The hybrid coil circuit in telephone transmission systems

Douglas A. Kerr

Issue 6
July 28, 2025

ABSTRACT

In applying amplification to a telephone circuit in which speech signals travel in both directions through a single electrical circuit (a so-called "2-wire" circuit), we must be able to separate the two directions of signal travel so that each can be amplified by a separate amplifier, while not allowing any path "around the loop" that could lead to oscillation. This need is traditionally met by an ingenious passive circuit known as a "hybrid coil circuit" (often called just a "hybrid").

In this article, I first give some context for that need. Then I examine the properties of the circuit as a "black box". Then I show traditional implementations of the circuit, and describe how they do what they do.

Finally, I discuss important applications of the circuit in telephone transmission beyond the original context, and in telephone sets themselves.

1 BACKGROUND

1.1 Early telephone circuits

The earliest telephones were used in what we might today call "intercom" service, perhaps proving communication between a ranch house and the bunkhouse, or from a sales office to the warehouse across the alley.

The earliest transmission medium was typically a single wire, operated "against ground" (that is, the circuit was completed through the earth). But this had numerous shortcomings from a transmission standpoint. Variations in the earth potential between the two stations would directly introduce interfering voltages into the circuit. And the unbalanced nature of the circuit meant that currents or voltages induced into the circuit (perhaps from its passing near electrical power wiring) could lead to noise on the circuit.

So very soon, it became almost universal to use a pair of wires for any telephone circuit, this being operated in a "balanced" mode with regard to the AC signals on the line.

1.2 Telephone exchange service

Before long, the concept of the *telephone exchange* arose. Here, telephones in a town's various businesses, farms, and residences were connected by pairs of wires to a central point, where a human operator working some type of switchboard could establish a connection from any customer to any other customer.

At first, the possibility only existed for a caller to call another customer in the same city. But eventually the concept expanded, next allowing calls between nearby cities, and eventually between cities that were not so nearby.

1.3 Attenuation

Of course, these passive, physical conductor circuits afforded a non-trivial signal attenuation ("loss") per unit of length. For calls within a geographically-small service area, this would not be a serious deficiency. But as interest grew in providing communication over a large metropolitan area, or between nearby cities, or between not-so-nearby cities, or even possibly across the entire nation, attenuation was a existential problem.

The major contributor to the attenuation was loss in the resistance of the conductors, and so a brute-force solution would be to reduce that, which essentially meant increasing the cross-sectional area of the conductors. But at best this greatly increased the cost of the conductors.

If the format was aerial open wire, the increased weight of the larger-gauge conductors meant that the poles would have to be closer together, again a matter of increased cost (as well as possible greater dissatisfaction of the nearby residents). In cables, larger gauge conductors meant that fewer pairs could be included in a cable of manageable diameter.

For open wire lines, the largest gauge widely used was a diameter of 0.165". In cables, the largest gauge fairly widely used was 7 AWG; a gauge of 9 AWG was very widely used in cables for longer circuits.

1.4 Inductive loading

A significant advance in this matter was the development, by American telephone engineer Michael Pupin, of a system of adding inductance periodically to the telephone lines. Combined with the inherent mutual capacitance of the conductors, this in effect created a distributed low pass filter, and it turned out that its insertion loss (for frequencies below its "cut off" frequency) was less than the attenuation of the conductors if used alone.

A further advantage was that, over the intended operating range of frequencies, the attenuation was more uniform with frequency than that of the basic conductor pair, desirable for realistic reproduction of the original speech..

This scheme came to be known in the U.S. as *inductive loading* (although, interestingly enough, elsewhere in the world it was often called “Pupinization”).

It may seem counter-intuitive that a passive “filter” would have (in its passband) less attenuation that that of the conductor pair around which it was constructed. The key to this conundrum is the matter of the *characteristic impedance* of the line (the parameter that represents the ideal ratio of voltage to current in the propagated signal). The characteristic impedance of the “loaded” circuit is significantly greater in magnitude than that of the basic conductor pair (perhaps twice as much).

As a consequence, for any given signal power on the loaded circuit, the signal voltage is higher, and the current lower, than for that same power carried over the basic pair (the voltage increasing as the square root of the impedance).

The “loss” in a conductor pair is preponderantly due to the series resistance of the conductors, and it goes as the square of the current. Thus, for the loaded pair, with its lower current for any given signal power, this loss is significantly less than for operation over the basic cable pair.

This technique was extensively used on circuits of many types, and allowed a considerable advance in the ability to have communications over longer distances than before. Even after the introduction of multiplex transmission for “intercity” circuits (see Section 18), for many years the *loaded cable pair* became the medium of choice for trunk circuits between the central offices in a metropolitan area. Often those cables had conductor gauges as small as 22 AWG.

But for longer distances, even with this technique, the attenuation of the circuit would be too great for desirable transmission performance.

1.5 The “repeater”

This has little real to do with our story, other than as the source of an important term, but it is such a great story I fell compelled to tell it.

At one time, calls across the country (say, between New York and San Francisco) were made practical by a rather bizarre facility. Such calls were routed, in Chicago, though one of a number of “booths” in which there was a telephone operator (invariably male) equipped with a high-efficiency telephone set, with headphones for receiving, whose

transmitter (microphone) could be switched to either the "eastbound" or "westbound" direction of the overall connection.

So when a subscriber in San Francisco said "to his attorney" in New York, "Arthur, I've been thinking about that contract extension", the operator heard that, switched the transmitter to the eastbound line, and said (loudly), "Arthur, I've been thinking about that contract extension."

Suffice it to see, this was not overall a big hit.

The special operator was called (understandably) a "repeater", and the term was actually applied to that whole installation.

1.6 A better repeater

The next stage of development was to take the human out of this scheme. In a system developed by H. E. Shreeve of Western Electric, the incoming signal moved an armature in a way equivalent to how the diaphragm of a telephone receiver was moved. The armature was directly connected by a push rod to where the diaphragm would go in a telephone set transmitter ("microphone").[1]

The transmitter was energized with a high DC current [2], and thus had a very high "acoustic-electric conversion efficiency". Thus the ongoing signal had significantly greater power than the arriving signal. This was very clever, although the apparatus was rather troublesome. But it was a step in the right direction.

Note that this was arguably the first audio amplifier (although we usually do not say "audio" in the field of telephone transmission, but rather "voice frequency" or "speech").

Not surprisingly, the apparatus that revolved around this mechanism came to be known as—a *repeater*. And in fact, even in modern times, an amplifier in a telephone transmission system is usually known as a *repeater*.

There was also an important fly in the ointment here, in fact a central premise of this article—but I'll delay addressing it for just a little bit.

---

[1] This was a variable resistance carbon transmitter, where the movement of the diaphragm changed the resistance of a chamber full of carbon granules (a "beefed up" version of those used in typical telephone sets). A DC current was run through this, and the variation of resistance generated an AC voltage.

