REFLECTOMETERS - A SIMPLE EXPLANATION OF HOW THEY WORK.   IT IS NOT MAGIC!!

What is SWR? How can you use a reflectomer to measure SWR? What does a Reflectometer actually measure and how does it to it?

Introduction

In early 1986, shortly after passing my extra class exam, I became obsessed with trying to understand how an SWR meter worked. After all, an extra class licensee ought to know how something works that is made of a couple of diodes, resistors and meters. But I didn't. In fact my extra class friends, many of whom are technically quite sharp, could not explain it to me either (maybe I was just too dense). My problem was that everything else in electronic circuitry, including RF, had been explained to me in terms of electron flow. But for some reason, when that RF energy went through the SO-239 at the back of my transmitter, all people wanted to talk about was waves; Forward power waves and Reflected power waves. There had to be more to it.

I need to point out that there is some confusion amongst the amateur community about the terms SWR Meter, SWR Bridge, directional coupler and Reflectometer, me included. The circuit and device to which I am referring is shown in Figure 1. This is the same circuit found in the Heathkit Model HM-11 and which Heathkit calls a “Reflected Power Meter and SWR Bridge” on the cover of its manual. So who am I to argue with Heathkit. The important point is that I am talking about devices which have the simple circuitry shown in Figure 1 and which have been used by amateur radio operators for decades to measure SWR, ever since it was first described by Norgorden[6].

I have been told that SWR is a ratio of the highest and lowest voltages along a transmission line.[1] How do you measure those voltages without moving the SWR meter? It has been said that if there is a mismatch at the antenna, then a portion of the power sent forward from the transmitter to the antenna will be reflected back from the antenna toward the transmitter [2]. SWR is a measure of the relative amounts of forward and reflected power. [3]   The more modern equipment has automatic shut down circuits to protect the transmitter, so you need a low SWR to get the most out of your system. Therefore, it is important to know the SWR of your antenna system, although a low SWR at the transmitter alone is not enough to ensure that you have a good antenna system.

OK, back to the question; what does your SWR meter measure and how does it do it? It measures forward power and reflected power. In my research to try to understand how the SWR meter works I have looked through a number of textbooks and journal articles. Some of them ask you to consider a forward power pulse and a reflected power pulse. A complex ratio of these power pulses gives you an SWR. A complex ratio of the voltages of these power pulses gives you a VSWR. But how do you measure that power or those voltages when both are present on the line at the same time?

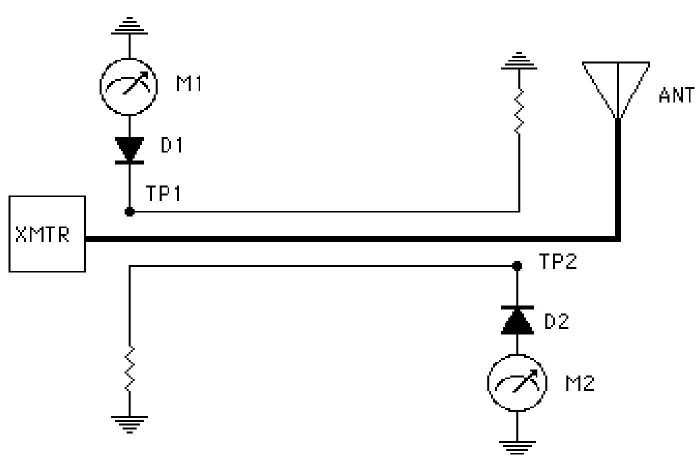
In his 1959 QST article[4], Bruene describes the inner workings of directional couplers.[5] I believe the description is accurate and I quote one key paragraph:

"The directional coupler can sense either the forward or reflected component by taking advantage of the fact that the reflected components of voltage and current are 180 degrees out of phase while the forward components are in phase. A small voltage derived from the current in the line is added to a sample of the voltage across the line. If these two samples have the right amplitude relationship, the two reflected components cancel. The sum then represents only the forward component. By reversing the phase of the current sample 180 degrees, the forward components cancel and the result is the sum of only the reflected components."

When I first read this I couldn't wrap my brain around the concept that the voltage and current in the reflected wave were 180 degrees out of phase. When I learned about voltage and current I was told about the analogy with water flow in a pipe, where voltage was the water pressure and current was the flow in the pipe. I can handle the fact that current and voltage are 90 degrees out of phase in capacitors and inductors, but if the current and voltage are 180 degrees out of phase, does that mean that the water in the pipe is flowing upstream against the pressure? We will come back to this.

