THE QUEEN'S AWARDS FOR ENTERPRISE 2001

BBC

R&D White Paper

WHP 132

March 2006

Dipole antennas

C. Gandy

Research & Development
BRITISH BROADCASTING CORPORATION

BBC Research & Development
White Paper WHP 132

Dipole antennas

Chris Gandy

Spectrum Planning Group

Abstract

Much of the work of Spectrum Planning Group is based on well established principles of 'antennas and propagation'. Propagation matters have received a lot of attention over the years and much has been written on the topic but the 'antennas' part, although arguably simpler, may lack a suitable written reference that gathers much of the relevant material into one place. The primary purpose of this White Paper is to take a step towards correcting this state of affairs, but a huge amount of underlying complexity comes to light when one 'scratches the surface' of a topic like this. A secondary objective, then, is to present some of the more interesting and useful facts, and to provide simple, physical explanations wherever possible.

In summary, this White Paper is intended to provide a compendium of well-known and obscure facts about dipole antennas and isotropic sources including effective aperture areas, coupling factors, radiation resistances and link budget equations. What began as a simple case of looking up the effective aperture area of a dipole developed into a quest for the whole story about dipole antennas and has resulted in this re-working of a small part of a very-old, but intriguing story.

Of course, spectrum planning for terrestrial broadcasting involves many more-complicated antennas such as log-periodic dipole arrays and Yagi-Uda arrays, not to mention the multitude of clever transmitting antenna designs based on slots. It wouldn't be possible to do these justice as well in a document of this size so I've deliberately limited the scope.

Additional key words: antenna, dipole, antenna gain, effective aperture area, link budget, radiation resistance, cage dipole, folded dipole, balun, monopole, biconical dipole, discone

White Papers are distributed freely on request.

Authorisation of the Chief Scientist is required for publication.

© BBC 2005. All rights reserved. Except as provided below, no part of this document may be reproduced in any material form (including photocopying or storing it in any medium by electronic means) without the prior written permission of BBC Research & Development except in accordance with the provisions of the (UK) Copyright, Designs and Patents Act 1988.

The BBC grants permission to individuals and organisations to make copies of the entire document (including this copyright notice) for their own internal use. No copies of this document may be published, distributed or made available to third parties whether by paper, electronic or other means without the BBC's prior written permission. Where necessary, third parties should be directed to the relevant page on BBC's website at http://www.bbc.co.uk/rd/pubs/whp for a copy of this document.

Dipole antennas

Chris Gandy

Spectrum Planning Group

1. Preface

The full title of this White Paper should be 'A compendium of well-known and obscure facts about dipole antennas and isotropic sources including effective aperture areas, coupling factors, radiation resistances and link budget equations' but I abbreviated it for the sake of presentation. It is mainly about dipole antennas, and VHF/UHF ones at that, but not just the types used for broadcasting.

There seems to be a long tradition of staff in BBC R&D (and Research Department before 1993 [1]) writing documents about antennas, often reproducing material that can be found in text books but usually presenting it in a different way in order to support some specific motive – not just education about the topic in general. This White Paper conforms to that tradition and I claim no originality for the content. Like Eric Morecambe's encounter with André Previn ("I'm playing the right notes sunshine …"), the parts of the story aren't always presented together in this order, and sometimes the order in which a story is unfolded can affect how easily it is comprehended. Also, I'll try wherever I can to add some value by justifying or challenging common 'knowledge' rather than just accepting it (e.g. can you explain why the majority of dipoles in use are folded ones?) although I won't attempt to derive everything from first principles. Unlike some of the well known authors on this topic, I will try my hardest *not* to leave derivations of important background as 'exercises for the reader'.

On a scale from dissertation to 'Haynes manual', I aim to pitch this document somewhere in the middle so its content will be accessible to people who make and use antennas. Much as I admire the universal applicability of Maxwell's equations, I won't use vector calculus here because I believe that would defeat this objective. Also, that degree of generality may not always be needed when the topic is this specialised. I will try as far as I can to present explanations in straightforward electrical engineering terms although, I'd have to admit, along the way I came across a few oddities that still baffle me!

Although most of what I'll present here is standard text-book material, for completeness I needed to deduce for myself some of the more obscure bits because either I didn't understand the explanations provided in the literature or no explanation was forthcoming. Even in these cases I'm confident that I've no more than stumbled upon paths many others have trodden before. An example is my derivation of the radiation resistance of a folded dipole and my treatment of the $\lambda/2$ transmission-line 'balun' that goes with it.

In addition, I'll mention some particularly interesting but, possibly, little known designs of dipole antenna that have stood the test of time, for forty or fifty years, but are now in danger of slipping, silently, into obscurity. The names by which I'll refer to these may not be the same as those that have been used in the relevant industry.

I will present this document in two parts, in a similar manner to my aged White Paper about DAB [2]. In the first part I'll deal with the use of dipole antennas and isotropic sources, particularly how their characteristics are represented in everyday link budget calculations, without getting too much into the physics of how antennas work. In the second part I'll delve deeper into the electrical engineering of dipole antennas, of several different types, and touch on matters of 'how' and 'why'. Unlike that earlier White Paper, I'll use a simple, sequential, section-numbering scheme.

There's not much here to enthral the antenna cognoscenti, but please do read on – one or two of the areas I'll venture into that are less well-covered in the standard text books might raise a smile!

PART 1: characteristics of isotropic sources and λ/2 dipole antennas

2. Introduction

There are many excellent text books on the topic of antennas and almost anything you could wish to know about simple forms of antenna, like the dipole, can be found – if you know where to look, or you have the time and inclination to read some of the books cover-to-cover. The book 'Antennas' by John Kraus [3] contains all the working needed to derive expressions for the effective aperture and the coupling factor (i.e. 'k-factor' or 'effective length') of a balanced, thin, linear, resonant $\lambda/2$ dipole antenna, although the full derivation takes a somewhat convoluted path through the chapters of the book. Kraus also provides working that can be applied to a hypothetical isotropic source (or 'point source'), used as a reference against which the gains of many types of antenna are measured. It is instructive to compare results for the two cases and to develop the working up to practical link budgets (actually, I will go as far as the equations on which link budgets are based).