[2] And had cooling fins!

1.7 Enter the vacuum triode

The development of the vacuum triode by Lee de Forest provided the final link in this branch of the story. It could be used in an amplifier that gave better performance than Shreeve's electromechanical amplifier element. And it led to the first really attractive telephone repeater.

2 THE FLY IN THE OINTMENT

2.1 Two-way transmission

Of course, the most common telephone circuit was expected to transmit in both directions over a single pair of wires (without any kind of "talk-listen switching"). And we must also provide for amplification ("gain") in both directions. Doing this in the straightforward way involves an amplifier for each direction.

But how would we set that up? As in Figure 1?

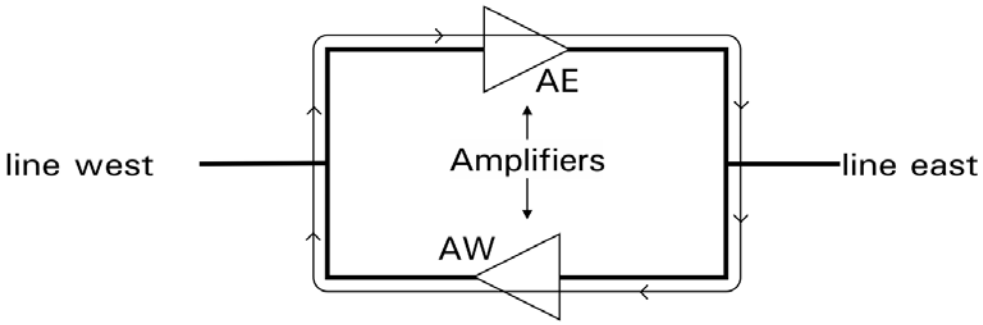


Figure 1. The feedback path

No, not really. The light line shows that there is a circulating "feedback" path through both amplifiers that would make this circuit just oscillate.

No, what we need is something that would in effect do what is seen in Figure 2.

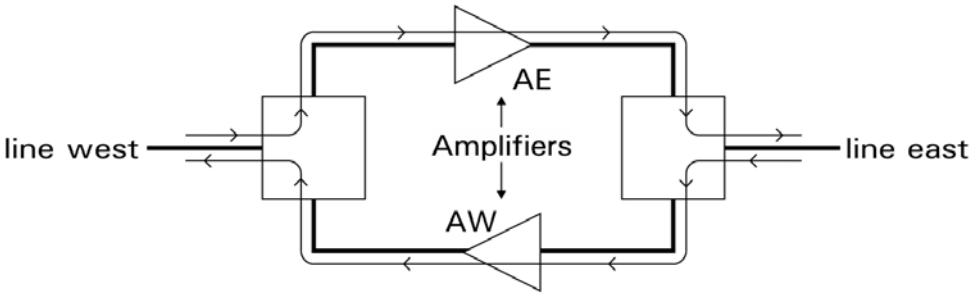


Figure 2. Magic boxes?

Here we connect the two amplifiers to the two lines through some kind of "magic boxes" that would route the signals in the way we

need without there being any "feedback" path. But what might be in those magic boxes?[3]

2.2 The hybrid coil circuit—behavior

Well, what we put there is usually a *hybrid*[4] *coil*[5] *circuit* (often just called for short a "hybrid"). Before we see how we would might actually make one, let's look at its behavior. We will look at Figure 3.

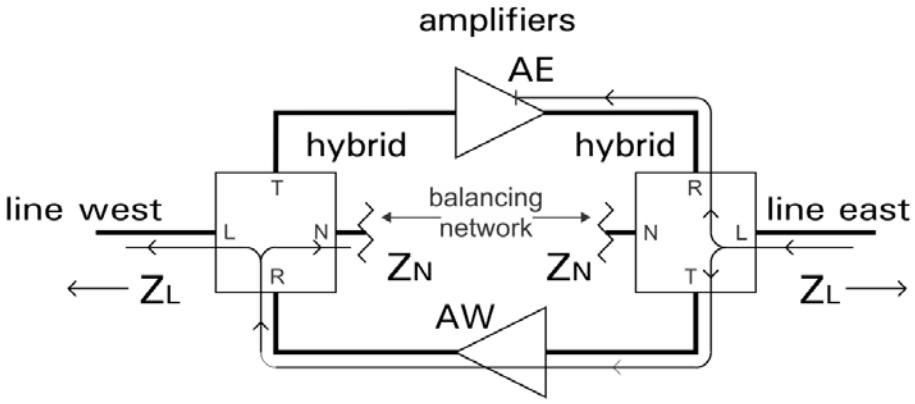


Figure 3.

The hybrid (I will often use its short name) is a symmetrical four-port network. Associated with is a *balancing network*.[6] This is a "single-ended artificial line"; we want its impedance (Z_N) to be the same as the impedance "looking into" the associated line (Z_L), at least over the passband of the amplifiers. (These have bandpass filters in them, not shown here, limiting their response to the range of frequencies we have decided we want to support in our transmission system.)

The impedance of the line varies over that range of frequencies, and is usually not purely resistive. Rather, at any given frequency it usually includes a reactive component, typically capacitive.

To emulate this, the balancing network is typically a *network* of resistors and capacitors, thus its name. It might be quite complicated.

---

[3] Those working in parts of the radio field might guess that these were "directional couplers", but their behavior is quite different.

[4] Yes, an odd name. I do not know how it was chosen.

[5] "Coil" here is short for *repeating coil* (sometimes just *repeat coil*.

[6] Often the term "hybrid" is used to mean the hybrid coil circuit itself plus the associated balancing network.

I have labeled the four ports of the hybrid circuit proper **L** (for line), **N** (for network), **R** (for receive), and **T** (for transmit).

We begin by considering the voice signal arriving from the east. It enters the "east" hybrid at its **L** port. The arriving power divides equally, one half exiting the hybrid's **R** port and the other half exiting its **T** port. Why do we want that?

Well, we don't **want** it; it is just part of the price we pay to get the property we really want, of which we will hear shortly.

The portion exiting port **R** tries to enter the output of the eastbound amplifier (AE). But of course it can't do anything there, and is just absorbed in the output impedance of that amplifier.

The portion exiting port **T** enters the input of the westbound amplifier (AW). It is amplified and goes from the output of that amplifier into port T of the west hybrid.

Now, if indeed, at all frequencies of interest, the impedance of the west hybrid balancing network (Z_N) is exactly equal to the impedance looking into the line west (Z_L), then half that power goes out the L port into the line, and half out the N port into the balancing network.

Why do we want the latter? Well, again we don't **want** it; it is just another part of the price we pay to get the property we really want, of which we will hear now.

Most importantly, in this case, none of the output power from amplifier AW exits from port **R** of the west hybrid. Thus the "feedback" path does not exist, and **that** is what we want.

