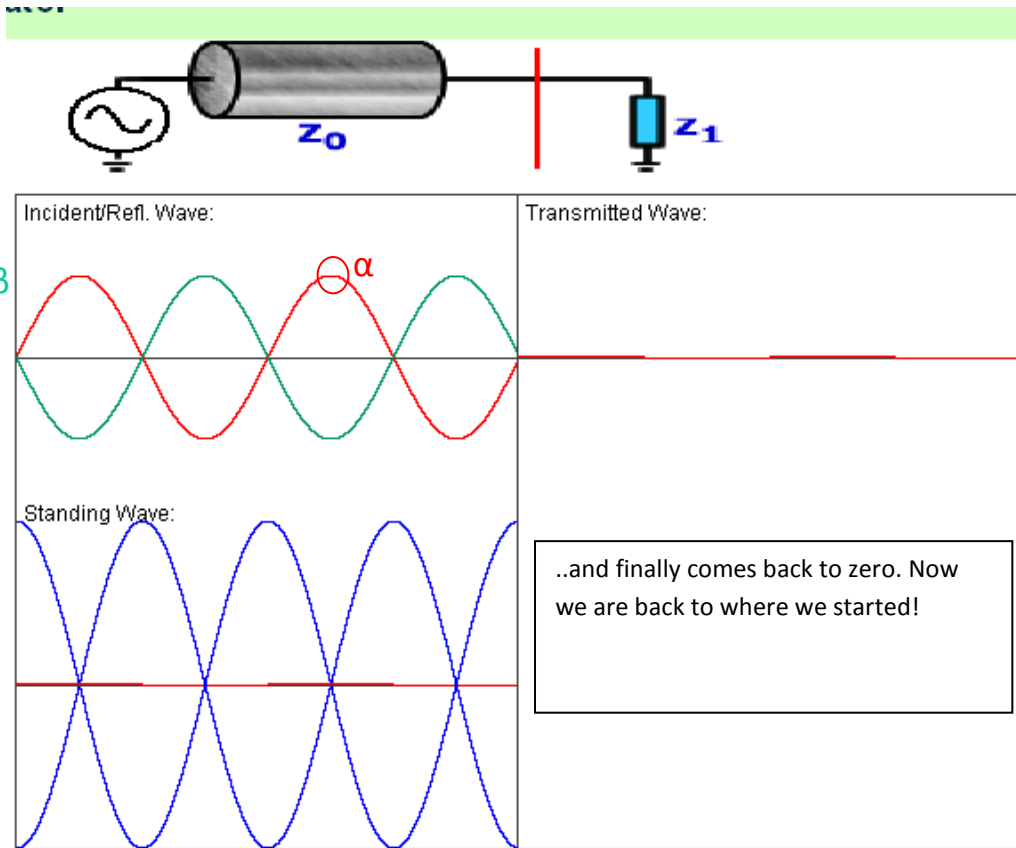


$$t = 0.33T \text{ s}$$



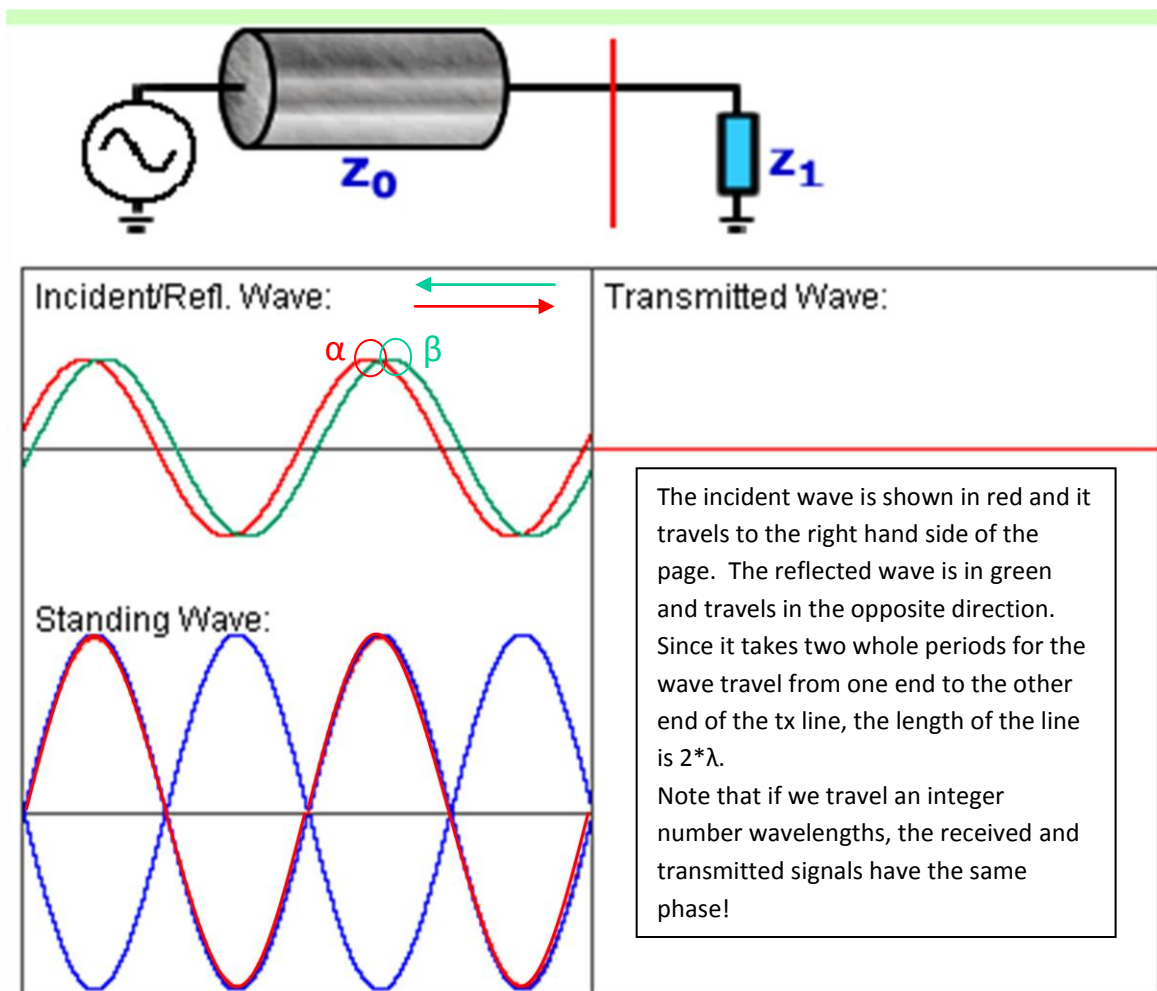
