The operation and design of baluns

Brian Collins, BSC Associates Ltd.

Summary

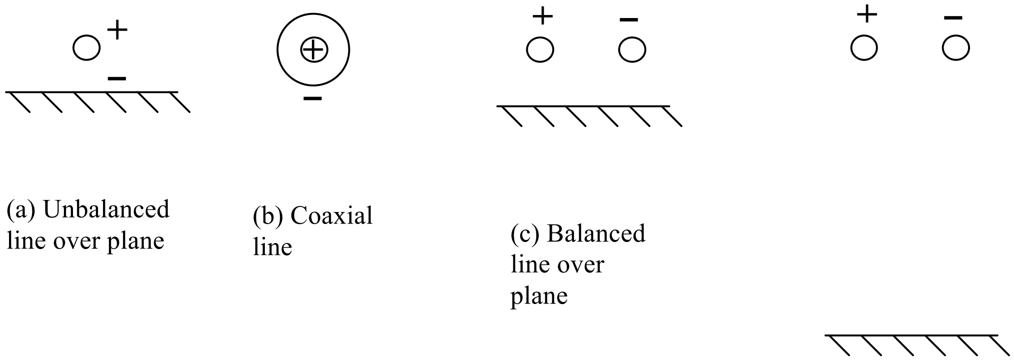
This paper explains in a non-mathematical way the principals and operation of a number of classes of balun used to connect antennas to transmission lines.

Introduction

Most RF engineers will have encountered the use of baluns, used to connect an 'unbalanced' coaxial feedline to an antenna or other device with a pair of 'balanced' input terminals. Over many years the author has noticed that there is a considerable mystique about the real function of these devices, which can be realised in many ways. This paper clearly defines the function of a balun and describes a number of possible physical realisations. The information provided here should allow a designer to create a variety of practical designs. Further variants can be created using one of the standard circuit design programs or electromagnetic simulation tools. This paper is not intended to be a guide to the whole literature of baluns, but is intended to provide some background understanding which is not explained in many published papers.

1 The concepts of balanced and unbalanced transmission lines

The concepts of balanced and unbalanced transmission lines are easy to understand if we regard the lines as having a relationship to a conducting plane.



(a) Unbalanced line over plane

(b) Coaxial line

(c) Balanced line over plane

(c) Balanced line with plane removed

Figure1: Some familiar transmission line geometries

Figure 1 shows some familiar balanced and unbalanced transmission line geometries.

(a) is a single conductor placed over a conducting plane. While this is not used for long lines, this geometry is frequently used within antennas. Current is carried in the elevated line, while an equal return current with opposite phase travels in the conducting plane. The transmission line can be regarded as comprising the physical line together with its image in the conducting plane. The field associated with the transmission of the wave is distributed in the space between the line and the plane, and extends beyond the conductor. The configuration is now familiar as a microstrip line, although the single conductor is usually (but not always) supported on a layer of dielectric whose permittivity constrains the field with the result that most energy is concentrated in a region close to the conductor.

(b) is obtained from (a) by wrapping the conducting plane round the single conductor. The return path is still provided by the conducting plane, but all the energy is now confined within the outer conductor. In a line made using perfect conductors there is no external field. Again, the inner conductor is usually supported by some physical dielectric material, but the presence of this dielectric has no effect on the field configuration within the line. (It does change the capacitance per unit length, and therefore the characteristic impedance).

(c) is less familiar. The two conductors are excited by equal voltages *relative to the conducting plane,* with opposite phases. This is probably not a very familiar configuration, but we can now remove the conducting plane from the immediate vicinity of the conductors, resulting in (d).

(d) maintains the same relationship between the line voltages and the conducting plane. It is a balanced twin line in the configuration that used to be used for HF and MF transmission. As the conductors carry equal and opposite currents, the line is balanced. There is no net current in the underlying plane and there is no net field once the distance from the line is large compared with the line spacing. Even if we remove the conducting plane completely the line is balanced. Furthermore, if we apply a voltage between the lines, the concepts of 'balanced' or 'unbalanced will have no meaning unless we refer to the now distant reference plane (which we now recognise as being 'ground'). The only point in the circuit at which this reference will almost inevitably be made is at the source at one end of the line and the load at the other.