Of course for an eastbound signal, arriving over the line west, the operation is just the same, but as if with the figure reversed.

2.3 Gain and loss

Again referring to a westbound signal, ideally the power arriving from the line east is split exactly in two, only half going to the input of amplifier AW, which represents almost exactly 3 dB of loss at the east hybrid. Similarly, the power from the output of amplifier AW is split exactly in two, only half of it going out to the line west, which represents almost exactly 3 dB of loss at the west hybrid. The total theoretical loss through the repeater from this "power splitting" is almost exactly 6 dB,

This it would seem that if we wanted this repeater to introduce a gain of 10 dB into the line, we would need to make the amplifiers themselves have a gain of 16 dB each $(10 + 3 + 3)$.

But in fact there is some actual "loss" in the hybrid (as its transformers are not "100% efficient"). This typically amounts to a loss of less than 1 dB. But for tutorial purposes, we often speak as if that loss were exactly 1 dB.

Thus for the incoming signal on its way to the input of amplifier AW, we consider the loss through the to bt 4 dB, and the same at the west hybrid on the way from that amplifier.

Thus, if we wanted this repeater to introduce a gain of 10 dB into the line, we would need to make the amplifiers themselves each have a gain of 18 dB (10 + 4 + 4).

3 THE HYBRID COIL CIRCUIT—IMPLEMENTATION

3.1 The "classical" implementation

3.1.1 *Introduction*

Figure 4 shows (in slightly simplified form the "classical" implementation of a hybrid coil circuit in a repeater such as I have discussed, used as the "west" hybrid of a repeater, and showing a signal arriving from the east.

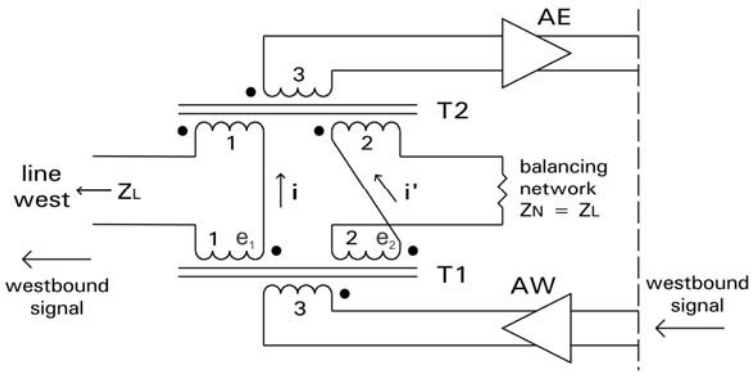


Figure 4. Westbound signal

The hybrid coil circuit proper comprises two *repeat coils* (transformers), T1 and T2. [7] (I will call them "transformers" here.) The black dots show the "corresponding" ends of the various windings of a given transformer (the ends that when one is positive the others will be positive).

[7] "Coil" here is short for *repeating coil* (sometimes just *repeat coil*), the traditional telephone term for an audio line transformer (but as I mentioned earlier we never say "audio" either, rather "voice frequency" or "speech")

3.1.2 *A westbound signal*

The westbound incoming signal as arriving over the line from the east (actually via the "east" hybrid, not seen here) passes through amplifier AW.

The output of amplifier AW goes to winding 3 of T1. It induces equal voltages (e_1 and e_2) in windings 1 and 2. Voltage e_1 propels a current, i, through winding 1 of T2 and through the line west. Voltage e_2 propels a current, i', through winding 2 of T2 and through the balancing network. Of course the current through the line propagates a signal westward through the line,

If, as we hope we have arranged, the impedance of the balancing network, Z_N, is exactly the impedance looking into the line west, Z_L, then i and i' will be identical. Those two currents pass through windings 1 and 2 of T2, but in opposite directions (observe the dots). Thus, there is no net magnetic effect on the core of T2, and accordingly, no voltage is induced in winding 3 of T2

Thus there is no signal into the input of amplifier AE, our wish for this circuit; ideally, none of the incoming signal is amplified and sent (through the east hybrid) to the line east.

Note also that, because there is a net zero magnetic effect on the core of T2, there is no voltage across windings 1 and 2 of T2 to disrupt the earlier picture of the currents caused by voltages e_1 and e_2.

3.1.3 *An eastbound signal*

In Figure 5 we see the same hybrid coil circuit in place in the repeater, but we now consider an eastbound signal arriving from the west over the line west.

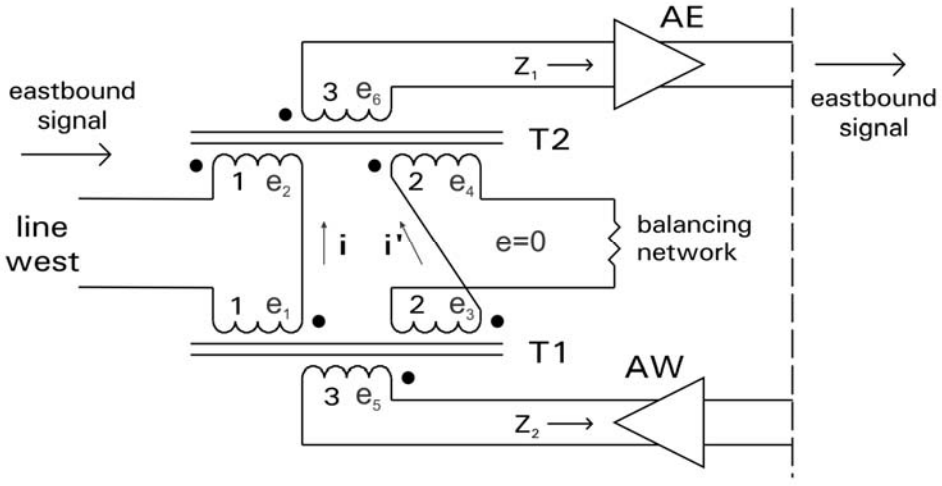


Figure 5. Eastbound signal

We first note that the input impedance of amplifier AE (Z_1) is reflected through transformer T2 and appears (scaled) as seen looking into winding 1 of transformer T2. Similarly, the output impedance of amplifier AW (Z_2) is reflected through transformer T1 and appears as seen looking into winding 1 of T1. And those two impedances are made equal by tailoring Z_1 and Z_2 to be equal in the circuit design of the amplifiers.

The arriving signal passes through windings 1 of T1 and T2 in series, Since the impedances seen looking into those windings are equal, the voltage across those two windings (e_1, e_2) will be the same. Thus equal voltages (e_3, e_4) will be induced in winding 2 of T1 and winding 2 of T2. Those two windings are connected, **in series opposition** ("dot to dot"), to the balancing network.