An important principle to keep in mind when reading books about antennas, and this document, is that this field of engineering is full of small approximations that usually don't matter much. In this Part I will print in *italic typeface* numerical values that contain small rounding errors that would be inconsequential in most circumstances. I will also use italics throughout for names of variables like E and A, to indicate *emphasis* and for occasional, well-known, bits of Latin.

Many stages of the development will require liberal interchange between the transmitting and receiving cases, assuming reciprocity without question, and where I present illustrations I'll hop between horizontal and vertical 'polarisation' to suit the page formatting.

3. Radiation resistance

In the transmitting case, power is fed into an antenna and, if it is well matched, most of it doesn't come back – it's either radiated as an electromagnetic wave or it's dissipated as heat. This is analogous to feeding power into a pair of resistors; let's say, connected in series. Their relative values determine the proportions of the power they absorb and, in a matched system, the sum of their resistances matches the internal resistance of the power source (which may be altered up or down using some kind of transformer in order to achieve a match). The value of the resistor that accounts for power transferred to the outgoing electromagnetic wave is the radiation resistance of the antenna. Generally, the smaller the radiation resistance of an antenna the less efficient it is because a greater proportion of the power is developed in the other, loss, resistance. Indeed, the *radiation efficiency* can be defined as the ratio of the radiation resistance to the total terminal resistance. However, large values of radiation resistance often go hand-in-hand with large rates of change of the value with frequency and the practical upper limit for efficient matching is probably little more than a kilohm.

When an antenna is receiving, an EMF is induced in it by the incoming electromagnetic wave in which the antenna is immersed, and this EMF is presented to the 'down-lead' connecting it to the receiver. Every generator has an internal resistance and in this case it is the radiation resistance (in series with the loss resistance, but in the following sections I'll assume this has a value of zero Ω).

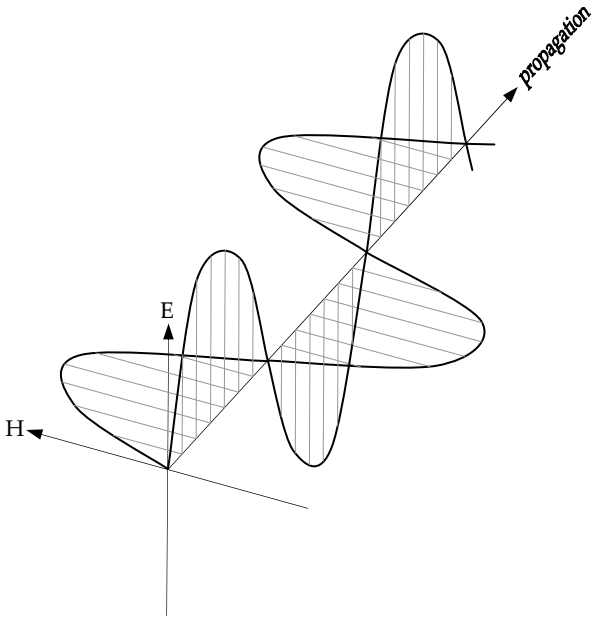
Of course, an antenna can be modelled just as well using shunt admittances rather than series impedances, in which case modelling the generator as a current source simplifies calculations – the principle of Thévenin and Norton equivalent circuits. However, I suspect I'm not the only engineer who considers constant voltage sources much more intuitive than constant current sources. After all, how do we most often encounter sources of electricity? Batteries and the mains: both well known voltage sources with nominally 'constant' values like 1.5 V, 9 V or 12 V, and 230 V RMS ±10%. When's the last time you encountered a practical current source, capable of developing whatever voltage it takes to maintain the specified constant current? I remember NiCd batteries liked them and I vaguely recall they might have been used for powering undersea co-axial cable repeater amplifiers.

In general, the impedance an antenna presents to its terminals contains reactive and resistive parts that both vary with frequency. Some types of antenna, for instance those involving conical elements, and generally "fat" antennas, are designed to minimise this variation over a wide band. Some other types, such as thin dipoles, exhibit distinct resonances and are usually operated at or around one of their resonant frequencies where the reactive part of their terminal impedance is small, in order to simplify matching.

4. Effective aperture area of an isotropic source

A practical transmitting antenna produces a multitude of different electric (E) and magnetic (H) field components distributed differently around the antenna, with different rates of decay with distance from it, oriented in different directions (e.g. the direction in which the E field would apply a force to a charge), and with different phase relationships relative to the alternating current fed to the antenna terminals. So, generally, the aggregate quantities E and H need to be represented by vectors in space and phasors in time and their magnitudes are quantified in units of V / m and A / m, respectively.

A large distance away from the antenna, in a region know as the 'far field', the only two remaining components of significant strength are the E and H fields that form a 'plane wave' – in phase with each other in time, oriented perpendicularly to one another and mutually perpendicular to the direction of propagation, radially out from the antenna. I have illustrated these to the right for a vertically-polarised wave using a pair of sinusoidal waveforms. The hatching is meant to identify the perpendicular orientations of the fields. If the plane in which I've shown E and H axes is visualised as moving in the direction of propagation, the instantaneous magnitude and sign of each field is indicated by where each sine wave cuts this transverse plane. Consequently a plane wave can also be termed a Transverse Electro-Magnetic or TEM wave. The magnitudes of E and H in a plane wave are related by Z, the intrinsic impedance of the propagating medium (see the box below: the intrinsic impedance of space). For simple antennas, the transition into the far field occurs at a distance from the antenna $r > 2\pi\lambda$, where λ is the wavelength in use.