$$t = 0.53T \text{ s}$$





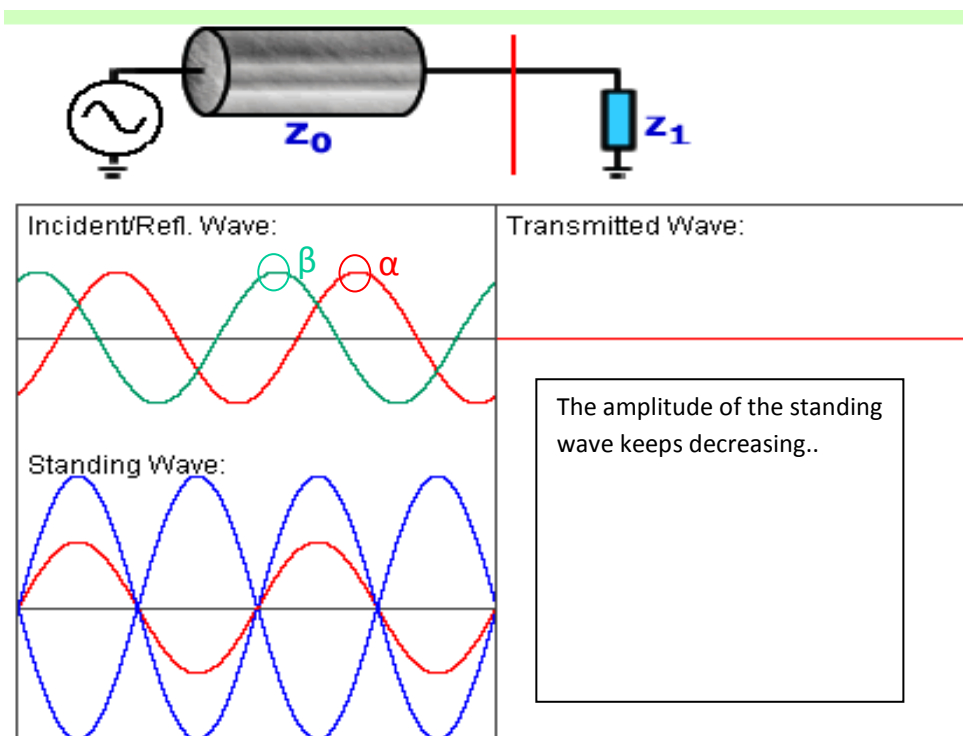
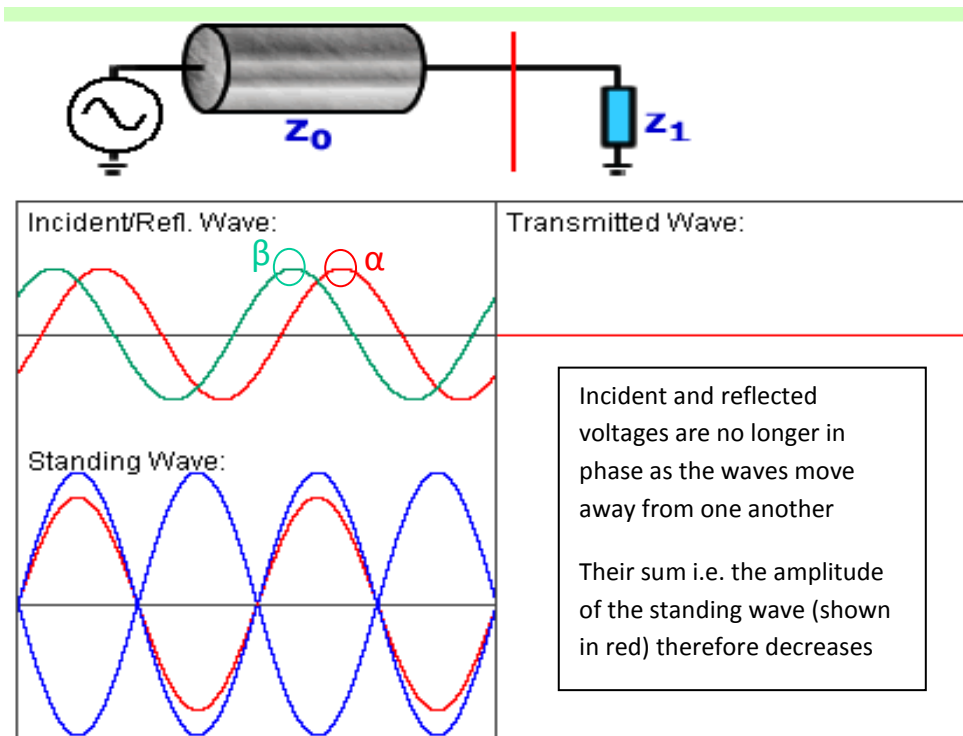
$$t = T s$$