Well let's take a look at a typical SWR meter as depicted in Figure 1.

Figure 1. A Simple SWR meter.



In this circuit the diodes D1 and D2 are the "detectors". The lines running parallel to the transmission line and connected to the diodes are the detector loops which are often called sampling loops. Now what I don't understand is how a diode "detects" a power pulse. The last time I looked diodes only did one thing (except for Zeners). Diodes conduct current in one direction and not in the other. And even if the diode could "detect" a power pulse, how does the diode in the "forward leg" know which is a forward pulse and which is a reflected pulse?

In this paper I will present a qualitative explanation of how an SWR meter (reflectometer) works. You will see how each diode "knows" which wave is the forward power and which is the reflected power. In fact you will see why the orientation of the diodes is irrelevant. In fact the diode is only there to make the meter work. You could replace the diode and meter with a light bulb and get the same results by measuring the bulb's brightness. It will also become obvious which leg measures the forward and which measures the reflected power. You may even be able to wrap your brain around the concept that voltage and current can be 180 degrees out of phase!

Before I get to the nitty-gritty I need to convince you of one simple principle and remind you of two others that you already know about.

Wave Properties

The first concept deals with the nature of the WAVE that carries the power from your transmitter to your antenna. Consider first how a sound wave travels through air (Figures 2A and 2B). Figure 2A shows

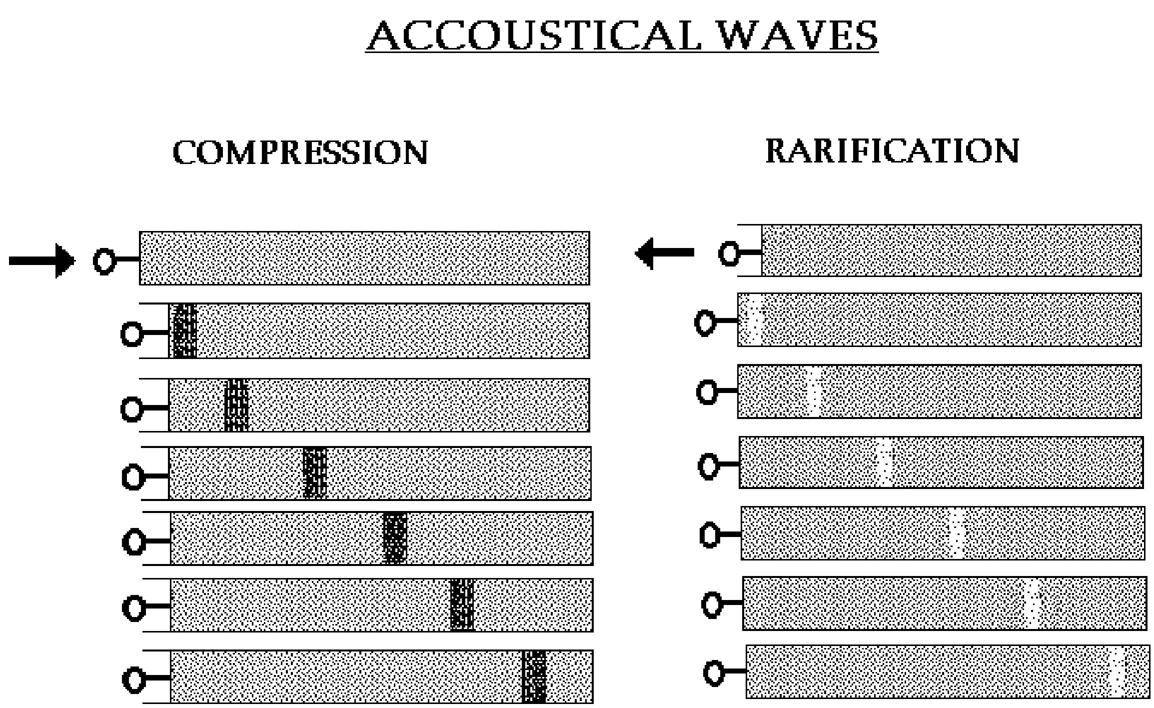


Figure 2A **Figure 2B**

how, as the diaphragm at the left is pushed quickly to the right, the molecules of air are compressed. Then that compressed "wave" moves to the right carrying the energy with it. In the second half of the wave, the diaphragm at the left in Figure 2B is quickly moved to the left. This results in a decompression, or rarefication, of the air molecules behind the diaphragm. This half of the energy carrying wave also moves to the right away from the source (the diaphragm). Notice that in figure 2A, the medium (air molecules) which carry the wave move in the same direction as the wave. However, in Figure 2B, the medium and the wave move in opposite directions; the wave to the right and the medium to the left.