At any distance from the antenna, the strength of the outgoing wave can be found by taking the vector cross product of whatever E and H components are present, with geometrical reference to the normal vector that radiates out from the antenna (i.e. a vector in the direction of propagation). This resolves those parts of E and H that are in time phase and oriented as above, and the result is another vector known as the 'Poynting vector'. The magnitude of the Poynting vector is the product of the strengths of the plane-wave E and H fields so it has the dimension W / m^2 and represents the power density $S = P / a$; a power P radiated through an area a. Considering magnitudes, $S = EH$ and $E / H = Z$ so $S = E^2 / Z$ and then $P = E^2 a / Z$.

For a hypothetical lossless isotropic transmitting antenna, at a distance r away the power is radiated uniformly through a sphere of surface area $4\pi r^2$ producing a uniform electric field of strength E_r related to P by:

$$P = \frac{E_r{}^2 4\pi r^2}{Z}$$

… and, because $E^2 \propto 1/r^2$ or $S \propto 1/r^2$ for a given value of P, this illustrates the principle known as 'spherical spreading'.

Now consider for a moment a lossless (uni-)directional transmitting antenna. Directional properties are achieved by interference between waves originating from multiple different sources, or points in a

the intrinsic impedance of space

An electromagnetic wave propagating in space, or any other medium, gives rise to electric and magnetic fields which can each be considered a manifestation of the same thing: the electromagnetic wave.

The relationship between the strengths of the electric and magnetic fields, $Z = E/H$, has the dimension of electrical resistance and is known as the intrinsic impedance of the medium [6]. For free space:

$$Z = \sqrt{\mu_0/\varepsilon_0} \quad \text{and} \quad c = 1/\sqrt{\mu_0\varepsilon_0} \quad \text{so} \quad Z = 1/(\varepsilon_0 c)$$

… where $\varepsilon_0 = 8.85\times10^{-12}$ (Fm^{-1}) and the velocity of light $c = 3\times10^8$ (ms^{-1}), both initially evaluated by measurement. So the intrinsic impedance of space, $Z = 120\pi$ (Ω) or $377\,\Omega$ and the error in applying this value to air instead of a vacuum is miniscule.

The value of $\mu_0 = 4\pi\times10^{-7}$ (Hm^{-1}) is often stated as being 'by definition' implying historically that the value of μ_0 followed from the definition of the expression for c.

continuous source distribution. Generally, the sources or continuous distribution are arranged in the direction of propagation, 'end fire' like a Yagi-Uda array, or normal to it, 'broadside' like the obvious aperture of a waveguide horn antenna or a dish reflector. The interference reinforces the strength of the resulting Poynting vector in some directions and reduces it in others.

For the case of a planar aperture, if the input power P is distributed uniformly throughout its area A giving rise to a uniform Poynting vector in the space immediately in front of the aperture (plus other field components that I'll disregard), the electric field component of this has a magnitude E_A, and:

$$P = \frac{{E_A}^2 A}{Z}$$

If the radiated plane wave has an electric field strength E_r at a distance r from the aperture in the direction of maximum radiation, Kraus [4] shows this is related to the aperture field E_A by:

$$E_r = \frac{E_A A}{r\lambda}$$

… which results from integration of contributions from elements of the aperture and is a development of Kraus's treatment of continuous arrays of discrete sources [5]. The E_A/r factor accounts for spherical spreading and the A/λ part represents the 'electrical' size of the aperture – interestingly, this factor cannot be A/λ^2. Working in the other direction, back towards the aperture:

$$E_A = \frac{E_r r\lambda}{A}$$

Of course, the field strength in each case is assumed to vary sinusoidally with time so it could be written[†] as $E_0 e^{j\omega t}$ where E_0 is the 'amplitude' or peak magnitude. However, the usual meaning of 'power' is average power which is proportional to the square of the RMS field strength, or current, or voltage, so in this context E_A and all the other variants of E are RMS values.

Equating the expressions for P: $\frac{{E_A}^2 A}{Z} = \frac{{E_r}^2 4\pi r^2}{Z}$ so $A = \frac{{E_r}^2 4\pi r^2}{{E_A}^2}$

… and substituting for E_A: $A = \frac{{E_r}^2 4\pi r^2 A^2}{{E_r}^2 r^2 \lambda^2} = \frac{\lambda^2}{4\pi}$

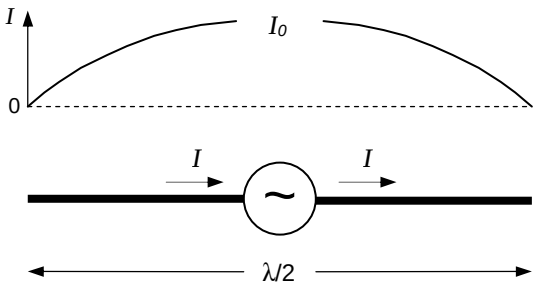
This means, hypothetically, if a planar aperture were assigned to an isotropic source in order to represent its radiation or reception characteristics in a manner suitable for link budget calculations, for example, its effective area would be:

$$A = \frac{\lambda^2}{4\pi}$$

This is a useful reference against which practical, non-isotropic antennas can be compared but I'm not suggesting a practical isotropic antenna could be designed, let alone one based on a planar aperture.