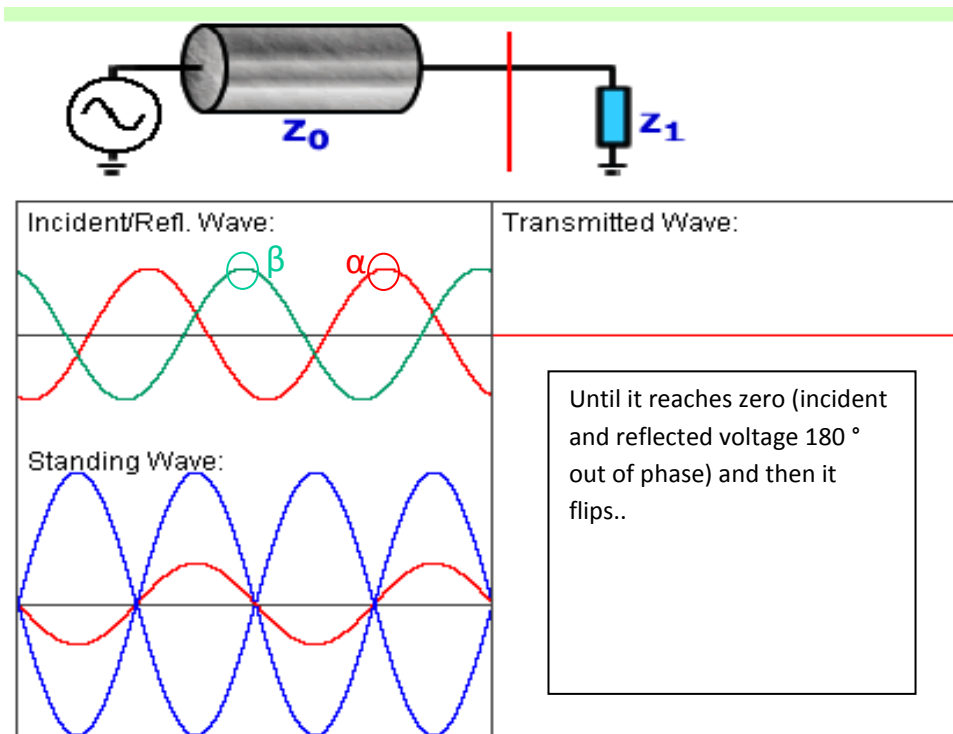
Short Circuit



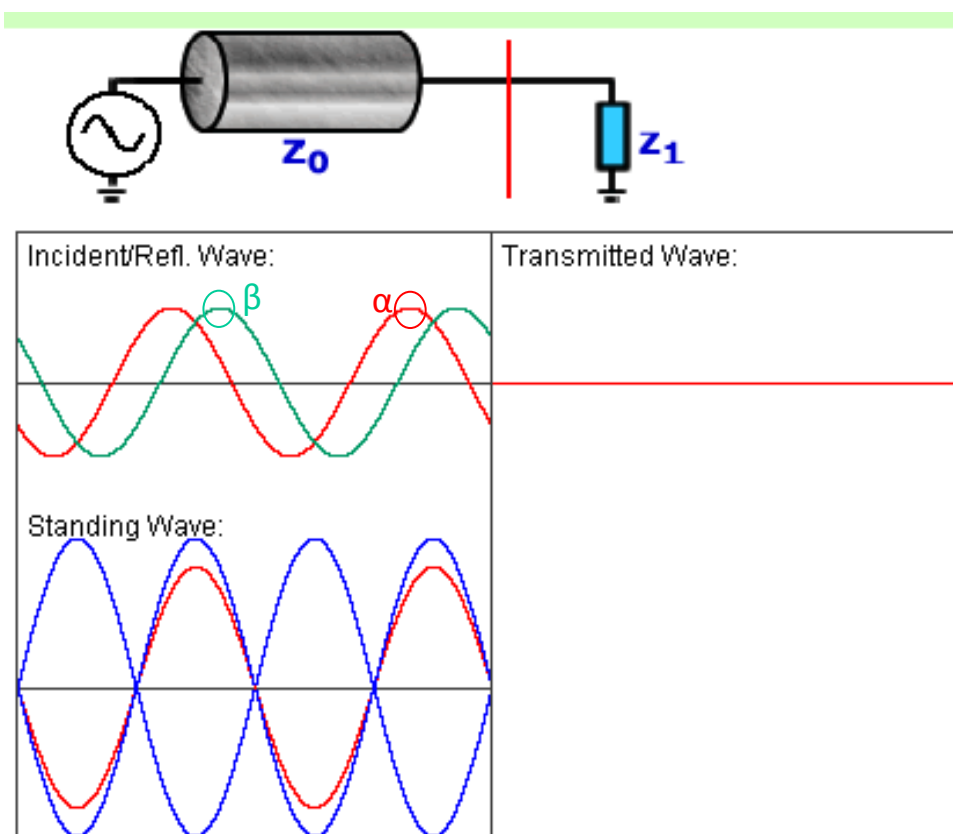
On the Standing Wave plot, the blue curve shows the envelope of the sum of incident and reflected waves, whereas the red curve shows the instantaneous sum of incident and reflected waves. At $t=0$ s, the incident and reflected are in phase hence their sum is maximum!

We have picked two points on the transmitted and reflected waves and we will now see how they move along a line terminated in an short circuit.