- **When the wave and the medium move in the same direction (Figure 2A) there is a compression of the medium in the wave.**
- **When the medium and the wave move in opposite directions, there is a rarification of the medium in the wave**.

Also notice that the individual air molecules only move a short distance. In fact, after the second half of the wave, they are just about back in their original positions.

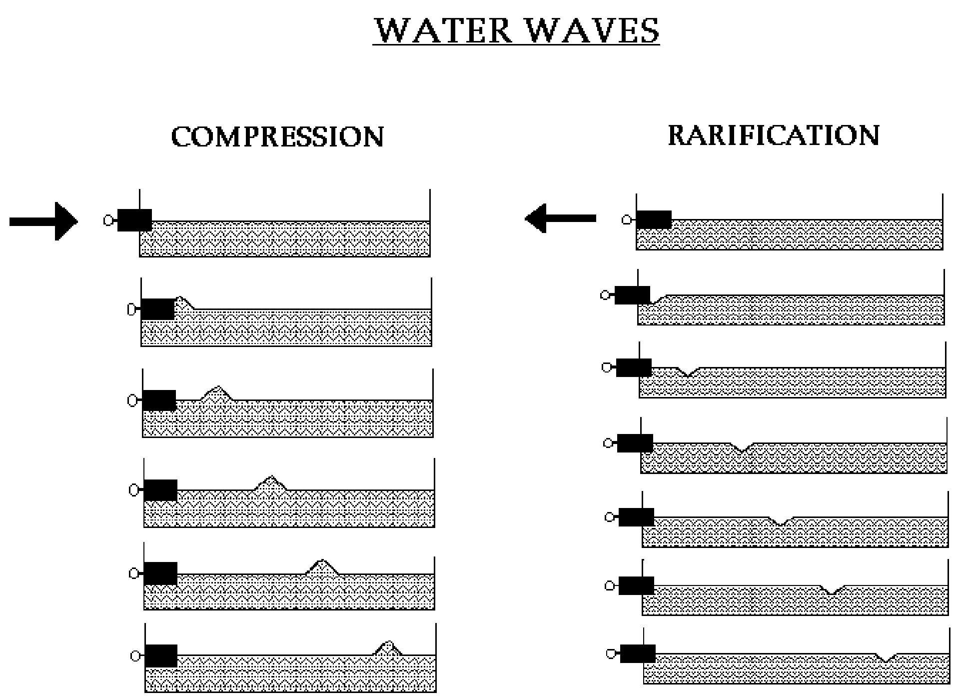


Figure 3A. Compressed Water Wave Figure 3B. Rarified Water Wave

A similar phenomenon is observed with wave propagation in water, as illustrated with the water tanks in Figures 3A and 3B. In figures 3A and 3B the water tanks are fitted with a piston on the upper left hand corner. When the piston is pushed to the right (Figure 3A), the energy is transferred to the water in the form of a wave, which moves to the right. Notice that the wave and the medium (water molecules) which carries the wave move in the same direction. This results in a wave with an excess of the medium. Since water can't be compressed, the excess water rises up to form the wave we are accustomed to seeing.

In Figure 3B, the piston is pulled back to the left to form the second half of the wave. This leaves a void which is filled in with water molecules moving to the left. The energy carrying wave, however, still moves to the right. Again, when the wave and the medium carrying the wave move in opposite directions, a deficiency of carrier particles (water molecules in this case) is generated in the wave.

This is a general phenomenon associated with this type of wave motion.

- **When the wave and the particles carrying the wave move in the same direction, an excess of the particles build up in the wave.**
- **When the wave and the particles carrying the wave move in opposite directions, a deficiency of particles develops in the wave.**

This is also true for the sine waves of electrical energy in a wire. The energy carrying particles in this case are electrons, which are negatively charged. The part of the wire which does not move, the atoms of copper for instance, are positively charged.