If we consider sources rather than loads, we can imagine that in a number of cases the output is presented in coaxial form – clearly 'unbalanced' because one terminal is grounded and the other is 'live'. In another case we may find a 3-terminal device, clearly –V, 0, +V with an exact 180° phase difference between the two output voltages relative to the central 'ground'. Sometimes there will be a nominal +V and –V, but the third terminal is not available, and it is far from clear as to whether the output voltage is equally divided between the terminals, or just what the relative phase may be.

In the case of loads the situation is perhaps more puzzling. A load may clearly be coaxial (like a coaxial termination) or unbalanced (like an MF mast radiator in which one terminal is clearly at ground potential). It may be obviously balanced (like a folded dipole whose centre point is explicitly grounded) or it may be far from clear what the mode of the termination may be…. and in any case, whether it matters.

So far we have recognised that transmission lines, generators and loads may exist in explicitly 'balanced' or 'unbalanced' forms. Now we examine what happens when these devices are connected together.

2 Connecting a balanced load to an unbalanced line

One of the most telling examples of the connection of an unbalanced line to a balanced load is the connection of a coaxial line to a simple dipole, a situation shown in Figure 2. In this case we simply apply the potential which exists across the end of the coaxial line across the two terminals of the dipole (which themselves are far from ground and have no innate concept of 'balance'. The question is: why is this wrong? In this case, note that even the 'mode' of the source has been left ambiguous, as it doesn't affect the question posed about the dipole.

The answer 'because there is no balun' is not an adequate answer. The question is: what is wrong with the operation of this circuit? In what sense does the dipole insist in behaving as a balanced load, and how does the system behave in an unexpected, and often unacceptable manner when they are connected as shown?

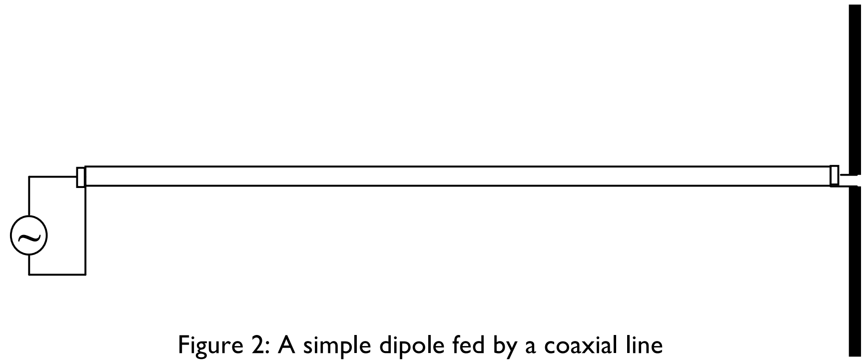


Figure 2: A simple dipole fed by a coaxial line

A clue to the situation is given in Figure 3. Here we see symmetrical stray capacitances existing between the two limbs of the dipole and the outer conductor of the feed cable. This situation seems innocent until we enquire into the voltages existing across these two equal capacitances: these are clearly not equal. In the case of the upper limb the whole feed voltage appears at the lower end of the limb, and even higher voltages as we move away from the centre point. In the case of the lower limb the voltage between it and the outer conductor is zero at the upper end, rising as we move down the limb, away from the point of connection to the coaxial line. The **currents** in these two equal stray capacitances (i_1 and i_2) are therefore **unequal.** Clearly the current emerging from the inner conductor of the cable and that flowing back along the inside of the outer conductor are equal (they could not be otherwise), so where does the difference current flow? The answer is that it must flow along the outside of the outer conductor. We will therefore find that the radiation pattern of the antenna will be affected by the exact position of the cable, and the impedance presented at the input end of the cable will also change as we re-position or coil the cable. Cables connecting different antennas (perhaps transmitting and receiving antennas) will be coupled together by these out-of-balance currents.