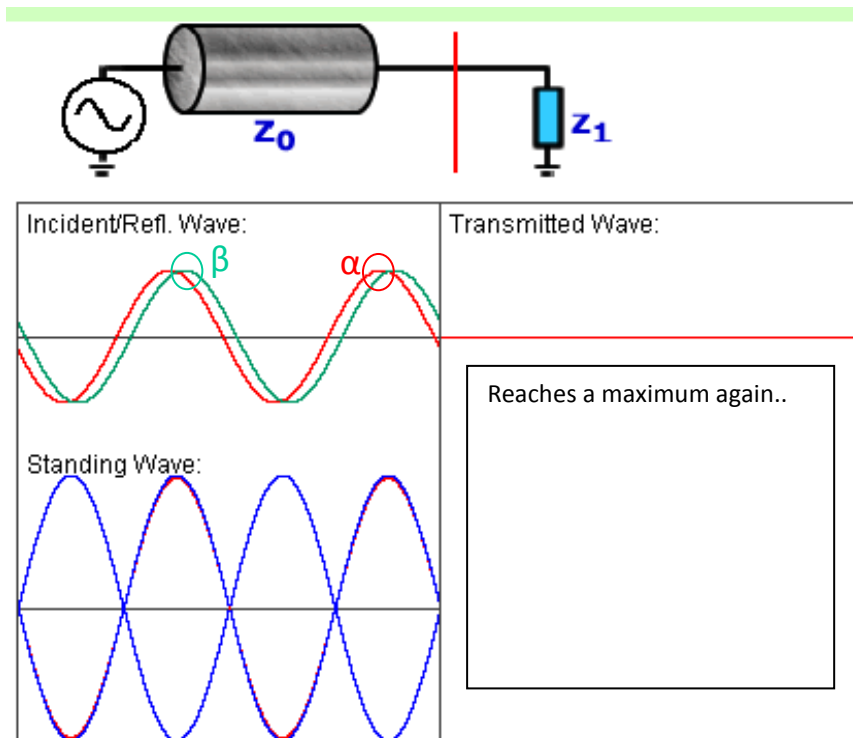




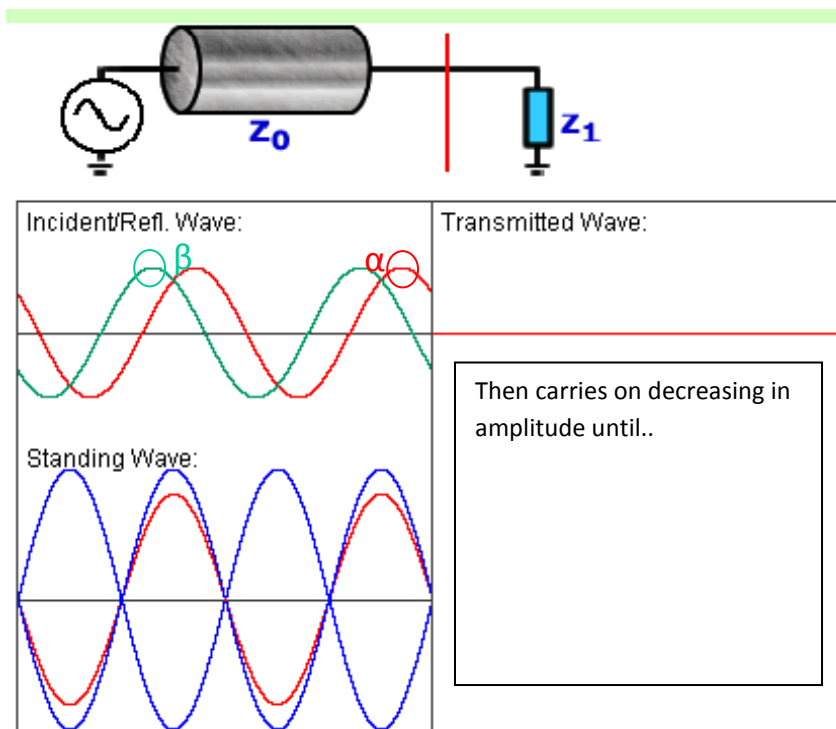
$$t = 0.28 T \text{ s}$$



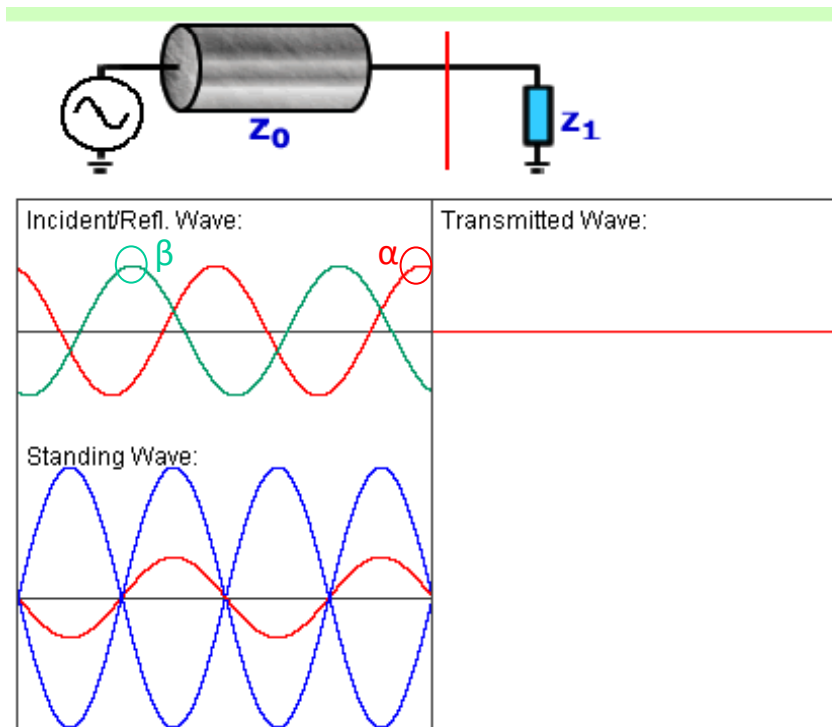
$$t = 0.37 T \text{ s}$$



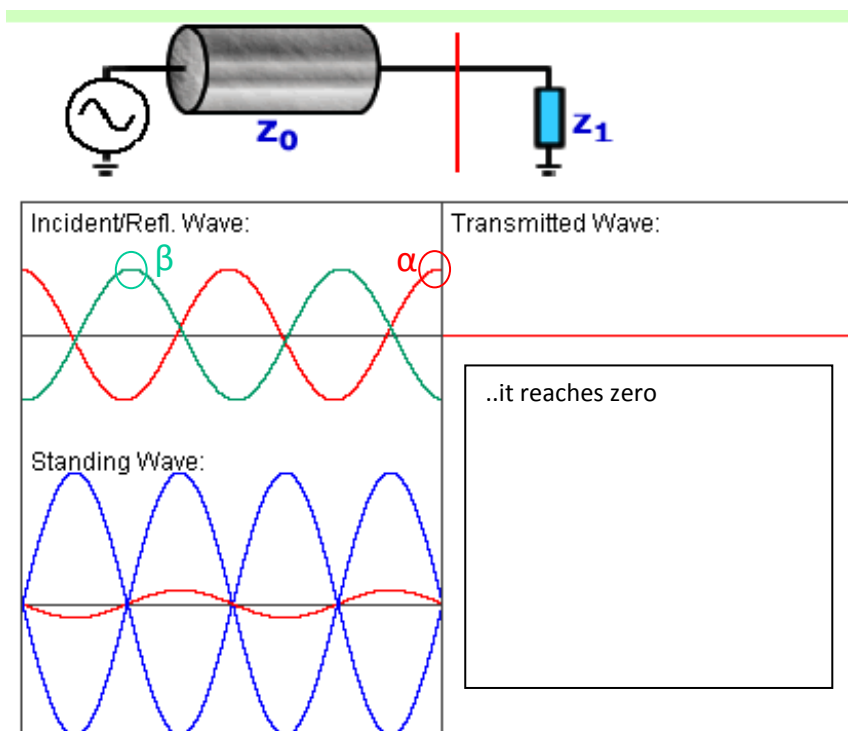
$$t = 0.52 T s$$



$$t = 0.58 T s$$

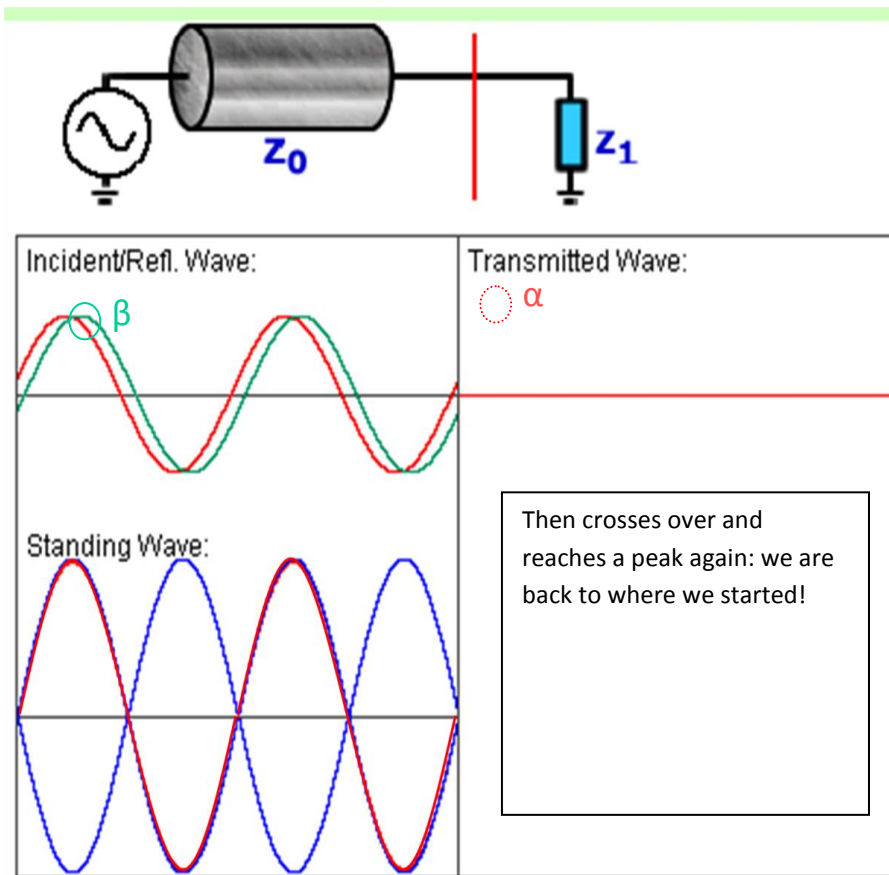