Figure 4A. Electron Flow in a Wire Touched by a Negative Voltage Source.

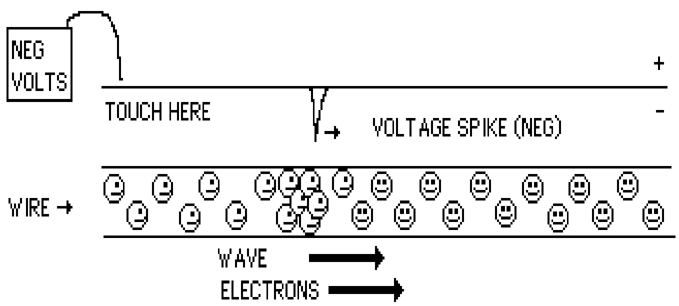


Figure 4B. Electron Flow in a Wire Touched by a Positive Voltage Source

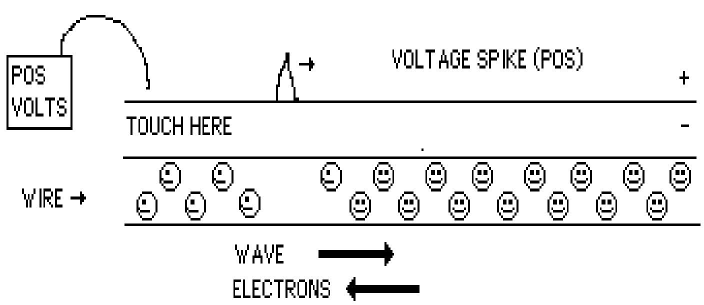


Figure 4A represents a wire, one end of which momentarily comes in contact with a negative voltage source at the left end. This results in the injection of a pulse of electrons into the wire and a negative voltage spike (wave) which travels down the wire to the right. The electrons in the wire will move from left to right in the same direction as the wave. The voltage spike is the result of a compression of electrons in the wave as it moves down the wire. It is a negative voltage because in any given space there are more negatively charged electrons than there are positive charges from the atoms.

In Figure 4B the wire is momentarily touched by a positive voltage source. This results in electrons being sucked out of the wire and a positive voltage spike which moves down the wire to the right as the electrons in the wire move to the left to fill in the void. It is a positive voltage because there are now fewer negatively charged electrons than positive charges from the atoms. Again, the wave and the particles are moving in opposite directions resulting in a deficiency of particles, in this case electrons, and thus a positive charge.

Figures 4A and 4B represent what happens in each half of a voltage sine wave. In the negative half of the voltage sine wave, the electrons and the wave move in the same direction - away from the source - resulting in a negative charge propagating down the transmission line. In the positive portion of the voltage sine wave, the wave again moves away from the source but the electrons move toward the source, resulting in a positive charge propagating down the wire.

Now let's relate this to the transmitter and the antenna and the conducting cable that runs between them.

We now have four separate cases to consider as illustrated in Table 1. In Cases I and II we have a forward power wave with the transmitter as the source. In Cases III and IV we have a reflected power wave with the antenna as the source. In both pair of cases the RF energy waves are made up of sine waves of alternating current at RF frequencies. Table 1 examines the direction of electron flow and the charge (sign of the voltage) associated with each of the four half waves.

- **In Case I, we have a forward (transmitter = source) power wave with electrons moving away from the transmitter resulting in a negative charge.**
- **In Case II, we have a forward (transmitter = source) power wave with electrons moving toward the transmitter, resulting in a positive charge.**
- **In Case III we have a reflected (antenna = source) power wave with electrons moving away from the transmitter, resulting in a positive charge.**
- **In Case IV we have a reflected (antenna = source) power wave with the electrons moving toward the transmitter, resulting in a negative charge.**

Table I. Electron Flow and Electronic Charge in the Forward and Reflected Waves.

CASE	WAVE TYPE	WAVE SOURCE	ELECTRON FLOW DIRECTION	ELECTRONIC CHARGE
I	FORWARD	XMTR	FROM XMTR	(-)
II	FORWARD	XMTR	TO XMTR	(+)
III	REFLECTED	ANTENNA	TO ANTENNA	(+)
IV	REFLECTED	ANTENNA	FROM ANTENNA	(-)

Notice that in both Cases I and III the electrons are moving away from the transmitter, but the waves have opposite charges associated with them. Also, in both Cases II and IV the electrons are moving toward the transmitter but the waves have opposite charges associated with them. The difference is that the source of the power wave is different. From Table 1 you can see that each case is unique. In no two cases are the direction of electron flow and the charge associated with the wave identical.