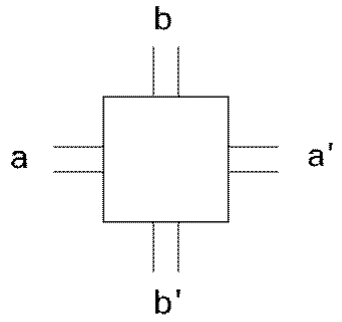
Thus the net voltage provided to the balancing network is zero—thus none of the signal power goes into the network. This is of course not really one of our objectives, but this illustrates the symmetry of the hybrid coil circuit (discussed in Section 3.3). And current i' will thus also be zero.

But current i causes voltage e_1 to be developed across winding 1 of T1, and voltage e_2 to be developed across winding 1 of T2. and thus voltages e_5 across winding 3 of T1 and e_5 across winding 3 of T2.

These voltages appear across the output of amplifier AW and the input of amplifier AE, respectively. Of course the voltage across the output of amplifier AW does nothing useful, but the voltage across the input of amplifier AE is amplified and (through the east hybrid of the repeater, not seen here) is sent into the line east.

3.2 The hybrid as a general four-port network

A hybrid can be treated as a generalized four-port network, shown in Figure 6.



a, a'; b, b": conjugate port pairs

Figure 6. The hybrid as a generalized four-port network

Relating this to Figure 4, we can consider the line (line West in the figure) to be on port a, the balancing network to be on port a', the input to amplifier AE to come from port b, and the output of amplifier AW to go to port B'.

3.3 The hybrid, rotated

But, treated in this "black box" way, it turns out that the black box has *quarter-turn symmetry*. That means that we could use port b for the line, put the balancing network on port b', and use ports a and a' as the connections to the two amplifiers. Said another way, we could rotate the hybrid circuit by 90° and it would still work in the same way.

That would lead to the actual circuit seen in Figure 7.

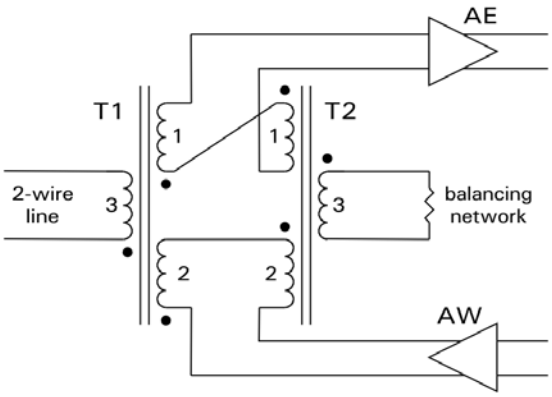


Figure 7. The classical hybrid circuit rotated

Actually, this is not often done in this kind of application. But a little later, we will see a case in which another hybrid implementation is indeed often used in its "rotated" form.

3.4 DC signaling

In many of the situations in which a hybrid circuit is used, DC voltages are placed on the two conductors of the line for signaling purposes. These DC voltages of course coexist on the line conductors with the AC voltages carrying speech.

Conceptually, this is done by connecting the AC signals to the line through a high-pass filter and the DC signaling voltages through a low-pass filter.

But in reality, this is sometimes done in a way that does not require explicit low-pass and high-pass filters. We see the principle in Figure 8 (which, for clarity as to this principle, does not involve a hybrid circuit, just a "line transformer", T1).

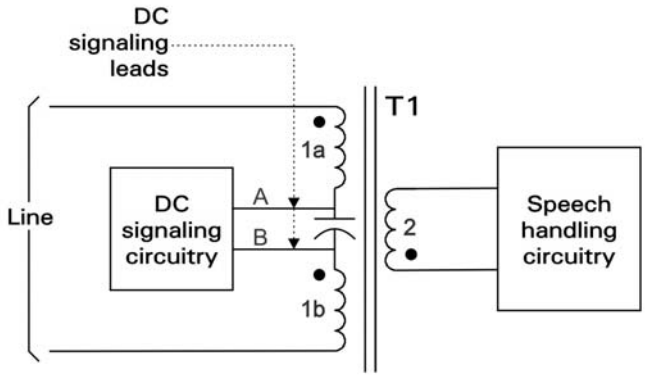


Figure 8. DC signaling paths–concept

We see that the line winding of transformer T1, winding 1, has actually been split into two parts (1a and 1b). The inner ends of the two parts are connected by a capacitor so the winding, insofar as AC signals are concerned, just appears to be a single winding.

But the inner ends also go over the "signaling leads" (I have labeled them "A" and "B", which in fact are sometimes their designations on telephone system circuit drawings) to the signaling circuit, which applies DC signals to the two leads (and thus over the line conductors) and/or responds to DC signals coming over the line conductors from the distant end.

The signaling leads, from an AC (speech signal) standpoint, are essentially both at "ground potential". But because of the symmetry of winding 1, from an AC standpoint the circuit still operates on a "balanced" basis, which as we noted earlier is obligatory to minimize the effects of noise from various phenomena.

Now when we get to a circuit involving the use of a hybrid coil circuit (a repeater, perhaps), doing this gets a little more complicated. We see a typical implementation in Figure 9.

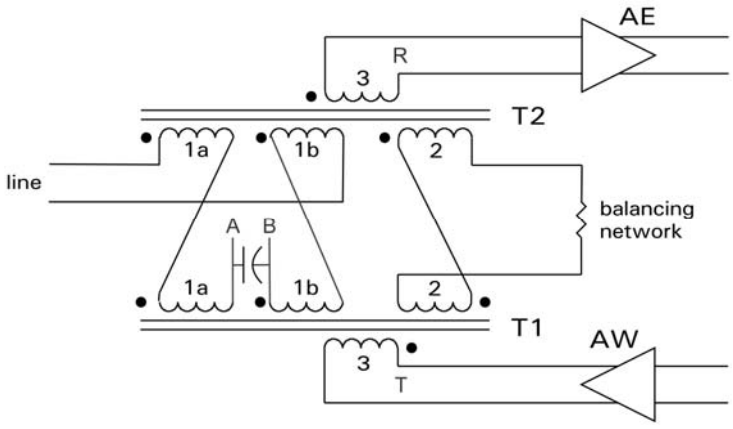


Figure 9. DC signaling paths–in a hybrid

Here, windings 1 of transformers T1 and T2 are “split” (again I label the two parts “1a” and “1b”). As in our earlier work, windings 1 of transformers T1 and T2 are in series, across the line. But here, the two parts of the windings are “interleaved” (for scrupulous symmetry).

Now the “center” capacitor is in the center of winding 1 of T1. As we saw before, the inner ends of that winding (which are the “inner ends” of the whole thing that is across the line) are also the signaling leads, again labeled A and B.

Again, the two signaling lends are, from an AC standpoint, at ground potential. But because of the clever way the line windings are connected, AC transmission is now still “fully balanced”.