$t = 0.66 T \text{ s}$



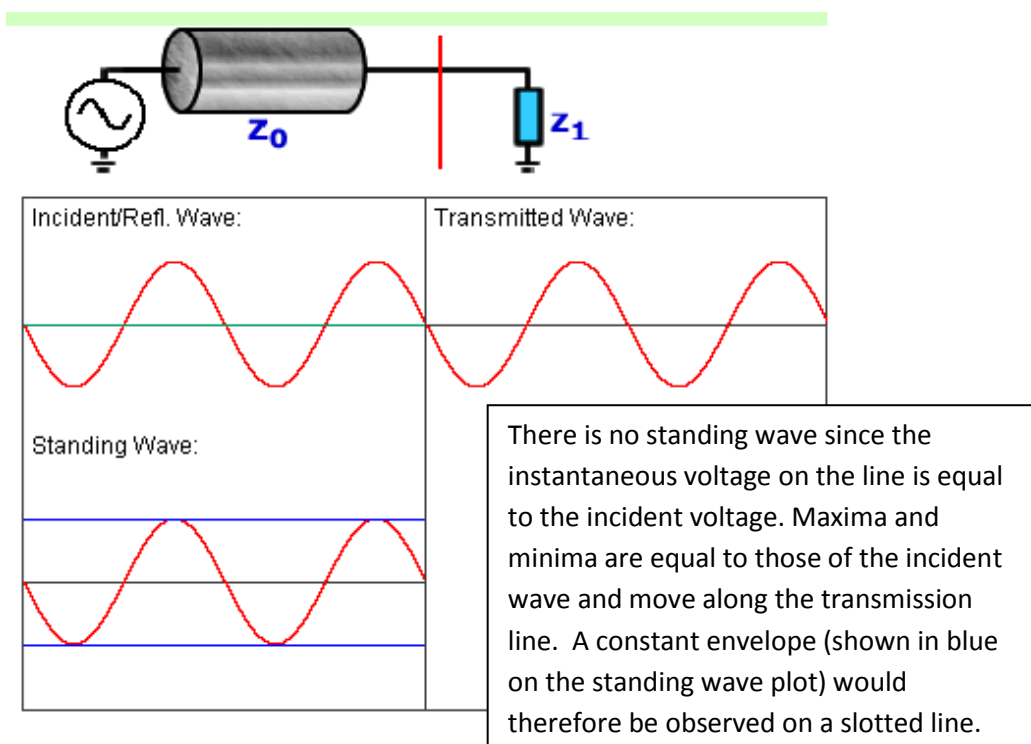
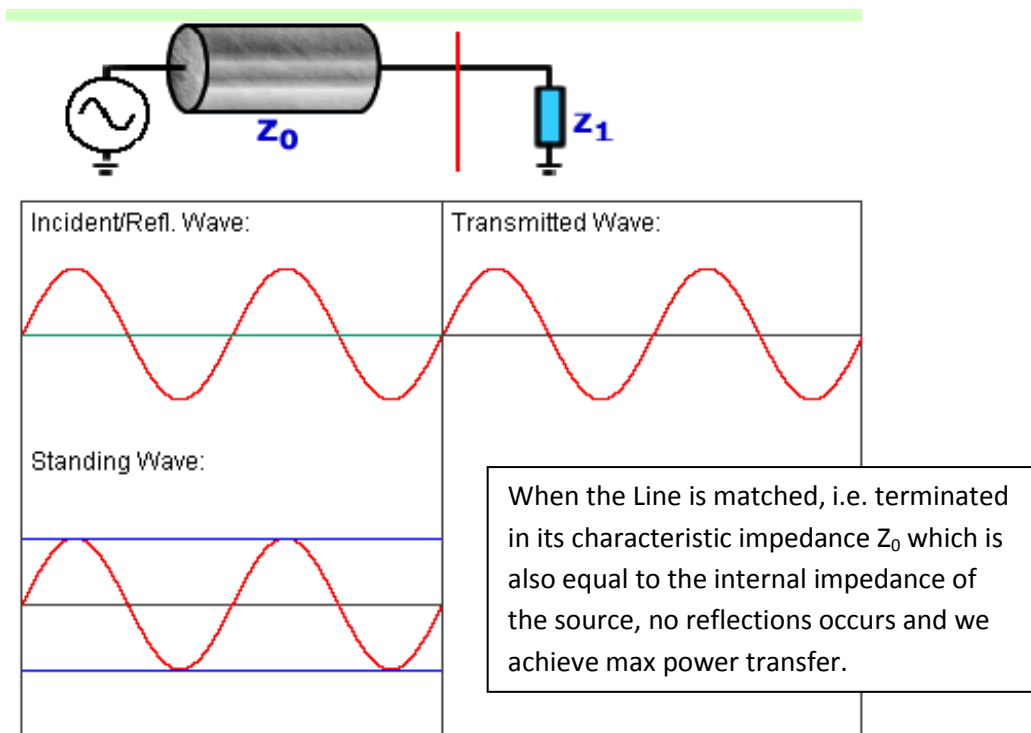
$t = 0.71 T \text{ s}$

Notice that the nulls this time are in a different position to the open-circuited case, although the shape of the standing wave is the same.