5. Effective aperture area of a λ/2 dipole

Kraus [7] considers a $\lambda/2$ dipole antenna illuminated by a uniform plane wave which he assumes to induce a current on the dipole with a sinusoidal distribution along its length. Apparently, this is a close approximation to most of the natural distribution for a thin antenna [8] and many of the standard text books on antennas state that this is supported by experimental results. In many respects a centre-fed dipole behaves like a balanced transmission line, open circuit at the far end(s), so a voltage applied (or induced) at the terminals gives rise to a standing wave which must have zero current at the open-circuit ends. Assuming a simple sinusoidal current distribution, the result would be as illustrated to the left. Note that the currents in both



[†] Actually, $E_0 e^{j\omega t} = \cos\omega t + j\sin\omega t$ but usually the real (more common) or imaginary part alone is taken to represent something like a field strength that cannot be instantaneously complex.

limbs of a balanced, linear $\lambda/2$ dipole flow, geometrically, in the same direction whereas, considered as a transmission line, these currents in the two conductors can be said to flow in opposite directions (towards and away from the source). The existence of two distinct current 'modes' operating simultaneously will be discussed further in Part 2.

For each infinitesimal element of the dipole's length Kraus assumes the elemental induced voltage and current are proportional to one another and that their ratio is constant (*viz* the average characteristic impedance of the 'transmission line'). He integrates over the length of the dipole and arrives at:

$$V = E\frac{\lambda}{\pi}$$

… where V is the EMF induced at the centre feed-point of the dipole by the appropriately-polarised incident field of magnitude E.

If the incident plane wave has a power density S and the dipole has an effective aperture area A, assuming perfect matching and no losses, the power extracted from the wave $P = SA$, so [9]:

$$A = \frac{P}{S} \text{ and, as before } S = \frac{E^2}{Z}$$

Perfect matching implies that reactive components are cancelled and the dipole is effectively terminated by a resistance equal to its own radiation resistance R_r. The induced EMF is then developed equally across the radiation resistance and the termination resistance so the power P developed in the termination is:

$$P = \frac{V^2}{4R_r}$$

Then:

$$A = \frac{V^2}{4SR_r}$$

Substituting for V and S:

$$A = \frac{E^2\lambda^2 Z}{4\pi^2 E^2 R_r} = \frac{\lambda^2 Z}{4\pi^2 R_r}$$

Also, Kraus [10] evaluates the radiation resistance R_r of an infinitely-thin $\lambda/2$ dipole as $73\,\Omega$. This begins [11] with derivation of an expression for the H field around the transmitting dipole, which requires knowledge of the current distribution along its length, which is assumed to be sinusoidal.

In the far-field, $E/H = Z$ so the Poynting vector can be expressed in terms of the H component alone [12]. The maximum H-field strength is related to the maximum current I_0 in the centre of the dipole by the expression $H = I_0/2\pi r$ [13] derived from the Biot-Savart law. Integrating the magnitude of the Poynting vector over a large sphere enclosing the antenna [14] then yields the power radiated which, assuming no losses, can be equated to the power input to the dipole $P = (I_0/\sqrt{2})^2 R_r$, and thereby R_r is evaluated. The $\sqrt{2}$ relationship between the RMS and peak values appears here because I_0 is the amplitude of the current as well as the peak in the distribution along the length of the dipole.

An alternative approach is known as the 'induced EMF method' [15, 16] in which two identical $\lambda/2$ dipoles are considered, parallel and separated, and an expression is developed for their mutual impedance, that is, the EMF induced in one (appearing between its terminals) for a given current driven through the terminals of the other. This involves calculation of the E-field radiated by the first dipole and then the EMF induced in the second dipole by this field, as before. The distance between the two antennas is allowed to approach zero and the ratio of the terminal EMF to the terminal current tends towards the terminal impedance of the (single) antenna. Apparently [17], it can be demonstrated that this method is equivalent to the 'integral of Poynting vector' method.

With the formulation Kraus uses, the actual value of $73\,\Omega$ contains a specific evaluation of the 'Cosine integral' as well as the 120 from Z so, unlike the value for Z, it is not just a simple combination of fundamental constants. The full terminal impedance of a lossless thin $\lambda/2$ dipole is $73 + j42.5\,\Omega$ [18] and Kraus mentions that in practice the length of the dipole can be reduced a little to remove the inductive reactance appearing in series with the radiation resistance, making the resulting value about $70\,\Omega$. Indeed, the length can be reduced further to provide a terminal resistance less than $70\,\Omega$ at the expense of introducing an effective series capacitive reactance and I'll return to this topic later.

It should be kept in mind that practical dipole antennas for VHF/UHF are often made of stout material and may not be so 'thin', so the relevance of these specific values is limited.

For an antenna of arbitrary shape and size, a numerical solution is often appropriate using one of the many computer programs available nowadays (see the box below: NEC and the 'method of moments').

Inserting values for Z and R_r :

$$A = \frac{\lambda^2 120\pi}{4\pi^2 73}$$

$$\therefore \quad A = \frac{120}{73} \cdot \frac{\lambda^2}{4\pi} \quad \text{or} \quad 1.644\frac{\lambda^2}{4\pi}$$

In comparison with the result for an isotropic source, the ratio 1.644 is the origin of the well-known 'gain' of a $\lambda/2$ dipole over an isotropic source.

$$10\log_{10}1.644 = 2.16 \text{ dBi}$$

The gain of an antenna relative to an isotropic source is usually written with the unit 'dBi'. For some reason, probably rounding, this particular value is given as 2.15 dBi (or a ratio of 1.64) in some textbooks [e.g. 22] and this number had been firmly etched into my memory. I'll write the ratio as 1.64 hereafter.

It may be of interest to note that since 'the induced EMF is then developed equally across the radiation resistance and the termination resistance', as I stated earlier in this section, an amount of power equal to P is 'dissipated' in the radiation resistance. Kraus explains [23] that this power is re-radiated or 'scattered' by the antenna and is an inevitable consequence of placing a single dipole ($\lambda/2$ or shorter) in an EM wave. As he writes "under conditions of maximum power transfer, as much power is dissipated in the generator as is delivered to the load" – evidently whether or not the generator has physical terminals.