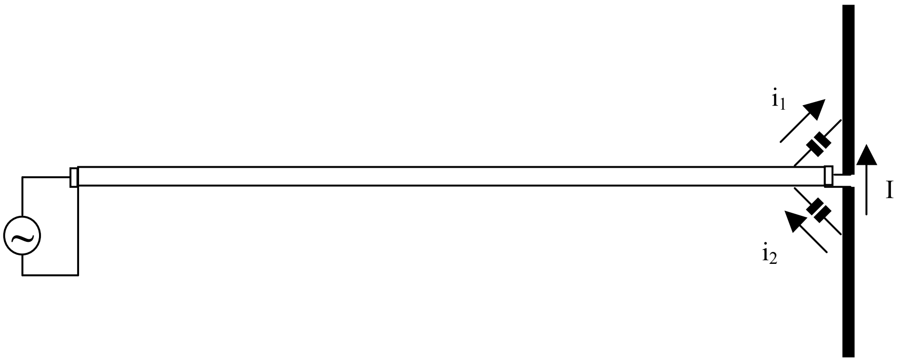


Figure 3: A the circuit of Figure 2 showing stray capacitances at the dipole

It is now clear that we need to introduce some device into the circuit to permit us to legitimately connect the coaxial cable (unbalanced) to the dipole (balanced). This device is a balun, often (but somewhat misleadingly) described as a balanced-to-unbalanced *transformer*. As we will see, a transformer is only one of the possible circuit elements which will perform the required function.

Another way of viewing the situation is by regarding the currents flowing in the limbs of the dipole as being the sum of currents with two different modes (Figure 4).

(a) Balanced mode (push-pull)

(a) Unbalanced mode (push-push)

Figure 4: The total current on a dipole may be regarded as being the sum of a balanced mode (a) and an unbalanced mode (b) which is usually of smaller magnitude

In some applications the push-push current is known as the common mode current. It creates an external field from an apparently symmetrical balanced line (cf Figure 1c), creating losses and reducing the noise immunity of the balanced line.

3 The formats of a balun

In general there are four common general formats for a balun. All perform the function we have defined, but in addition most of the devices have specific impedance transformation characteristics and/or impedance/frequency characteristics. The impedance characteristics of the balun when combined with the impedance/frequency characteristics of the load, often determine the choice of a balun design for a particular application

3.1 Transformers

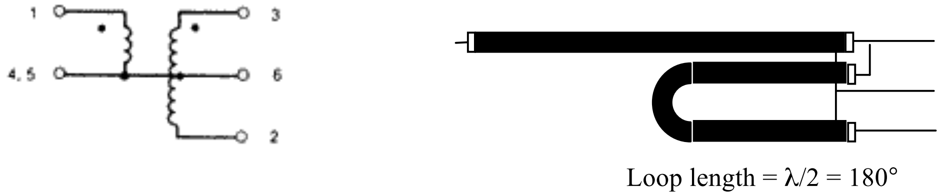


Figure 5: Typical circuits for transformer baluns

This may have the general lumped form (Figure 3a) or an equivalent transmission line form (Fig 3b). In both cases the single-ended unbalanced feed voltage is transformed into a bi-phase voltage with an explicit ground terminal. This ground terminal is connected to the cable sheath or local ground, and the two bi-phase voltages to the terminals of the balanced load impedance. The form of Fig. 3a is very common in the HF band where ferrite-cored transformers are used. There are various forms using one or two windings, and by appropriate choice of the turns ratio a very wide variety of impedance transformations can be obtained, typically in the range 1:1 to 1:12, with the balanced port generally requiring the higher impedance to match a variety of antenna types..

By using low loss ferrites with high saturation flux densities and placing the transformer in a tank of oil, these units can be designed to handle high power levels and the accompanying high peak voltages. Careful design provides exceptionally wide impedance bandwidths, units with operation over the whole HF band from 2 to 30MHz being readily available.

For very high impedances the problem becomes one of limiting or compensating the stray capacitances associated with the windings and the connections. At high frequencies the designs are increasingly limited by losses and stray reactances, and above about 50MHz the coaxial form of the balun (Fig 3b) becomes more common.

The coaxial bi-phase balun of Fig 3b provides a 4:1 impedance transformation between balanced and unbalanced terminals. Its impedance bandwidth and balance ratio bandwidth are both limited. A more general form of the balun has path lengths of l and ($l + \lambda/2$), and the input impedance can be manipulated by the right choice of l and the characteristic impedance of the lines. This balun has a limited bandwidth, as both the input impedance and the balance ratio will change as the input frequency moves away from the design frequency (where the loop is exactly a half wavelength long).

3.2 Devices which apply the feed voltage across an isolated balanced terminal pair

This class of balun creates a balanced physical structure such as a short-circuit balanced stub, and applies the feed voltage across the stub. The coaxial feedline can be attached to one side of the stub, entering from the grounded end, while the balanced load is fed from the two sides of the stub. This structure is usually known as a Pawsey stub or a twin-line balun (Fig 6), The length of the stub is conventionally a quarter wavelength, but this is not essential as long as the susceptance of the parallel stub is taken into consideration – the stub can be used to compensate the reactance/frequency characteristics of the load; by careful choice of parameters very low VSWRs can be obtained from a dipole fed with this style of balun over wide bandwidths. The structure provides a balance ratio which is essentially independent of frequency, as the intrinsic balance does not depend on lines being any specific fraction of a wavelength. (By inspection we can see that the structure meets the test described in Section 2.) The balanced conductors may be made from flat strips, tubes, channels or material of any other cross section provided the mechanical form of the two arms of the sub – including the transmission line – are symmetrical or nearly so.