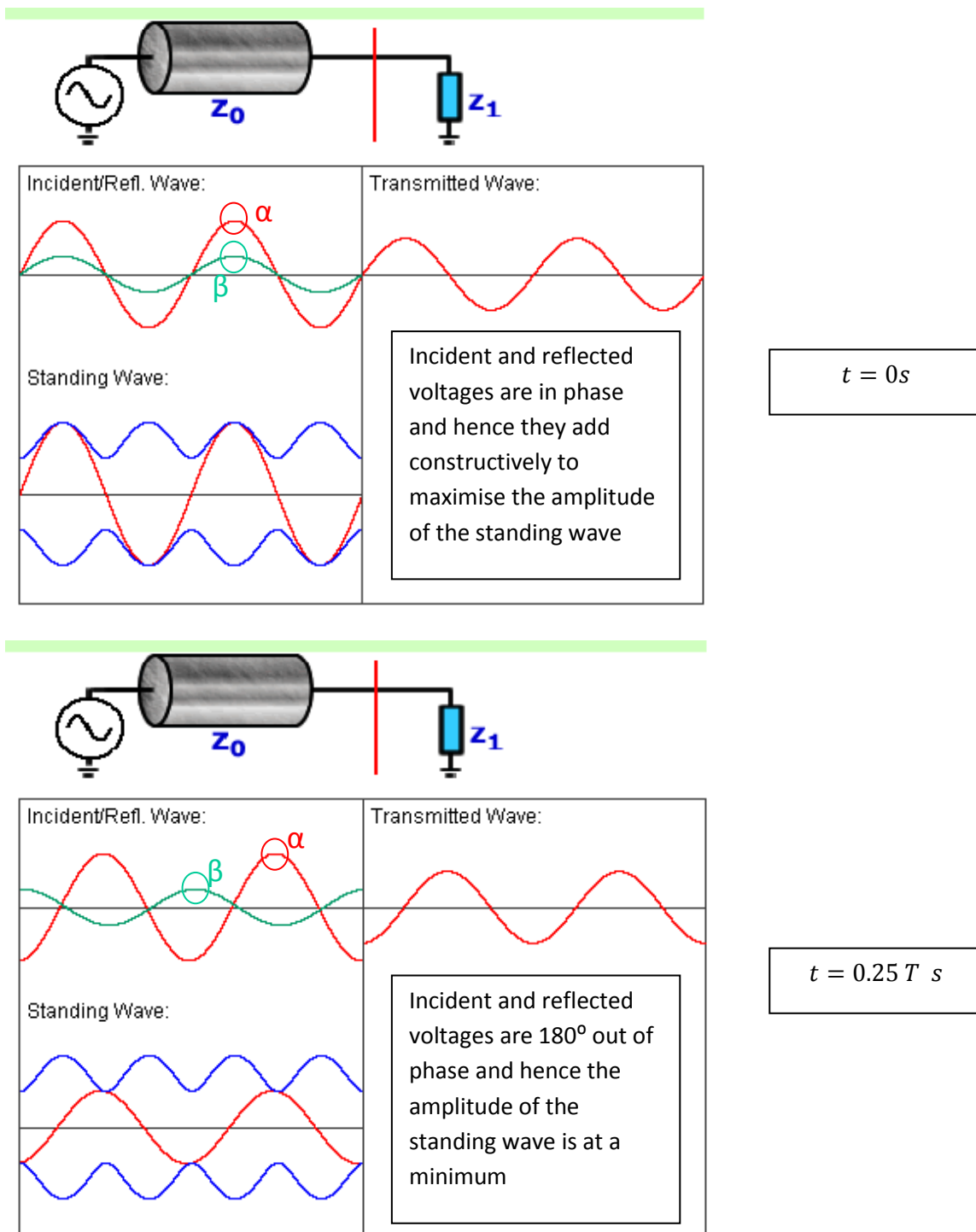
$$t = T s$$

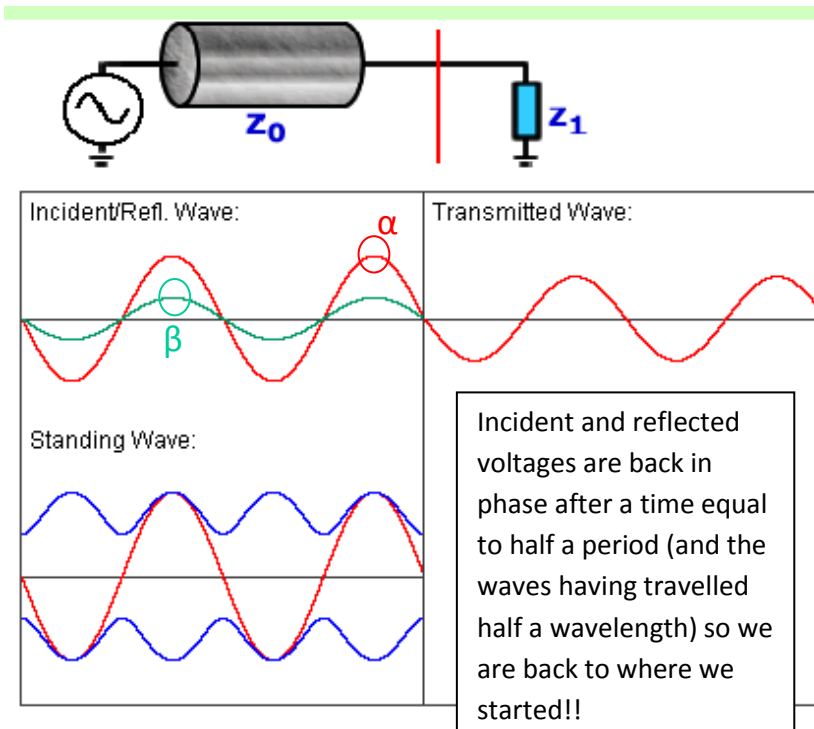
Matched Load ($50\ \Omega$ in this case)



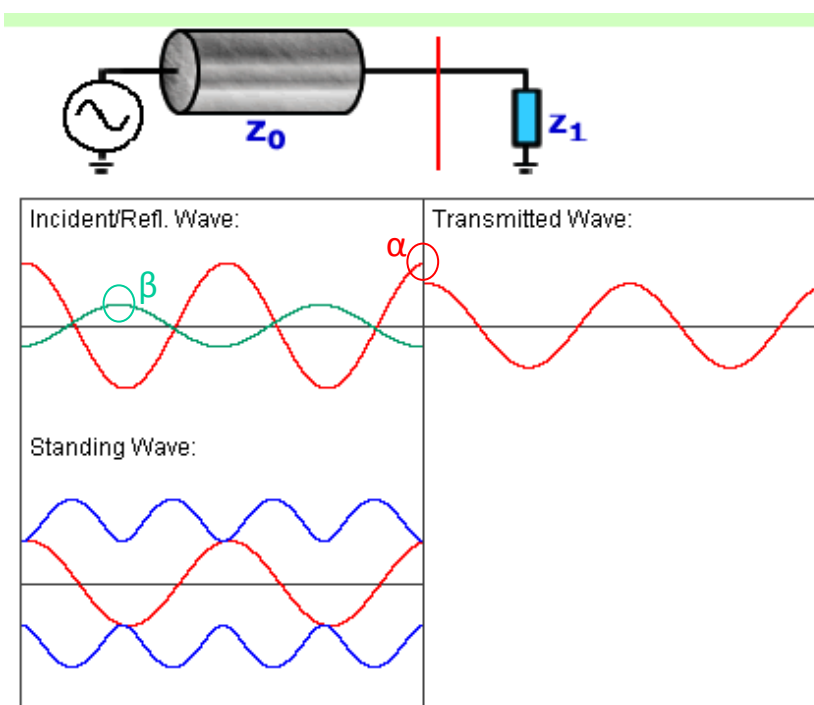
Mismatched $Z_L < Z_0$ (25 Ω in this case)

When $Z_L < Z_0$ some of the power is reflected back to the source and hence a standing wave is created along the line. Note that, in this case, the incident and reflected voltage DO NOT have the same amplitude as in the short and open circuit terminations case. Instead of peaks and nulls we will therefore observe minima and maxima. Part of the wave will however be transmitted to the load and its amplitude will be equal to the amplitude of the incident voltage minus the amplitude of the reflected voltage.