Kraus mentions [24] that this scattering phenomenon does not apply in general but goes on to show [25] that it does apply, equally, to a large sheet of RF-absorbing material backed by a conducting sheet, $\lambda/4$ behind it. He states [26] "the incident wave front is disturbed and the energy flow redirected over an area twice the area A". This isn't surprising because, considering diffraction, some effect on the incident wave-front would be expected in a region surrounding the perimeter of the sheet – or immediately above a 'knife edge'; a case well-known to spectrum planners. Cases where this principle does not apply include an absorbing sheet of infinite size and, presumably (and hopefully!), a fully anechoic range.

NEC and the 'method of moments'

For the general case of a dipole of any thickness, the current distribution cannot be assumed sinusoidal but can be found by solving an 'integral equation', where the integral of an unknown quantity equals a known quantity. Analytical solution is only straightforward for a filament. A much-favoured numerical approach is the 'method of moments' on which software tools like NEC (the Numerical Electromagnetics Code) [19] and many derivatives are based. In this method [20], the antenna is divided along its length into a number of 'segments' over each of which the current is taken to vary in accordance with some simple relationship to distance along the segment; a so-called 'basis function' (e.g. uniform or triangular, at the simplest).

The effect of the current is represented by a single 'source point' on each segment, rather like the centre-of-gravity of an object represents its distributed mass at a single point. Then, to some arbitrary 'observation point' the effect of the current in each segment is a simple function of the distance from the source point and the effective value of the current – hence the term 'moment' (as encountered in mechanics).

A set of 'observation points' is then defined, rather like gathering a 3-D view of an object by looking at it from several different viewpoints. If this set has the same number as the set of 'segments', the resulting matrix, of what are effectively simultaneous linear equations, is usefully made square. Solution of the integral has then been transformed into matrix inversion which is amenable to computation for sets up to 10,000 or so [21].

Successful solution also requires application of a 'boundary condition' to one side of the matrix equation. If this describes the current distribution in an antenna when immersed in a uniform plane wave then a candidate may be that the value of E is constant for all observation points in the same plane transverse to the direction of propagation. Another is that the tangential E-field strength is zero on the surface of a conductor.

Solution is then achieved numerically, and accuracy can be traded for processing time by increasing the number of 'segments'. However, the practical implementation is much, much more complicated than this!

6. Coupling factor of a λ/2 dipole

Taking the $\lambda/2$ dipole first because the solution has already been given in the previous section, $V = E\lambda/\pi$ where V is the EMF induced at the centre of the dipole by a uniform incident field of strength E with appropriate polarisation.

What seems to have become known in BBC R&D as the 'effective length' of the dipole, which Kraus [27] and others call the 'effective height' h (probably from consideration of monopoles built on the ground), is then given by:

$$h = \frac{V}{E} = \frac{\lambda}{\pi} \text{ (m)}$$

… and this facilitates calculation of the signal voltage presented to a receiver from a given incident field strength, usually both RMS values.

Assuming no losses, the signal voltage V_s presented to a perfectly matched termination (e.g. receiver) is half the EMF, so:

$$V_s = E\frac{\lambda}{\pi}\cdot\frac{1}{2}$$

It is often useful to work with expressions like this in decibel form, and this is made straightforward by always dealing with quantities directly related to power, hence:

$$V_s^{\,2} = E^2\left(\frac{\lambda}{\pi}\cdot\frac{1}{2}\right)^2$$

In decibels: $20\log_{10} V_s = 20\log_{10} E + 20\log_{10}\lambda - 20\log_{10} 2\pi$

Then, for example (because, alternatively, one might wish to express the field strength in dBmV/m or some other unit):

$$V_s\,(\text{dB}\mu\text{V}) = E\,(\text{dB}\mu\text{V/m}) + 20\log_{10}\lambda \; - 16$$

… with λ expressed in metres. Any known power gains or power losses, expressed in dB, can easily be included in this sum.

7. Link budgets involving antennas related to a λ/2 dipole

In link budgets involving dipole antennas or antennas based on arrays of dipoles, as apply to terrestrial radio and television broadcasting at VHF and UHF, it is conventional to express antenna gains in dBd, meaning decibels above the 'gain' or coupling factor of single $\lambda/2$ dipole. This dimensionless 'gain' can be interpreted as the degree to which the antenna beam is further confined (e.g. in terms of solid angle) or as the relative size of the effective aperture area.

Since the effective aperture area of a $\lambda/2$ dipole is:

$$A = 1.64\frac{\lambda^2}{4\pi}$$

… the relative gain G of an antenna with some arbitrary effective aperture area A_e is given by:

$$G = \frac{4\pi}{1.64\,\lambda^2}A_e$$

Assuming no incidental losses, if a power P_t were radiated by an isotropic source (uniformly throughout a solid angle of 4π steradians), at a distance r away the power density would be:

$$S = \frac{P_t}{4\pi r^2}$$

If the same amount of power were radiated by a $\lambda/2$ dipole, the maximum power density at a distance r away would be increased by the factor 1.64:

$$S' = \frac{1.64P_t}{4\pi r^2}$$

If, in practice, the transmitting antenna has a gain G_t relative to a $\lambda/2$ dipole, the solid angle over which the energy is radiated is further reduced, proportionately, and the maximum power density is correspondingly increased:

$$S_r = \frac{1.64 P_t G_t}{4\pi r^2}$$

If the receiving antenna has an effective aperture area A_e, the power P_r it collects is:

$$P_r = S_r A_e = \frac{1.64 P_t G_t A_e}{4\pi r^2}$$

… and if this receiving antenna has a gain G_r relative to a $\lambda/2$ dipole, then its maximum effective aperture area is:

$$A_e = \frac{1.64\lambda^2}{4\pi} G_r$$

$$\therefore \quad P_r = P_t G_t G_r \left(\frac{1.64\lambda}{4\pi r} \right)^2$$

This is an adaptation of the Friis transmission formula [28]. The dimensionless factor $(1.64\lambda/4\pi r)^2$ represents the free-space 'spherical-spreading' transmission loss between two $\lambda/2$ dipole antennas which is multiplied by the cascaded gains (and any incidental losses) to yield the power received for a given power fed to the transmitting antenna.