3.5 An alternate implementation

An alternate implementation, sometimes called a “single-coil hybrid”, as the name suggests uses only one transformer. Its principle can be seen in Figure 10. I show an unbalanced implementation for simplicity.

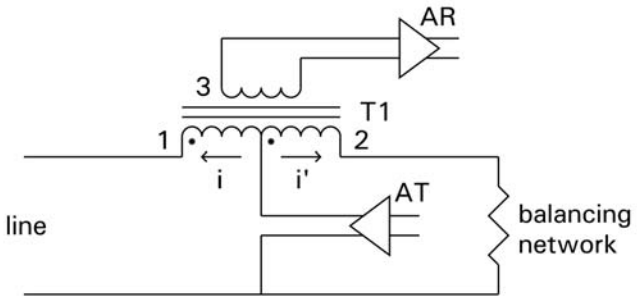


Figure 10. Single-coil hybrid

We will look at the pivotal issue: the output from one of the amplifiers, intended to go to the line. The output from amplifier AT (T for “transmitting”) goes into the “centertap” between identical windings 1 and 2 of the transformer (T1), against the lower conductor of the line.

Currents i (into the line) and i’ (into the balancing network) will be equal if the impedance of the balancing network is the same as the impedance looking into the line (which we hope to have arranged).

But those currents flow in opposing directions in windings 1 and 2 (again. note the dots), and thus their magnetic effect is canceled out. Accordingly, no voltage is induced in winding 3, and no signal passes into the input of amplifier AR (R for receiving), as we desire.

Although it is harder to see at a glance, and I will spare the reader the proof, a signal arriving from the line will have its power equally divided between the input to amplifier AR and (fruitlessly) the output of AT.

Of course in reality, given the need for balance, in actual transmission equipment, winding pair 1-2 is usually actually accompanied by an alter ego in the other side of the line (much as we saw for the 2-coil hybrid in Figure 9 in Section 3.4.

3.6 Rotated

Of course, as we saw in Section 3.6, we could "rotate this hybrid clockwise by 90°", leading to this:

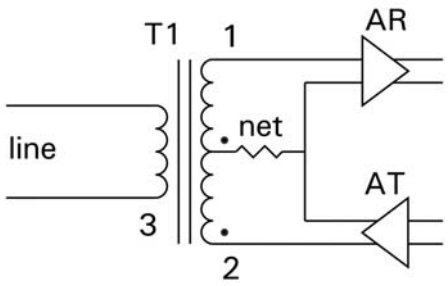


Figure 11. Rotated single-coil hybrid

This could be very advantageous in some applications. Note that one output lead of amplifier AT is common with one input lead of amplifier AR, convenient in transistor electronic circuitry (that might be "circuit ground").

4 QUANTIFYING HYBRID "ISOLATION"

4.1 Transhybrid loss

If we wish to determine the "degree of isolation" afforded by a specific hybrid *in situ*, we can send a carefully-calibrated signal[8] typically with a power of exactly 1 mW) into port **R** and measure the power emerging from port **T**.

The difference in level is called the "transhybrid loss", which can be taken as indicative of the isolation afforded. The greater the loss, the greater the isolation. If the isolation were perfect, the difference in power (expressed in dB) would be infinite (there being zero power out of port **R**).

But this metric is not the one most often cited in telephone transmission work. The next section describes a more meaningful alternative metric.

---

[8] If not stated to the contrary, at a given frequency. One variant uses a "simulated voice" signal extending over the voice band, actually white noise with its frequency distribution shaped, and its bandwidth limited, by a filter.

4.2 Return loss

We can think of an "imperfect" isolation by the hybrid as being equivalent to the hybrid sending, from its **L** port, a certain amount of power into the line, then part of that being "reflected" in the line, returning to the hybrid's **L** port, where it would be evenly divided between the **T** and **R** ports. This is the outlook behind the concept of the "return loss" of the line, as dealt with by the hybrid of interest. [9] The *return loss* is defined numerically as the difference between the "sent" power and the "returned" power, expressed on dB.

The worst possible degree of isolation of the hybrid is equivalent, under this outlook, to all the power sent to the line being reflected back to the hybrid, which would by definition be described as a return loss of 0 dB. But how do we determine that metric?

Even under conditions of a substantial imperfection in isolation, the power sent into the line is less than the power sent into port **T** by (we commonly assume) 4 dB (as a consequence of both "power splitting" and actual loss in the transformers). And similarly, the power sent to port **R** is less than the power reflected back from the line by (we commonly assume) 4 dB.

So in this case, equivalent to the worst possible isolation of the hybrid (a return loss of 0 dB), the power seen at port **R** is 8 dB less that the power sent into port T (that is, by the *transhybrid loss*).

So to get the return loss in this exampole,, we take the 8 dB transhybrid loss and subtract from it twice the "through" loss of the hybrid (here taken to be 4 dB). That gives this "world's worst hybrid" a *return loss* determination of 0 dB, as seems fitting.

4.3 "Precision" and "compromise" balancing networks

If we are interested in getting the best isolation in a hybrid, we will configure the balancing network to closely track the impedance expected to be seen "looking into" the line of interest. This in practice often involves the placing of "straps" on a can full or resistors and capacitors to create (from an empirical "recipe" for the kind of line involved) the optimum network. This kind of balancing network is called a "precision network".

Bu in other situations, we are not dependent on such a high degree of isolation. Then we may use a very simple and universal balancing network that emulates the expected line impedance "closely enough" for many kinds of line for many situations.

---

[9] This outlook fits well with one of the major reasons that, in overall transmission planning, we are interested in hybrid behavior.

For example, for use on a common transmission facility used for trunks between central offices in a metropolitan area (paired 19 or 22 AWG cable conductors with a certain standard inductive loading system), the impedance (over the band of interest) tracks fairly closely with 900 ohms of resistance in series with 2 μF of capacitance.

Accordingly, for this class of transmission facility we use a universal "compromise network" consisting of (nominally) 900 ohms in series with a (nominally) 2 μF capacitor. Of course, there is not a close tolerance expected of these values, since this is in any case a compromise approximation to the expected impedance of the line.

Now, in several common types of Western Electric capacitors used in these networks, in the applicable series of "preferred values" (the "2% tolerance" series) used at one time, the nearest value to 2 μF was 2.16 μF, and so that value was widely used in actual circuits.

Perhaps to avoid any mental distress over this by the designers, the parameters of the compromise network became often stated as "900 ohms in series with 2.16 μF".

More recently, the series of preferred values for Western Electric capacitors were readjusted to match the general electronic industry norms in which (for the 2% tolerance series, called "E48") the closest capacitance to 2 μF is 2.15 μF. But the description of the compromise balancing network stayed (usually) as "900 ohms in series with 2.16 μF".