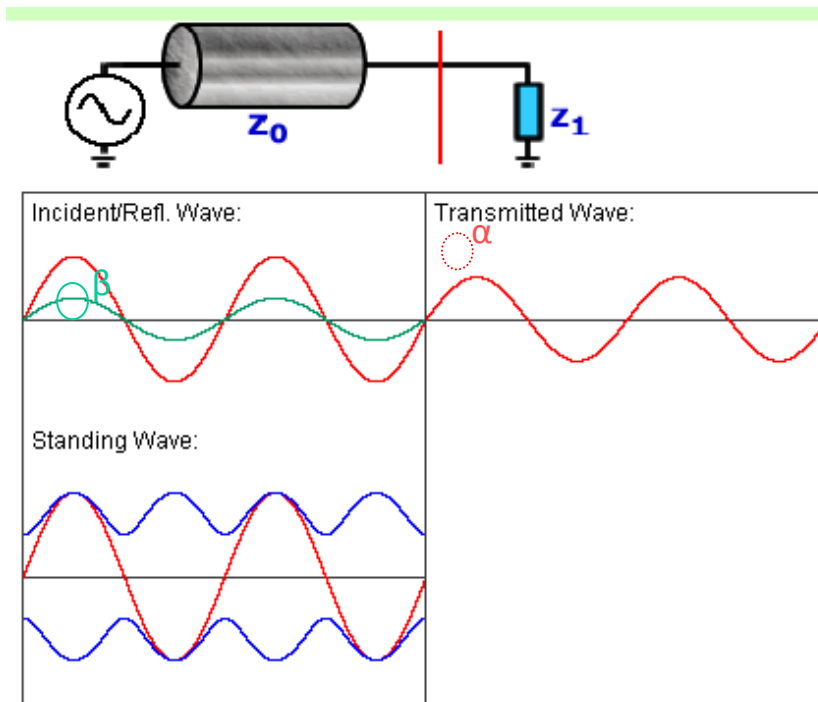




$$t = 0.5 T \text{ s}$$

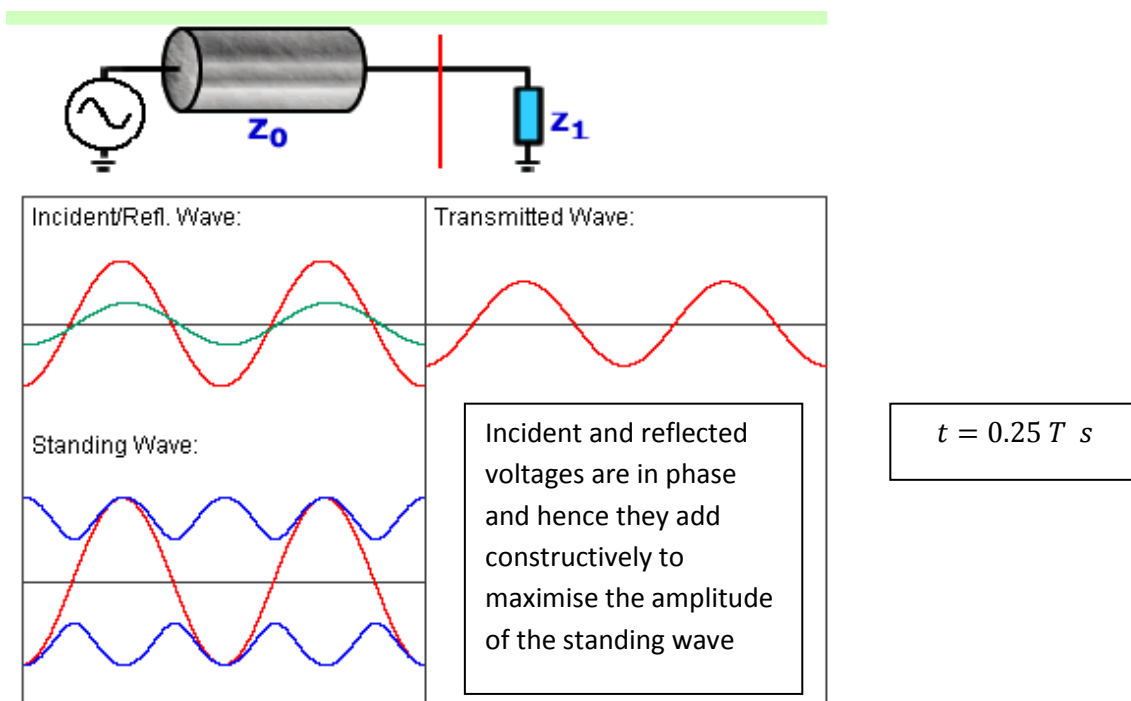
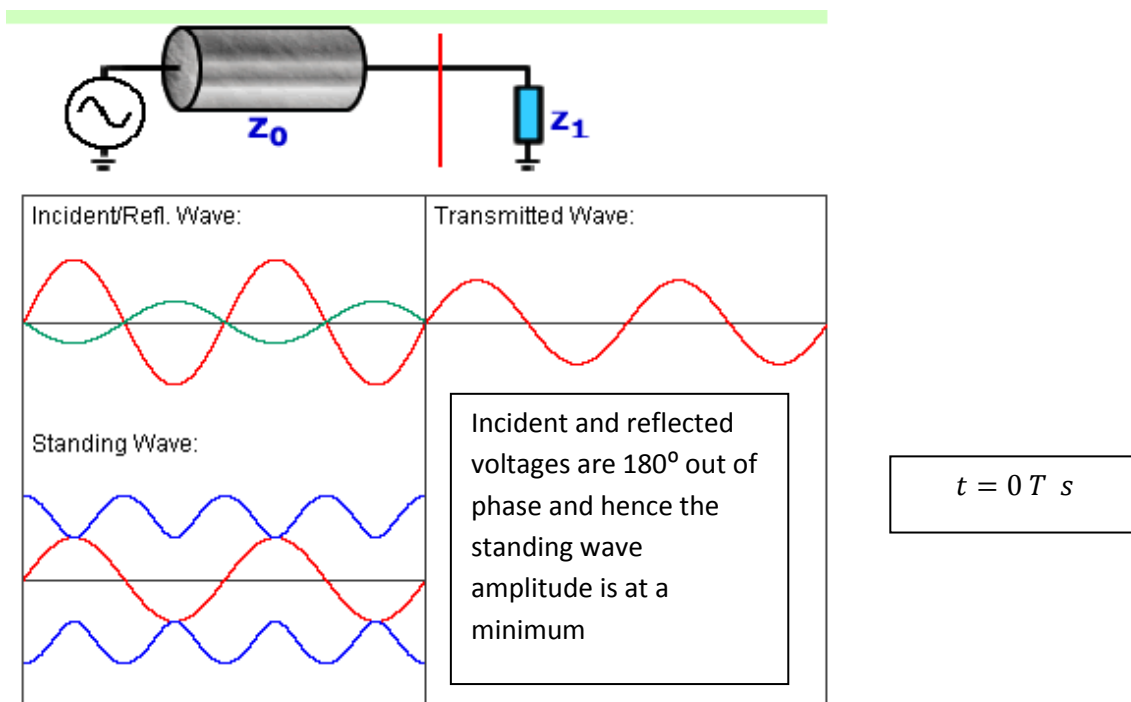