It's probably incorrect to say that this transmission loss depends on λ because this factor has only appeared to relate the aperture area to the gain of the receiving antenna. The ability of a transmitting antenna to create a certain power density at a chosen radius clearly depends on its gain which, equally, is related to its aperture area by λ^2, so if the above expression were written containing the aperture areas of both antennas A_{et} and A_{er} it would become:

$$P_r = \frac{P_t A_{et} A_{er}}{r^2 \lambda^2}$$

… *reversing* the apparent sense of the dependence on λ . . . QED?

Returning to the version containing both antenna gains, in decibels:

$$10\log_{10} P_r = 10\log_{10} P_t + 10\log_{10} G_t + 10\log_{10} G_r + 20\log_{10}\lambda - 20\log_{10} r + 20\log_{10}(1.64/4\pi)$$

For example:

$$P_r\,(\text{dBm}) = P_t\,(\text{dBW}) + G_t\,(\text{dBd}) + G_r\,(\text{dBd}) + 20\log_{10}\lambda - 20\log_{10} r + 12.3$$

… with λ and r expressed in the same unit. Any incidental losses can be taken into account by subtracting their decibel values from the right-hand side.

The product of P_t and G_t is often represented by a single factor known as the ERP (Effective Radiated Power) with the unit of watts or dBW, which describes the strength of the signal radiated by a terrestrial transmitter.

However, for reasons that are probably clouded by history, the sensitivity of domestic television receivers, VHF broadcast radio receivers, communication receivers, VHF/UHF transceivers, etc., is often specified in terms of the input signal *voltage* required to yield a certain signal-to-noise ratio. $P_r = V_s^2/R$ where R is the input resistance of the receiver (and the characteristic impedance of the matched antenna and down-lead) so $V_s = \sqrt{(P_r R)}$. Evidently, this can introduce inaccuracy because the input signal *power* is the important parameter and specifying the sensitivity this way makes it dependent on R.

Nevertheless, the type of link budget often encountered in broadcasting can be presented in two stages. First, from a given transmitter ERP $P_{ERP} = P_t G_t$, the electric field strength is derived at the potential location of the receiving antenna:

$$S_r = \frac{1.64 P_{ERP}}{4\pi r^2} \quad \text{and} \quad S_r = \frac{E^2}{Z}$$

… so:

$$E^2 = \frac{1.64 P_{ERP} Z}{4\pi r^2} \quad \text{or} \quad E = \frac{7\sqrt{P_{ERP}}}{r}$$

This remarkably-simple expression for E is used as a 'rule of thumb' in Spectrum Planning Group.

In decibels: $20\log_{10}E = 10\log_{10}P_{ERP} - 20\log_{10}r + 10\log_{10}(1.64\times120\pi/4\pi)$

For example: $E\,(\mathrm{dB\mu V/m}) = P_{ERP}\,(\mathrm{dBW}) - 20\log_{10}r + 137$

… with r expressed in metres. Of course, this simple version is only valid for a single line-of-sight path and the practical case for paths diffracted over hills and other obstacles would be much more complicated.

The second step is to apply the coupling factor for a receiving antenna based on a dipole to yield the signal voltage appearing across the input terminals of the receiver. Recapping from the beginning of the previous section, the EMF induced in a $\lambda/2$ dipole is given by:

$$V = E\frac{\lambda}{\pi}$$

… and the signal voltage V_s presented to a perfectly matched receiver is half V so:

$$V_s = E\frac{\lambda}{\pi}\cdot\frac{1}{2} \quad \text{or} \quad V_s^{\,2} = E^2\left(\frac{\lambda}{\pi}\cdot\frac{1}{2}\right)^2$$

Including the power gain of the receiving antenna G_r, relative to a $\lambda/2$ dipole, and the power loss L of the down-lead connecting the antenna to the receiver, we have:

$$V_s^{\,2} = \frac{E^2G_r}{L}\left(\frac{\lambda}{\pi}\cdot\frac{1}{2}\right)^2$$

In decibels: $20\log_{10}V_s = 20\log_{10}E + 10\log_{10}G_r - 10\log_{10}L + 20\log_{10}\lambda - 20\log_{10}2\pi$

For example: $V_s\,(\mathrm{dB\mu V}) = E\,(\mathrm{dB\mu V/m}) + G_r\,(\mathrm{dBd}) - L\,(\mathrm{dB}) + 20\log_{10}\lambda - 16$

… with λ expressed in metres.

Putting the two steps together, for example:

$$V_s\,(\mathrm{dB\mu V}) = P_{ERP}\,(\mathrm{dBW}) + G_r\,(\mathrm{dBd}) - L\,(\mathrm{dB}) - 20\log_{10}r + 20\log_{10}\lambda + 121$$

… with r and λ expressed in the same unit.

I've given an example using some of these expressions in the box below: a practical example – the gainless repeater.

a practical example – the 'gainless' repeater

The on-channel repeater principle is useful for extending transmitter coverage without need for additional spectrum but unwanted feedback always imposes an upper limit on the amount of gain that can be used. In some circumstances, though, a small but worthwhile extension of coverage might be achieved with no gain at all – just a receiving antenna, located where there is plenty of field strength, connected by a cable to a transmitting antenna in a region that would otherwise lack sufficient field strength; inside a shop, for example.

The main limitation of this approach is on account of spherical spreading from the new, secondary transmitting antenna. Whereas the original, primary transmitting antenna is some distance away and moving a portable receiver one metre further away from it has little effect, moving the receiver one metre further away from the secondary antenna when it's only one metre away to start with can reduce the available field strength by 6 dB.