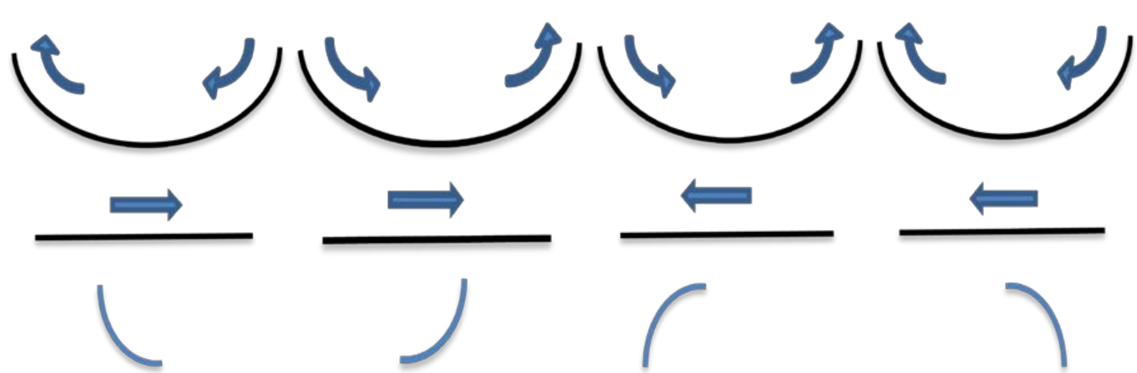
AHA!! Here it is! In cases I and III the electrons are flowing in the same direction but the charges (and thus voltages) are opposite. In one case, the forward wave, the voltage and current are in phase, Case I, and in the other, Case III the reverse wave, **the current and voltage MUST BE 180 degrees out of phase!**

Capacitive and Inductive Coupling

Now we have a handle on the relevant properties of the waves in a transmission line. But before I can put everything together I have to remind you of two simple principles of which you are probably already familiar.

The first is the magnetic or inductive effect which a changing current flow in one wire has on a nearby and parallel wire. Imagine a sinusoidal alternating current flowing in a wire. As shown in Figure 5, a current of negatively charged electrons, which is <u>increasing</u> and moving from left to right in the primary line (the solid straight lines in Figure 5) will induce a current which flows from right to left in an adjacent secondary line (Depicted by the curved line above the primary line). The induced current flows in the opposite direction as the source current as long as the source current is <u>increasing</u>, (as indicated by the ¼ sine wave shown below the primary). At the bottom of the sine wave, the source current stops increasing and the induced current in the adjacent line drops to zero. As the current in the primary then starts to <u>decrease</u>, as shown in Figure 5B, but continues to flow to the right, the induced magnetic field begins to collapse and the current in the adjacent line begins to increase and flow to the right, the same direction as current in the main line.

Figure 5. Inductive or Magnetic Effect



A B C D

As the current in the primary drops to zero and then starts to increase in the opposite direction, to the left, the current in the adjacent sample line reaches a maximum but continues to the right. As the left going current in the primary reaches its maximum value the induced current in the sample line goes to zero again. As the main current to the left begins to decline the current in the sample line changes direction, to the left, and begins to increase, until the current in the main line falls to zero. Then the cycle repeats itself.

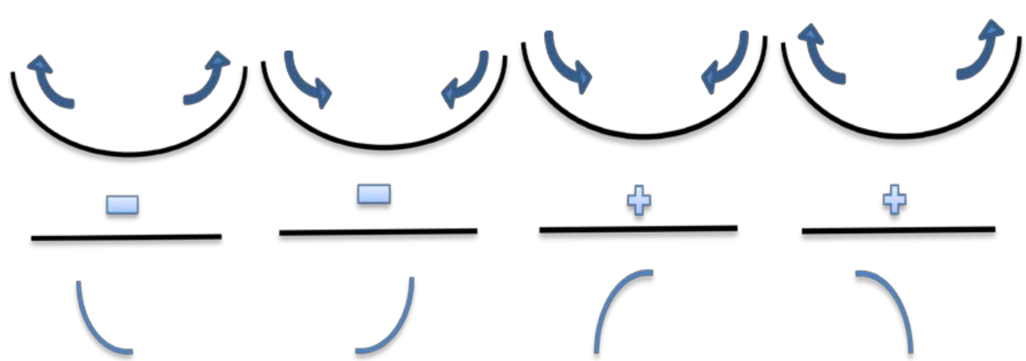
From this analysis, it is clear that the current in the sample line lags the current in the primary line by 90 degrees.