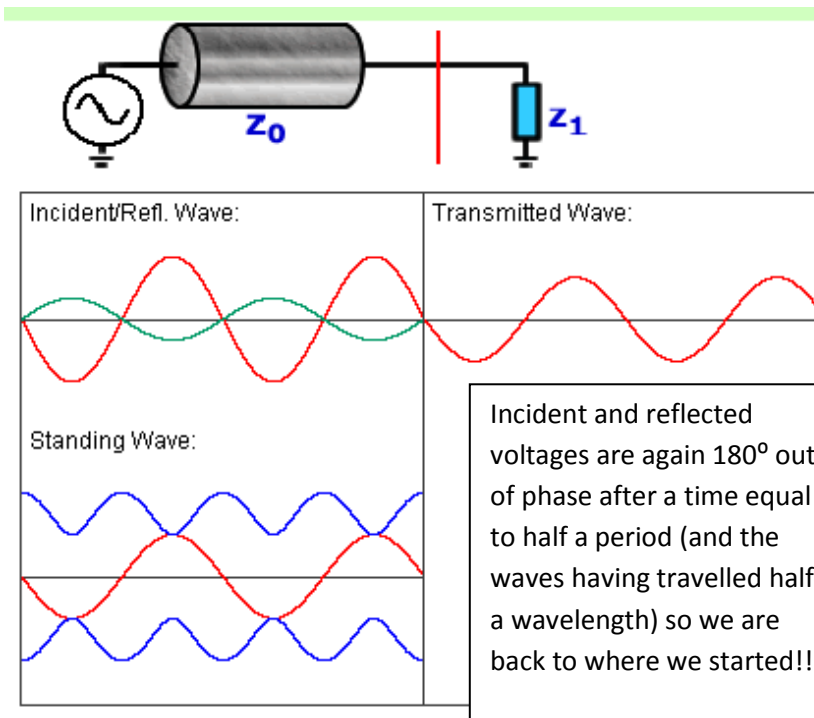
$$t = 0.75 T \text{ s}$$



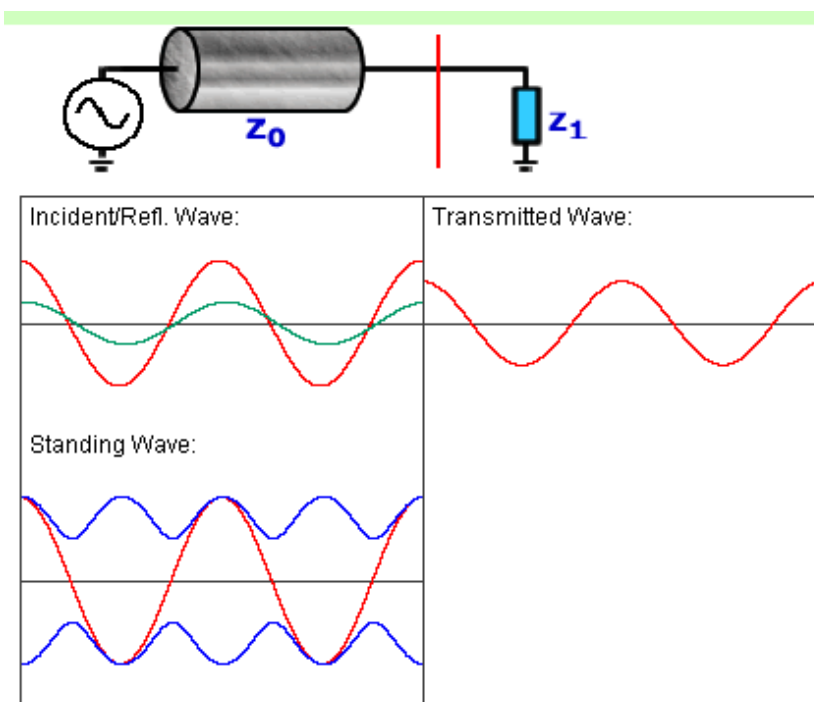
$$t = T s$$

Mismatched $Z_L > Z_0$ (100 Ω in this case)

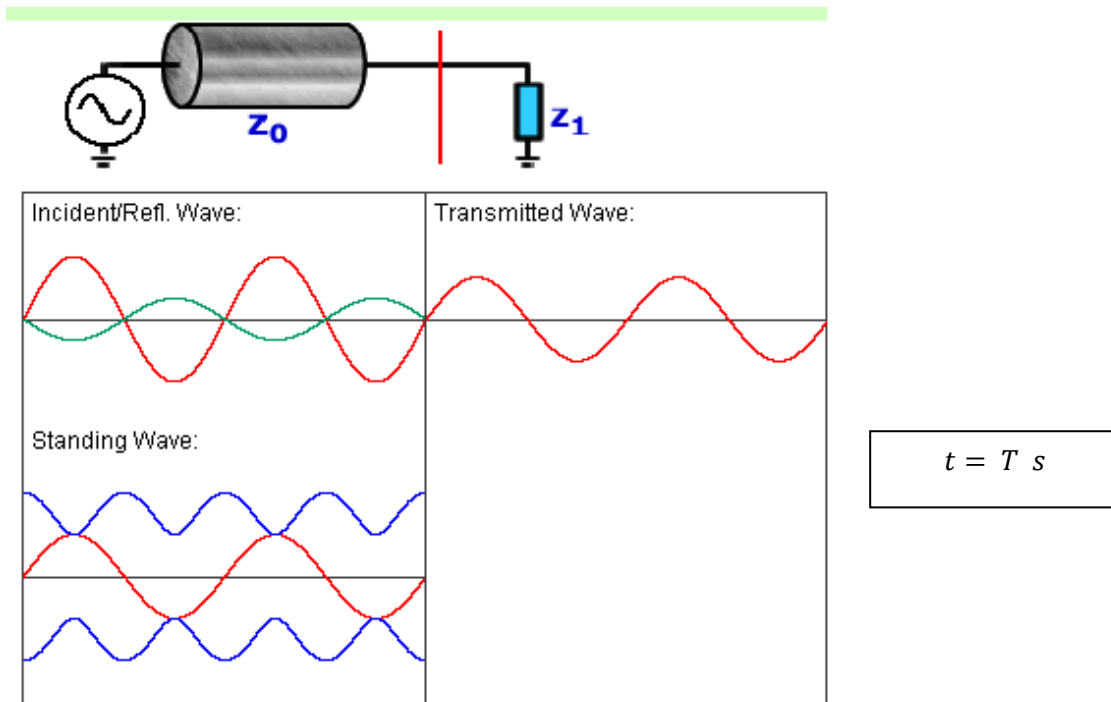




$$t = 0.5 T \text{ s}$$



$$t = 0.75 T \text{ s}$$



As in the case $Z_L < Z_0$, when $Z_L > Z_0$, some of the power is reflected back to the source and hence a standing wave is created along the line. The incident and reflected voltage DO NOT have the same amplitude as in the short and open circuit terminations case. Instead of peaks and nulls we will therefore observe minima and maxima. Part of the wave will however be transmitted to the load and its amplitude will be equal to the amplitude of the incident voltage minus the amplitude of the reflected voltage.

Note that, although the absolute values of minima and maxima are the same for load resistance of 25Ω and 100Ω , and so is the shape of the standing wave, the location of maxima and minima along the line is different!

A useful applet to look at is the Besser Reflectometer which may be found at

<http://www.bessernet.com/Ereflecto/tutorialFrameset.htm>

This applet was used to produce most of the images presented in this section.

References

[1] Allaboutcircuits.com – ‘Volume II – AC’, Chapter 14 – ‘Transmission Lines’

http://www.allaboutcircuits.com/vol_2/chpt_14/index.html

[2] Besser Reflectometer – Bessernet.com

<http://www.bessernet.com/Ereflecto/tutorialFrameset.htm>

Contacts

Dr Francesco Fornetti

francescofornetti@gmail.com

f.fornetti@bris.ac.uk