Suppose the building-penetration loss of a shop is 20 dB, from outside on the roof to the display area, the field strength outside E_1 = 55 dBμV/m and the value required for solid operation of receivers is E_2 = 40 dBμV/m. From expressions presented in Section 7 it can be deduced that:

$$\frac{E_2^{\,2}}{E_1^{\,2}} = \frac{\lambda^2}{r^2}\cdot\frac{G_rG_t}{LR}\cdot\left(\frac{7}{2\pi}\right)^2$$

… where G_r and G_t are the gains of the receiving and transmitting antennas at the ends of the cable for which L is the loss, R is the characteristic impedance of the matched system and r is the distance between the secondary transmitting antenna and the receiver. If λ = 1.33 m (DAB), G_r = 4 (Yagi), G_t = 1 (dipole), L = 2 (50 m of H100 co-axial cable plus connectors) and R = 50 Ω, this gives $E_2^{\,2}/E_1^{\,2} = 0.09/r^2$. The requirement in this case is $E_2^{\,2}/E_1^{\,2} \geq 0.03$ (-15 dB) so $r \leq 3$ m and the secondary antenna could be mounted above a suspended ceiling, for example (but not directly above whip antennas).

Of course, this 'repeater' is bi-directional so the principle could be used to improve mobile telephony in an otherwise-screened enclosure like a modern railway carriage.

8. Coupling factor of an isotropic source

In the derivation of the effective aperture of an isotropic source, we found:

$$P = \frac{{E_A}^2 A}{Z} \quad \text{and} \quad A = \frac{\lambda^2}{4\pi} \quad \text{so} \quad P = \frac{{E_A}^2}{120\pi} \cdot \frac{\lambda^2}{4\pi}$$

Now we can drop the subscript from E_A because we are dealing solely with the incident (i.e. aperture) field. Also, in the derivation for a $\lambda/2$ dipole we covered the general result for a perfectly matched receiving antenna terminated in a resistance equal to its own radiation resistance R_r, that:

$$P = \frac{V^2}{4R_r} \quad \text{so} \quad \frac{V^2}{4R_r} = \frac{E^2}{120\pi} \cdot \frac{\lambda^2}{4\pi} \quad \text{or} \quad V^2 = \frac{E^2\lambda^2}{\pi^2} \cdot \frac{R_r}{120}$$

$$\therefore \quad V = E\frac{\lambda}{\pi}\sqrt{\frac{R_r}{120}}$$

Unlike the $\lambda/2$ dipole, the radiation resistance R_r of an isotropic source cannot be evaluated uniquely. Indeed, if an isotropic antenna could be realised its radiation resistance would probably be so small as to make its radiation efficiency close to zero. However, the isotropic source is only intended to be a hypothetical reference so if it were postulated that an isotropic receiving antenna could be made with R_r = 73 Ω, the same as a $\lambda/2$ dipole, then:

$$V = E\frac{\lambda}{\pi}\sqrt{\frac{73}{120}} = E\frac{\lambda}{\pi} \cdot \frac{1}{\sqrt{1.64}}$$

In comparison with the result for a $\lambda/2$ dipole, the voltage ratio $1/\sqrt{1.64}$ yields the expected -2.16 dB fractional relative 'gain' of an isotropic source. Generally, considering an isotropic source in the same way as one would a dipole, the effective length or 'effective height' is then:

$$h = \frac{\lambda}{\pi}\sqrt{\frac{R_r}{120}} \text{ (m)}$$

It may not be widely appreciated that most of the apparent gain of a $\lambda/2$ dipole is *not* on account of its physical length but simply a consequence of the shape of its radiation pattern, which is mainly on account of its cylindrical symmetry. An infinitesimal electric doublet, that is a dipole of negligible length, has a radiation pattern similar to that of a $\lambda/2$ dipole but a slightly larger half-power beamwidth (90° as opposed to 78°) and consequently a gain of 1.76 dBi [29]; only 0.4 dB smaller. The greater length of the $\lambda/2$ dipole increases its radiation resistance and, therefore, its radiation efficiency. I will say more about the choice of this specific $\lambda/2$ length a little later, in Section 11.

Wherever I refer to beamwidth hereafter, I will always mean the angle between the half-power points.

9. Link budgets involving antennas related to an isotropic source

The most common application of an isotropic source as a reference is in link budgets involving aperture antennas that are not dipoles or arrays of dipoles, such as parabolic dishes. It is conventional to express the gain of such antennas in dBi, meaning decibels above the coupling factor of an isotropic source. Again, this gain can be interpreted as the degree to which the antenna beam is confined or as the relative size of the effective aperture area.

Since the effective aperture area of an isotropic source is:

$$A = \frac{\lambda^2}{4\pi}$$

… the relative gain [30] of an antenna with some arbitrary effective aperture area A_e is given by:

$$G = \frac{4\pi}{\lambda^2}A_e$$

Assuming no incidental losses, if a power P_t were radiated by an isotropic source (uniformly throughout a solid angle of 4π steradians), at a distance r away the power density would be:

$$S = \frac{P_t}{4\pi r^2}$$

If, in practice, the transmitting antenna has a gain G_t relative to an isotropic source, the solid angle over which the energy is radiated is reduced proportionately from 4π and the power-density is correspondingly increased:

$$S_r = \frac{P_t G_t}{4\pi r^2}$$

If the (separate) receiving antenna has an effective aperture area A_e, the power P_r it collects is:

$$P_r = \frac{P_t G_t A_e}{4\pi r^2}$$

If this receiving antenna has a gain G_r, then its effective aperture area is:

$$A_e = \frac{\lambda^2}{4\pi} G_r$$

$$\therefore \quad P_r = P_t G_t G_r \left(\frac{\lambda}{4\pi r} \right)^2$$

… and this is the adaptation of the Friis transmission formula most often used in satellite link budgets. The factor $(\lambda / 4\pi r)^2$ is sometimes referred to as the free-space 'isotropic loss' – the transmission loss on account of spherical spreading between two isotropic antennas – which is multiplied by the cascaded gains, and any incidental losses, to yield the power received for a given power fed to the transmitting antenna. This is convenient because it makes the transmission loss dimensionless, like most other losses and the antenna gains. As before, the dependence on λ has more to do with what happens at one end than the propagation path itself.