The impedance ratio of this form of balun is essentially unity, but by creating a tapered balanced line from the twin line conductors, wideband baluns have been created for the HF band with an impedance ratio of around 5:1. They are used from the HF band up to 10GHz and beyond.

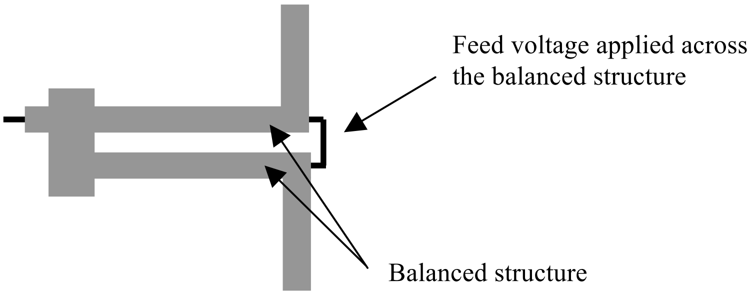


Figure 6: The Pawsey stub balun

A further version of this balun, described by Oltman as the Compensated Balun (3) has an open circuit coaxial stub on the side opposite the coaxial cable. A number of related configurations that use flexible coaxial cables are described by Jasik in his patent of Feb 1960 (5) and by Phelan (6).

The split tube or three-wire balun

A related but quite different form of balun is the split tube balun (Fig. 7). Here the symmetrical stub is excited from within; the stubs form two halves of a coaxial line, and the inner conductor is grounded to one side of the stub. The internal fields are clearly not symmetrical, but the symmetrical impedances between the ends of the stubs and ground creates symmetrical balanced driving voltages at the open-circuit end of the stub. A consideration of the stray capacitances as shown in Figure 2 clearly leads to the conclusion that this device meets the requirements for a true balun.

In many text books this balun is shown as a split coaxial line (Figure 7b), but the format of the line can vary; examples exist using two flat strips with the third conductor emerging between them (Figure 7c), or even as a three-wire configuration in which the 'outer' conductors are of the same dimensions as the 'inner' conductor (Figure 7d).

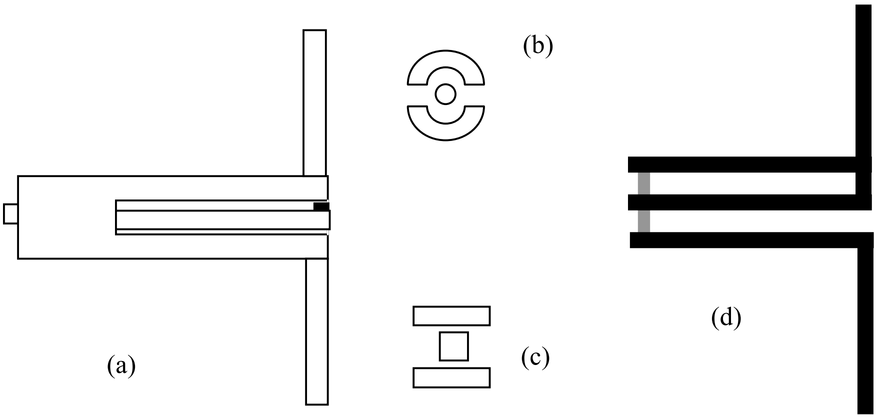


Figure 7: Split-tube and three-wire baluns

The split tube balun has an intrinsic impedance ratio of 4:1 and a wide balance ratio bandwidth. As with the previous form, this balun may be modified to suit the application – the balanced conductors may project beyond the internal feedpoint, and the length and Zo of the balanced lines can be manipulated to provide reactance compensation. Examples of this format can be found from around 50MHz to 10GHz. The method of construction will be chosen to meet the mechanical constraints imposed, perhaps by the weight of the element, or the ease of maintaining the geometry when dealing with very small components.

The Roberts balun

The Roberts or hairpin balun is a derivative of the Pawsey stub which requires no direct connection between the unbalanced feedline and the opposite side of the parallel line stub.