This is the same inductive effect which operates in transformers.

The second concept I want to remind you about is the capacitive effect which an electrical charge has on a nearby conductor. As shown in Figure 6A, when an increasing negative charge (increasing negative voltage) on the primary wire is brought near a secondary wire loop, electrons will flow away from that charge up both legs of the loop. When the negative voltage (charge) on the primary wire reaches its maximum value, current in the secondary loop goes to zero. As the negative voltage begins to fall, Figure 6B, the current in the sample loop reverses direction as electrons are drawn back into the wire. When an increasing positive charge (increasing positive voltage), Figure 6C, is brought near the loop, electrons will flow toward the charge in both legs of the loop. When the positive voltage on the primary reaches its maximum, the current in the secondary loop goes to zero. As the positive voltage begins to fall, Figure 6D, the current reverses direction as the electrons start to flow out of both legs of the sample loop.

In this case the source voltage (a measure of charge density) in the primary line lags the current in the secondary line and is also 90 degrees out of phase with the current in the secondary line.

Figure 6. Capacitive or Electrostatic Effect



A B C D

Both the electrostatic and magnetic effects are being detected in an SWR meter. The sampling loops in the SWR meter of Figure 1 are the same as the secondary loops in Figures 5 and 6. All we have to do now is put together our understanding of the relevant characteristics of wave propagation in a transmission line with the resultant inductive and capacitive effects which the waves exert on the sampling loops in the SWR meter.

Figure 7A A Simple Reflectometer type "SWR Meter"

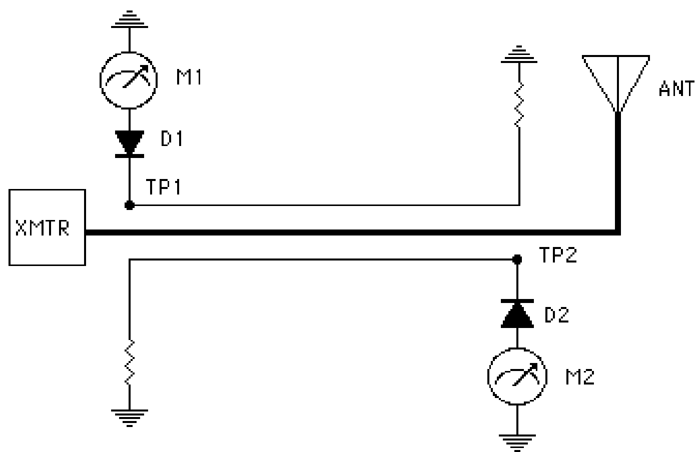
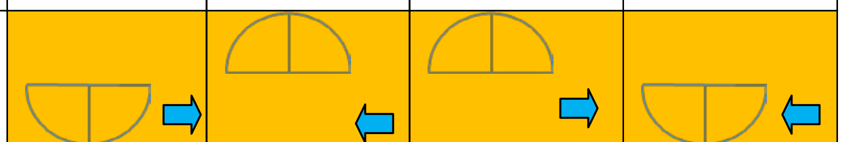
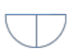


Figure 7A shows a simple SWR meter with test points TP1 and TP2. In order to get a reading on meter M1 or M2, diode D1 or D2 must conduct a current. In order for diodes D1 or D2 to conduct a negative voltage must be developed at TP1 or TP2, respectively. Another way to look at it is that electrons only flow through the diodes as shown when electrons flow into the diodes from the sample loop toward the meter. A positive or zero voltage at these test points will result in no meter reading.

What we have to do now is to consider each of the four quarter sine waves described in Table 1. We need to determine the results of the inductive effect and the capacitive effect of these four quarter waves on the detector loops and the resulting voltages at TP1 and TP2. In some cases the inductive effect and capacitive effect will produce opposite voltages which will cancel. In other cases the two voltages will add. Only when those two voltages are negative will D1 or D2 conduct.