But because of all that, especially outside of the Bell Telephone System, we sometimes see the specification for this kind of compromise network stated as "900 ohms in series with 2.15 μF". But, in some papers (even from the Bell System),for some reason the parameters are spoken of as "900 ohms in series with 2.14 μF".

As to the resistor, the value 900 ohms was forcibly included in the Western Electric 2% preferred value series, even though it did not follow the series algorithm, because it was so often needed as a terminating resistor (600 ohms likewise), and the designers wanted to do that "exactly". Thus no peculiarity flowed back into the specification of the resistance in the universal compromise network discussed above.

5 OTHER APPLICATIONS IN THE TELEPHONE NETWORK

5.1 Introduction

So far, the discussion of the hybrid coil circuit has largely been in the context of its first major application: in a telephone repeater,

But evolution of the telephone transmission network led to other uses, in a sense an extension of its original task.

In this section, I will discuss an important one of these further uses of the circuit.

5.2 4-wire transmission

The "directional separation" afforded by hybrid coil circuits in the kind of repeater I discussed above is, in practice, imperfect. For one thing, it is not economically practical to measure the impedance looking into individual transmission circuit pairs, or even individual "sets" of pairs, over the working frequency range, and then setting the balancing networks to have that.

In addition, the impedance seen looking into a transmission pair varies with the distribution of temperature along the length of the pair, which of course varies by the hour, the day, and the season.

These departures from ideal performance by the hybrids do not (hopefully) lead to oscillation at the repeaters. But they do represent a source of echo at the repeaters. We see that in Figure 12.

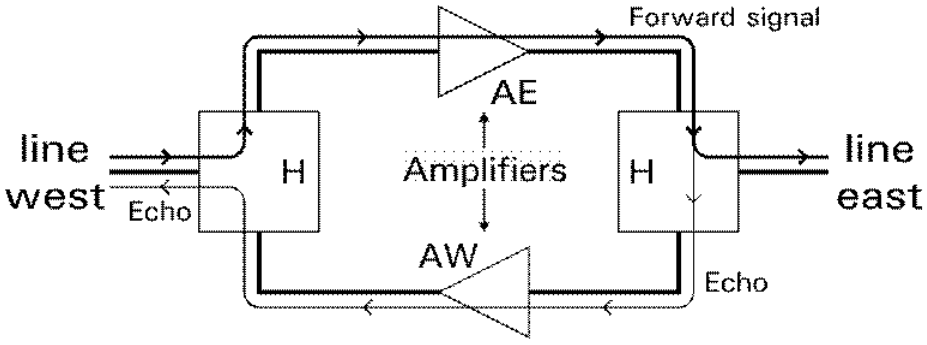


Figure 12. Echo paths

Typically, the amount of such echo contributed at one repeater may not be consequential. But for longer distance circuits (there might be a repeater every 45 miles of so), the accumulation of this echo can represent a serious transmission impairment.

The decision was taken by the system planners to meet this challenge head on. For longer circuits, two pairs would be allocated, one used exclusively for each direction of transmission. As we can readily imagine, the repeaters are now much more straightforward—notably, there will be no hybrids and the associated balancing networks, just one amplifier for each direction of transmission. And there is no opportunity for echo at the repeaters.

Of course, the cost of the transmission media proper (pairs) is now greater—basically twice as much. But reduction in maintenance and

adjustment labor, and the improved performance, was deemed to make this a good choice.

Not surprisingly, this new format was called “4-wire” transmission. And then, in contrast, the term “2-wire” transmission” came into use for the earlier (once universal) arrangement.

Today, those terms have broader, but related, implications.

5.3 A new role for the hybrid

Imagine that a lengthy trunk connecting two central offices in the toll (long distance) network is implemented as a 4-wire circuit, with distinct paths for the two directions of transmission. Originally these paths would each have been implemented by a cable pair, hence the name.

But the switching systems at those central offices work on a 2-wire basis.

So, at these central offices, we need to join a 4-wire transmission path to a 2-wire path. This is conceptually equivalent to what happens inside a repeater in a 2-wire circuit (where the two amplifiers are in a very short “4-wire path”). And of course our go-to circuit for doing that kind of task is the hybrid.

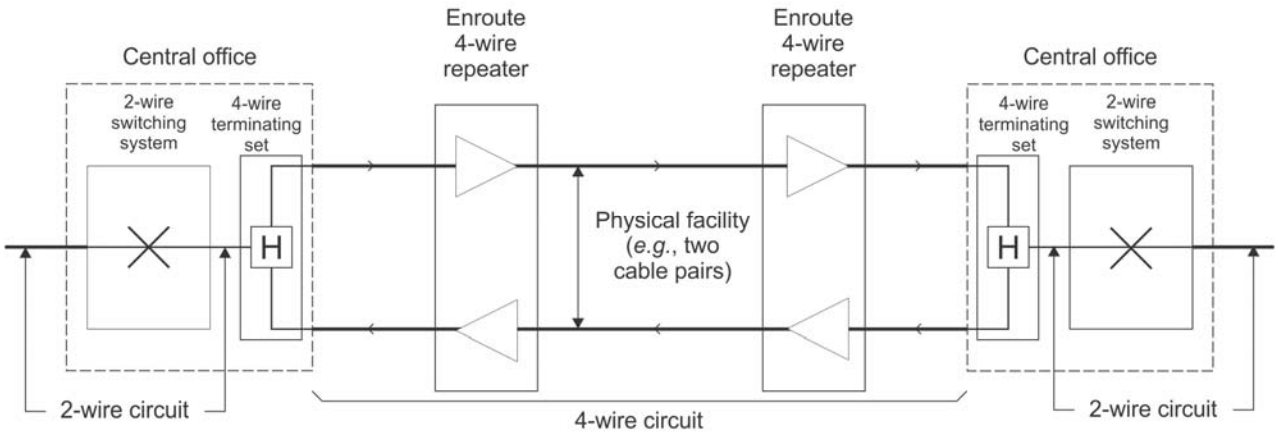


Figure 13. 4-wire physical transmission line

Figure 13 shows the situation I described. The hybrids needed are contained in what are known as “4-wire terminating sets”, which we assume here are located in the two central offices involved. (The symbol for the hybrid, labeled “H”, implies both the hybrid coil circuit proper and the associated balancing network.)

5.4 Multiplex transmission

Starting in about 1930, initially for longer circuits (in the long distance network), multiplex transmission was introduced. Here a pair of wires,

or more likely, two pairs of wires, could carry perhaps 12 transmission circuits, using analog frequency division multiplex. Basically, the speech signal for one channel (in one direction or the other) modulated a carrier frequency, usually using *single sideband suppressed carrier amplitude modulation*.

The object of course was to share the substantial capital and maintenance cost of the pairs (including their supporting infrastructure) over several circuits.

Much was made when presenting this development of the parallel with radio transmission, in which the audio signal also modulated a carrier frequency. And in turn, this new transmission concept was spoken of as "carrier current telephony", and the systems came to be known as "carrier systems".