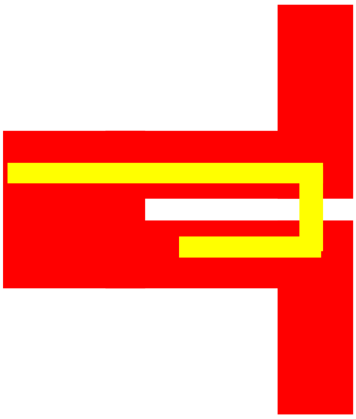


Figure 8: A Roberts 'hairpin' balun applied to a microstrip dipole

The line crosses at the feedpoint, but instead of being directly connected to the other stub conductor, a virtual short-circuit is created by running the line for a quarter wavelength along the stub. This arrangement provides further possibilities for the reactance compensation of the balanced load. The arrangement can be constructed from coaxial line, as originally suggested by Roberts, but it is widely used to feed microstrip dipoles because the arrangement requires no soldered or plated-through connection.

The Marchand balun

This balun is constructed from coupled transmission lines and is often realised as a microstrip device.

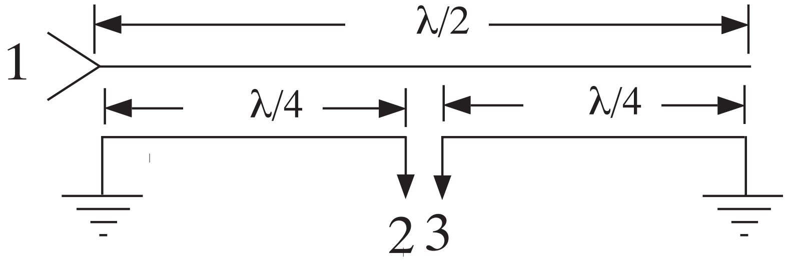


Figure 9: The Marchand balun

There is a subtle relationship between the Marchand balun of Figure 9, the Pawsey stub of Figure 6 and the Roberts balun of Figure 8. The Marchand balun is sometimes realised in a variant in which the balanced output line is a co-planar waveguide which can be used to drive a Vivaldi antenna.

Antennas with no apparent balun

Some antennas appear to contain no balun, but maintain proper conditions for balanced excitation. Examples are the excitation of a folded dipole by a cable passed through half of the element, and the excitation of log-periodic antennas and spirals. Each of these examples are versions of the class of baluns just described but it is the antenna itself which forms the symmetrical balanced structure. In each case the feed line is routed along one arm or branch of the structure and the feed voltage is applied at the symmetrical terminals. To obtain exact balance a dummy cable is often fed along the other arm of the antenna so the profile of the conductors on both sides of the antenna are maintained in exact physical symmetry.

3.3 Devices which suppress the out-of-balance current

Instead of transforming the applied voltage into a balanced mode, an unbalanced excitation can be used as long as the impedance presented to any unbalanced currents is high enough to reduce their amplitude so only a negligible current can flow.

The most obvious example of this form of balun is a coaxial cable wound round a ferrite core. Other examples are the quarter-wave coaxial sleeve (Bazooka) and the bifilar inductor which presents a low series reactance to a balanced (push-pull) current, but a high impedance to any unbalanced (push-push) current.

3.4 Progressive transformation of mode

A further class of balun depends on the property of transmission lines that if changes are made progressively along the line and there is no single point of disjunction, the line appears as a smooth transition. This is especially true when the transition is made progressively over several wavelengths. Tapered line sections are familiar as impedance transformers, they can have very wide frequency bandwidths and are frequently used to drive antennas such as wideband spirals..

In a similar way a microstrip line can be transformed from an unbalance device to a balanced device by tapering the ground conductor until it has the same dimensions as the upper conductor. Or a twin balanced line can be transformed by progressively widening one conductor and wrapping it round the other until it embraces it completely as a coaxial outer conductor (Ref. 4). Both these formats create impedance transformations which depend on the line geometries used, and if they are shortened both the impedance and balance-ratio bandwidths will be reduced.

4 Balance ratio

The effectiveness with which a balun performs its task may be described by its balance ratio. Referring to Figure 10, this is defined as:

Balance ratio $B = 20\log_{10} |(\mathbf{V}_1 + \mathbf{V}_2)/(\mathbf{V}_1 - \mathbf{V}_2)|$
where V1 and V2 are complex voltages

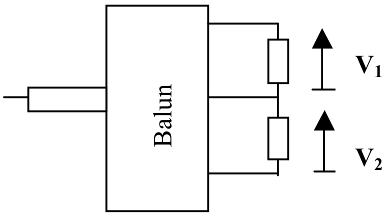


Figure 10: Definition of balance ratio. The two load impedances are of identical value.