Figure 7B. Analysis of Induced currents at TP1 and TP2 in Fig. 1.

	FORWARD WAVE		REFLECTED WAVE	
	Case I	Case II	Case III	Case IV
Center Conductor				
TP1 Cap TP1 Ind Net	← → ← → ⇦ ◎	→ ← → ← ◎ ⇦	→ ← ← → ◎ ◎	← → → ← ◎ ◎
TP2 Cap TP2 Ind Net	→ ← ← → ◎ ◎	← → → ← ◎ ◎	← → ← → ◎ ⇨	→ ← → ← ⇨ ◎

 Represent the voltages of the sine waves in the "Center Conductor" and emphasizes the fact that we are considering the ¼ sine wave which is increasing in absolute value and the ¼ sine wave which is decreasing in absolute value, separately.

 Indicates the direction of current flow in the center conductor (the primary).

← Indicates the direction of the force acting on the electrons in the sample loop (the secondary) near each Test Point in Fig. 1 due to the isolated inductive or capacitive effect.

◎ Indicates that the sum of the inductive and capacitive effects cancel each other or that the resultant force results in a positive voltage at the corresponding Test Point in Fig. 1 and the diode does not conduct.

⇨ Indicates that there is a net force creating a negative voltage at the corresponding Test Point in Fig. 1 and that the diode will conduct.

Figure 7B shows the results of the inductive and capacitive effects on the voltages at TP1 and TP2. Since the inductive and capacitive effects depend upon whether the voltage and current are increasing or decreasing, we must consider what is happening during each quarter of the sine wave for both the Forward Wave and the Reflected Wave. Consequently, there are eight sets of conditions in which we need to consider both the inductive and capacitive effects. The half wave above the midline represents a positive voltage on the center conductor and the half wave below the midline represents a negative voltage on the center conductor.

Case I

Column 1 represents the half of the forward wave with the electrons flowing out from the transmitter. With the wave and energy carrying medium (electrons) moving in the same direction (away from the transmitter) there is a net negative charge associated with the wave and thus a negative voltage. The direction of electron flow is our clue to the inductive effect and the negative charge is our clue to the capacitive effect.

If we consider TP1 we see that the inductive effect from the **increasing** current flow, results in electrons in the detector loop flowing toward TP1. This leads to a negative charge build up at D1 and results in a negative voltage there. The capacitive effect of this portion of the forward, wave also pushes electrons toward TP1 and also produces a negative voltage. In this case both effects generate negative voltages which reinforce each other. This portion of the forward wave will cause D1 to conduct and a reading will be obtained on M1. In the **decreasing** portion of the wave (i.e. the absolute value is approaching zero), both the results of the capacitive and inductive effects are reversed, resulting in electrons in the detector loop flowing away from TP1, forming a positive charge at TP1. In this portion of the wave, the diode does not conduct and no reading will be seen on the meter.

Now let's look at column 1 again and see what the effect of this portion of the forward wave has on TP2. Again we have a negative charge in the wave and the electrons are moving toward the antenna. In the **increasing** portion of the wave, the inductive effect will result in electrons being pulled away from TP2 and thus a positive charge will develop. The capacitive effect will push electrons toward TP2 and generate a negative charge and thus a negative voltage. In this case the two voltages are of opposite sign and will cancel each other resulting in a zero voltage at TP2. Diode D2 will not conduct and no reading on M2 will be observed. In the **decreasing** ¼ section of the wave, both the inductive and capacitive effects are reversed. The resulting forces again cancel each other, diode D2 will not conduct and no reading on M2 will be observed.

Case II

Column 2 of Table 1 represents the second half of the forward wave in which the electrons are now coming back toward the transmitter (clue to the inductive effect). In this case the wave (Forward) and the wave carrying electrons are moving in opposite directions. This results in a net positive charge (clue to the capacitive effect) in the wave.

During the increasing portion of the wave, the inductive effect generates a positive voltage at TP1 and the capacitive effect also generates a positive voltage. The resultant is a net positive charge and voltage at TP1. D1 will not conduct and M1 will show no reading. During the decreasing portion of the wave the capacitive and inductive effects are both reversed resulting in a negative voltage at TP1 and now the diode will conduct and M1 will indicate current flow.