Ongoing development eventually made this approach economical for shorter circuits (such as for the trunks running between the central offices in a metropolitan area, or the long distance circuits between cities perhaps 35-75 miles apart.

Later, digital multiplex systems were introduced (but they were still nomenclatured as "carrier" systems out of historical inertia).

In all of these systems, the two directions of transmission for a circuit ere entirely separate, and did not inherently interact in any way. Because those were also the key properties of the earlier "4-wire" transmission circuits, the circuits created by these multiplex systems were spoken of as "equivalent 4-wire" circuits (bit often just "4-wire").

And so, just as in the previous example, when a multiplex system was used to provide circuits between 2-wire switching systems, we had the same need to join a 4-wire (albeit "equivalent") path to a 2-wire path through the switching system and to the 2-wire trunk beyond. Thus we have constructed an "equivalent 2-wire circuit".

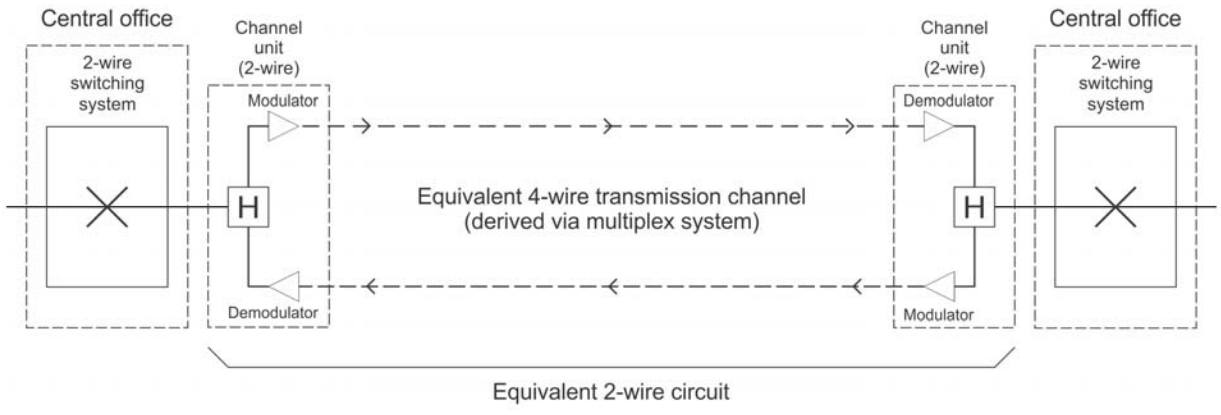


Figure 14. Multiplex transmission circuit

To do this, we called upon our old friend, the hybrid, now as we see in Figure 14.

Here we see a module called a *channel unit*, which is the portion of an analog multiplex system that provides for the separation of the channel's frequency band from the overall transmitted signal, and the modulation of the outgoing channel signal and demodulation of the incoming channel signal.

These modules are available in several types, to accommodate differences in application, and when the channel is used to create an "equivalent 2-wire" circuit (to interface with a 2-wire switching system), it includes a hybrid coil circuit. By the way, here that hybrid is likely implemented with the "single coil" implementation seen in Figure 10.

6 TODAY

Today, of course, the hybrid coil circuit function may be implemented in ways much different from the "classical" implementations I discussed. But these often, at the bottom line, turn out to be to some degree "equivalent" in function to those classical versions.

7 IN A TELEPHONE SET

7.1 The basics

7.1.1 *Common battery operation*

In early telephone sets (and in fact for over half a century), the transmitter (microphone) was almost invariably of the *carbon variable resistance* type. It had to be energized by a DC current before it could convert acoustic waves into an AC electrical signal.

In the earliest telephones, that DC current was most often provided by a battery of dry cells (yes, often of the infamous "number 6" size, each about the size of a small bottle of milk). Of course, the periodic replenishment of these was quite a nuisance to the telephone set owner.

When the telephone companies generally adopted the policy that they, not the user, must own any telephones "on their networks", of course they inherited that nuisance. (There is a fabulous painting of a horse-drawn wagon with zillions of these dry cells, driven by a telephone repairman.)

Eliminating that nuisance was one of the motive for the introduction of the *common battery* concept. There, the telephone line was provided with a DC voltage applied at the telephone central office. The resulting current (when the telephone was "off hook", that is, in use) energized the transmitter in the telephone.

A second advantage (not of direct interest here) was that now the central office could tell, with a relay, when the telephone set on a certain line was "off hook", as only then did any current flow in the line from the applied voltage. This was of course a key to the development of highly efficient telephone switching systems (all be they in the early days manually operated).

7.1.2 *A telephone set circuit*

Figure 15 shows a simplified form of the basic "talking circuit" of an early common battery telephone set.

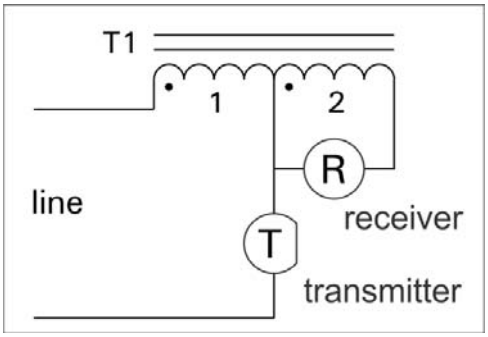


Figure 15. Basic common-battery telephone circuit

We see the transmitter (microphone), T, and the receiver (earpiece), R. We note that DC current (propelled through the line by an applied DC voltage at the central office) can pass through winding 1 of transformer T1 [10] and through the transmitter, thus energizing it. The AC voltage developed across it by the incident acoustic wave from the talker causes an AC current to flow into the line, this signal ultimately being heard by the person at the other end of the connection.

When the distant party spoke, the signal of course arrives over the line as an AC component of the line current. That was coupled through transformer T1 to the receiver R, where is was turned into an acoustic wave heard by the person at this end.

The use of transformer coupling to the receiver accomplished two important things (as compared with just putting the receiver itself in series with the transmitter). Firstly, it avoided the DC component of the line current from flowing through the receiver, where it would have biased the electro-magneto-acoustic mechanism. This could have limited the output of the receiver, and led to distortion at signal amplitudes likely to occur.

---

[10] For historical reasons, this transformer was actually nomenclatured as an *induction coil*.

Secondly, it provided for impedance transformation between the impedance of the line and the impedance of the receiver, allowing the receiver to be made with a lower impedance, less costly to manufacture as it would require fewer turns of wire.

7.1.3 *Sidetone*

It does not take much imagination to recognize that when the person at this end spoke into the transmitter, the AC current in the line that this caused was nicely coupled, by transformer T1, into the receiver in this telephone set. In acoustic terms,. the talker's voice came out of the receiver, a phenomenon called *sidetone*.