It is unusual to see balance ratio explicitly specified, but is it the only measure of the 'goodness' of a balun. A balance ratio in excess of 20dB would usually be considered to be acceptable for most purposes, but when operating at very high powers (as with HF broadcast antennas) a higher value may be desirable to avoid losses occurring in conductors that were not intended to carry significant RF currents.

It is relatively simple to measure the balance ratio of a metal-cased balun with an available ground reference on the balanced side; it is less easy to measure the balance ratio of a device like a bazooka balun where the balanced terminals are far from ground. This may be a case where ratio of balanced and out-of balance currents may be more appropriate, but it is difficult to make such a measurement without disturbing the currents that are being measured.

5 Baluns in antenna measurements

When measuring the performance of a small antenna, especially an unbalanced antenna, it is desirable to connect the measurement cable via a sleeve choke (which as we have seen is a form of balun). This suppresses radiating currents in the feed cable, which would otherwise change the characteristics of the antenna under test [7]. This is not a case of the lack of a a balanced to unbalanced transformation, but the device delimits the position at which radiating currents may flow in the groundplane of the DUT. In the absence of this, the unbalanced feed can excite radiating currents in the outside of the cable and these create instability in antenna properties.

A ferrite ring or bead can also be used to suppress the currents in the cable outer conductor, but at higher frequencies the loss caused by the ferrite (usually not accurately known) can make gain measurements pessimistic.

6 Balanced screened loop

Figure. 11: Balanced screened loop

The balanced screened loop is a useful device for measurement of currents and radiated fields of antennas and of complete electronics assemblies. There are several possible configurations for this, but the one I prefer [7] is shown in Fug. 11. A piece of semi-rigid UT047 carefully wrapped round a pencil provides a loop useable up to around 2.5GHz. The loop is held with its open end close to and co-planar with the conductor on which the field is to be measured (so the H-field passes through the loop). To measure the currents in the elements of an array the distance from the open end of the loop to the antenna element should be around 10mm – easily controlled with the aid of a small block of PS or PE foam. Because the feed cable is also exposed to the fields being measured, it is prudent to add a choke balun on the cable close to the loop to suppress induced currents.

If used to probe currents on PCBs the open end of the loop, which may come in contact with circuit conductors, should be insulated to avoid short-circuit damage to the DUT.

7 Conclusion

Baluns appear in many situations. This short article has explained some of the considerations involved in balun design and provided a basic understanding of their function. New configurations will continue to emerge as techniques evolve, but most of them will conform with one of the formats that has been described.

Brian S Collins BSc(Eng), CEng, FIET, MIEEE, G3MXA
BSC Associates Ltd, Engineering Consultancy Services, Ely, UK.
brian.collins@ieee.org

Version 3, January 2015 (Section 6 and [8] added.)

Web: www.bscassociates.co.uk

References

1. N Marchand, Transmission Line conversion transformers, *Electronics,* vol 17, pp 142 – 145, December 1944.
2. W K Roberts, A new wideband balun, *Proc IRE,* Vol 45, Dec 1957, pp1628 – 1631.
3 G Oltman, The Compensated Balun, *IEEE Trans MTT*, Vol. MTT-14 No 3, March 1966, pp 112 – 119.

4. J W Duncan, 100:1 Bandwidth balun transformer, *Proc IRE*, Feb. 1960, pp 156 – 164.
5. H Jasik, Broad band balun transformers, *US Patent 2 925 566*, Feb 1960.
6. H R Phelan, A wide-band parallel-connected balun, *IEEE Trans MTT*, Vol MTT-18 No 5, May 1970, pp 259 – 263.
7. B S Collins and S Saario, The use of baluns for measurements on antennas mounted on small groundplanes, *IEEE International Workshop on Antenna Technology: Small Antennas and Novel Metamaterials,* iWAT 2005, Singapore, pp 266 - 269.
8. H. Whiteside, R. W. P. King, The loop antenna as a probe, *IEEE Trans. AP, Vol. AP12, May 1964, pp. 291-297*.

I would be very pleased to hear from an author who originated any of the devices I have mentioned without attribution. BSC