At TP2, during the increasing portion of the wave, the inductive effect generates a negative voltage while the capacitive effect generates a positive voltage. These two voltages cancel each other resulting in zero voltage at TP2. Diode D2 will not conduct and no reading will be observed on M2. During the decreasing portion of the wave, the capacitive and inductive effects are again reversed and the resulting voltages still cancel. Diode D2 will not conduct and no meter reading will be observed.

Case III

In column 3 we have that portion of the reflected wave (from antenna to transmitter) in which the electrons are flowing toward the antenna (inductive clue). In this case the wave and electrons are flowing in opposite directions resulting in a positive charge in the wave front (capacitive clue).

During the increasing portion of the half wave, the inductive effect generates a negative voltage at TP1 and the capacitive effect generates a positive voltage. These voltages cancel and diode D1 does not conduct and M1 is silent. During the decreasing portion of the wave, the inductive and capacitive effects are reversed and still cancel each other. Diode D1 does not conduct and M1 shows no current flow.

At TP2, during the increasing portion of the half wave, both the capacitive and inductive effects generate positive voltages. D2 does not conduct and M2 is also silent. However, during the decreasing portion of the half wave, the capacitive and inductive effects are reversed and a negative voltage build up at TP2. Diode D2 conducts and M2 will indicate a current flow.

Case IV

In column 4 the electrons in the second half of the reflected wave are flowing toward the transmitter (inductive clue) and both wave and electrons are moving in the same direction resulting in a negative charge in the wave front (capacitive clue).

During the increasing portion of the wave the inductive effect produces a positive voltage at TP1 while the capacitive effect produces a negative voltage. These voltages cancel, D1 does not conduct and M1 is lifeless. During the decreasing portion of the half wave, the capacitive and inductive effects are reversed, the voltages still cancel, D1 does not conduct and M1 is lifeless.

At TP2, during the increasing portion of the half wave, both effects the inductive and capacitive effects produce negative voltages.. D2 conducts and M2 finally comes to life, indicating current flow. During the decreasing portion of the half wave, the capacitive and inductive effects are reversed resulting in a net positive voltage at D2. Diode D2 will not conduct and M1 is unresponsive.

Conclusions

From inspection of Figure 7B we can see that M1 only gives a reading during half of the forward wave. Also, M2 only gives a reading from half of the reflected wave.

Several additional observations can be made from inspection of Figure 7B.

- First, with calibration, M1 will indicate the forward power and M2 will indicate the reflected power. If you know the forward and reflected power accurately, you can calculate the SWR or devise a circuit to display the SWR on a meter.
- Second, the need for careful placement of the SWR meter detector loops is the result of the need to carefully balance the inductive and capacitive responses to the electron flow and charge density.
- Third, the orientation of the diodes is irrelevant. For instance, if we reversed D1 in Figure 7A, then the diode would still only conduct in response to the forward wave, but during the other half of the forward wave. It would still measure the forward power only. It would however be necessary to reverse M1. In fact the diodes are not even necessary. You could replace the diodes with light bulbs and get an indication of the forward and reflected power by measuring the bulb's brightness. This was beautifully demonstrated by W8YZ, Dave Smith, at the 1987 ARRL Great Lakes Division convention in Saginaw, Michigan. More recently, the same demonstration is being presented by William Hays, AE4QL, in his presentation, “Standing Up for Standing Waves”, http://home.windstream.net/whays101/index.html.

The whole of transmission line propagation is much more complex than what I have described here. However, I hope this article has given you a little insight into transmission line wave propagation and how a reflectometer, direction coupler can tell the difference between a forward wave and a reflected wave on the same transmission line.

1. The ARRL Handbook for Radio Communications, 2010 Edition, pg 20.4.
2. M. W. Maxwell, Reflections III Transmission Lines and Antennas, 2010, p 3-1.
3. R.E. Bloom, “Inside the Bird”, 73 Magazine, July, 1977, p 44..
4. Bruene, Warren B, W0TTK, “An Inside Picture of Directional Wattmeters”, QST, April 1959, pg 24.
5. F.E. Terman presents a brief description of how the directional coupler works in terms of magnetic and electric field coupling but again, in terms of waves on a primary line, not in terms of electron flow. F.E. Terman, “Electronic and Radio Engineering”, McGraw-Hill Book Company, 1955, Chapter 4-13, p 115
6. O. Norgorden, “A Refelectometer for H-F Band, NRL Report 3538. September 15, 1949.

Dennis K. Klipa, N8ERF
June 28, 2020