Now a certain degree of that has been found desirable. For one thing, although the talker probably does not realize this, the sidetone makes the telephone seem "live" to the talker.[11] For another thing, it provides a type of negative feedback that tends to regulate the potency of the talker's speech, which otherwise would vary more widely among different people and with their state of mind when speaking on a given connection.

But too great a degree of sidetone was not desirable. For one thing, the users just found it annoying. For another, the strong negative feedback it provided tended to regulate the potency of the talkers' speech to too low a level, not desirable for effective communication over the connection.

7.2 The anti-sidetone circuit

7.2.1 *Introduction*

This led to the development of circuits for the telephone set that would reduce the degree of sidetone without diminishing the effectiveness of either outgoing of incoming transmission. There were hundreds of circuits developed for such "anti-sidetone circuits", resulting in a veritable plethora of patents (and of course patent suits).

But it turns out that almost all of them really tried to implement, in a away that would fit into telephone set circuitry, a hybrid coil circuit. Figure 16 shows this in a "black box" way.

---

[11] If the line is interrupted, the loss of the DC current causes the transmitter to stop working, resulting in a loss of the sidetone, and the user will, without really thinking about it, sense that the telephone has "gone dead", which indeed it has.

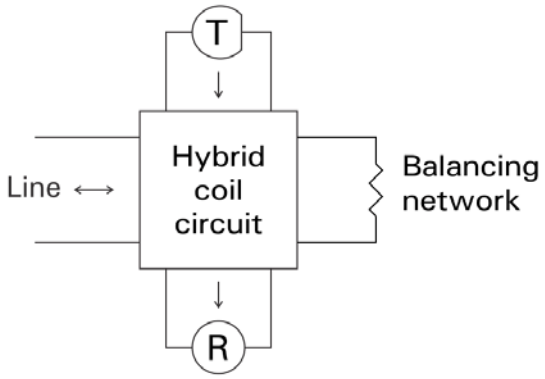


Figure 16. Anti-sidetone operation–concept

7.2.2 *A widespread implementation–the basic circuit*

A particular implementation of that, originally attributed to G.A. Campbell (in perhaps 1906), has been the basis of the anti-sidetone circuits in essentially all Western Electric telephone sets (and many others) since the concept was introduced. Figure 17 shows that circuit concept:

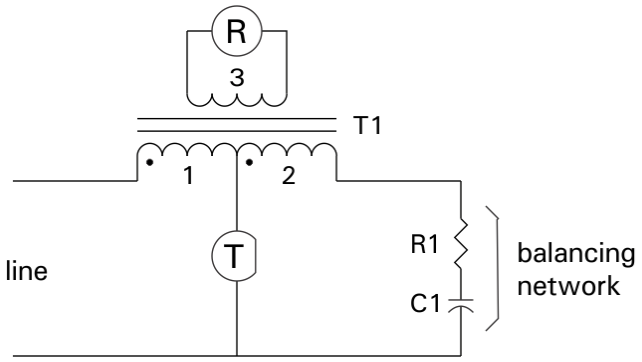


Figure 17. Anti-sidetone circuit

We see that this directly follows the "single coil" hybrid circuit concept seen earlier in Figure 10 (in its basic, "unbalanced", form). We also see that the DC current in the line can still flow through transformer winding 1 and through the transmitter, energizing it.

Note the unsurprising appearance of the *balancing networ*k, in this case a "compromise" network consisting of merely a resistor (R1) and capacitor (C2) in series, that simple approach being apt given that the actual impedance of the millions of telephone lines varies so widely.

Capacitor C1 also plays another important role. It prevents any of the DC current flowing through the line from traveling through transformer windings 1 and 2 and then through the receiver. This would be undesirable for two reasons: it would shunt away from the transmitter part of the DC current from the line on which transmitter operation

depends (thus reducing the "conversion efficiency" of the transmitter), and it would bias the electro-magneto-acoustic system of the receiver.

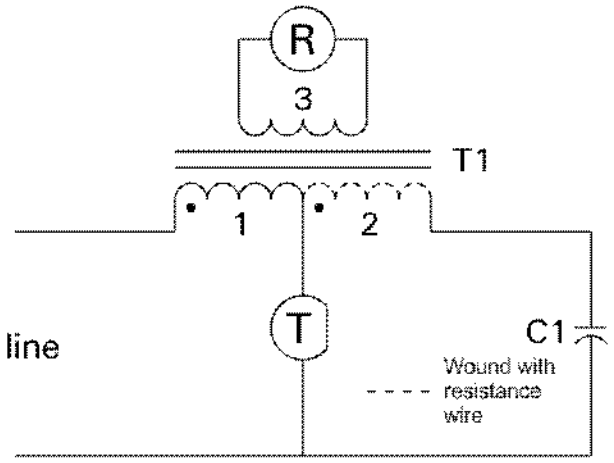


Figure 18. A simplification

7.2.3 *An important simplification*

In the actual implementations of this, several clever ploys were employed to minimize manufacturing cost. One was to make winding 2 of the transformer of resistance wire. The resulting series resistance takes the place of discrete resistor R1. We see this in figure 18.

7.2.4 *A further simplification*

A further simplification was not to have the transformer have a winding 3 at all. Rather, a portion of winding 2 was used (in "autotransformer" fashion") to drive the receiver. We see this in Figure 19.

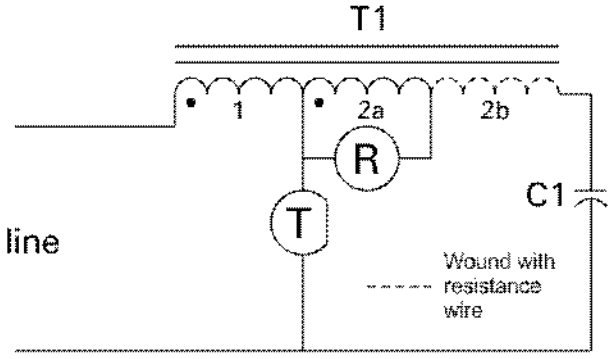


Figure 19. A further simplification

For continuity I have still labeled one winding as winding 2, but now it is in two parts, labeled 2a and 2b.

Of course we don't want any extra resistance included in the circuit to the receiver. So only part 2b of winding 2 is wound with resistance wire.

7.2.5 *Nomenclature*

The "transformer" seen here, when a separate component, just as for the transformer used in earlier telephone sets (not of the "anti-sidetone" variety), was (by "historical inertia") formally nomenclatured as an "induction coil".

In the 500-type telephone set and beyond, this transformer was swept into a module that included the entire transmission circuit (now more complicated but still revolving around the hybrid coil concept described just above) and a separate capacitor used in the ringer circuit. These modules were nomenclatured as "networks" (since they comprised many circuit elements).

-#-