In decibels: $10\log_{10} P_r = 10\log_{10} P_t + 10\log_{10} G_t + 10\log_{10} G_r + 20\log_{10} \lambda - 20\log_{10} r - 20\log_{10} 4\pi$

For example: P_r (dBm) = P_t (dBW) + G_t (dBi) + G_r (dBi) + $20\log_{10} \lambda - 20\log_{10} r + 8$

… with λ and r expressed in the same unit. Any incidental losses can be taken into account by subtracting their decibel values from the right-hand side.

The product of P_t and G_t is often represented by a single factor known as the EIRP (Effective Isotropic Radiated Power) with the unit watts or dBW, which describes the potential strength of a satellite's down-link signal, for example. Sometimes, G_r is omitted from the calculation and then P_r can be termed the 'incident isotropic power', that is the power that would be delivered by an isotropic receiving antenna – to be increased proportionately by whatever receiving antenna gain is available, relative to an isotropic antenna.

I've given an example in the box below: a practical example – waves and photons. I had wondered if the DSat signal received by a domestic dish antenna was the weakest RF signal used by any domestic appliance but nowadays that credit probably belongs to the signals received by GPS receivers.

a practical example – waves and photons

Consider one of the Astra 2D satellites we use presently to broadcast DSat. Its EIRP is about 51 dBW at beam-centre (near Liverpool) and Transponder 45, used for BBC One, etc., has a down-link frequency of about 10.8 GHz, so λ = 27.8 mm. The height of the geostationary orbit is about 35.7×10^6 metres and the gain of a typical '75 cm' offset receiving dish antenna is about 36 dBi. Putting these values into the relevant equation yields P_r = -87.2 dBW or 1.9 nW fed into the low-noise front-end of the satellite receiver.

1.9×10^{-9} J.s^{-1} is probably the weakest signal used in the domestic environment but, in terms of photons, it is still a very substantial signal. The energy in a single photon is given by $E = h\nu$ where Planck's constant $h = 6.26\times10^{-34}$ J.s and the frequency $\nu = 10.8\times10^9$ Hz, so $E = 6.8\times10^{-24}$ J. Therefore, about 280 trillion photons are received each second, or nearly 26,000 per cycle – assuming one's allowed to consider photons and cycles, or waves, at the same time!

An alternative way to represent a satellite link budget is to leave in the effective aperture A_e of the receiving antenna as we had earlier:

$$P_r = P_t G_t A_e \frac{1}{4\pi r^2}$$

The factor $1/4\pi r^2$ is sometimes referred to as the free-space 'spreading loss' – the transmission loss on account of spherical spreading between an isotropic transmitting antenna and a receiving aperture of unit area – multiplied by the actual area and whatever other cascaded gains and incidental losses to yield the power received for a given power fed to the transmitting antenna. In this case the 'spreading loss' has the dimension m^{-2}, which counteracts the dimension of A_e.

In decibels: $10\log_{10} P_r = 10\log_{10} P_t + 10\log_{10} G_t + 10\log_{10} A_e - 20\log_{10} r - 10\log_{10} 4\pi$

For example: P_r (dBm) = P_t (dBW) + G_t (dBi) + A_e (dBm^2) – $20\log_{10} r$ + *19*

… with r expressed in metres and any incidental (dB) losses subtracted from the right-hand side. Here, A_e is expressed in decibels relative to one square metre and must take account of the 'aperture efficiency' of the antenna; the fraction of the physical aperture area that accounts for received power. This can have a value anywhere from about 50%, for a symmetrical paraboloid dish where the feed or sub-reflector blocks part of the aperture, to 85% or more for an offset paraboloid.

There can't be many cases in spectrum planning where it's necessary to relate field strength to EIRP – helicopter measurements at HF is one – in which case for a given EIRP $P_{EIRP} = P_t G_t$:

$$S_r = \frac{P_{EIRP}}{4\pi r^2} \quad \text{and} \quad S_r = \frac{E^2}{Z}$$

… so:

$$E^2 = \frac{P_{EIRP} Z}{4\pi r^2} \quad \text{or} \quad E = \frac{\sqrt{30 P_{EIRP}}}{r}$$

Another remarkably simple 'rule of thumb' used in Spectrum Planning Group and in ITU-R texts, but note that the number 30 is *inside* the root in this case.

In decibels: $20\log_{10} E = 10\log_{10} P_{EIRP} - 20\log_{10} r + 10\log_{10} 30$

For example: E (dBμV / m) = P_{EIRP} (dBW) – $20\log_{10} r$ + *134.8*

... with r expressed in metres.

PART 2: electrical engineering of dipole antennas

10. Antenna or transmission line?

Back in the 1980s a television series was broadcast in the UK about the Royal Navy submarine fleet. One of the officers interviewed came forth with the eternal line "There are two kinds of vessel: submarines and targets". Whilst obviously incorrect, this sort of statement nevertheless displays 'painting the world black and white' to a remarkable extent. In the same vein it could be stated that:

there are two kinds of electrical conductor: transmission lines and antennas

The simplest forms of transmission line have two terminals at each end and a uniform characteristic impedance throughout the length of the line, meaning the ratio of distributed inductance to capacitance (per unit length) is held constant. This impedance determines the ratio of voltage to current at any point in the line, and when terminated by a load of the same impedance, the line will pass signals to the load without anything being reflected. A two-conductor transmission line is just a pair of conductors closely coupled to one another in an electromagnetic sense, capacitively and inductively, and because of this strong coupling very little power 'leaks' out.

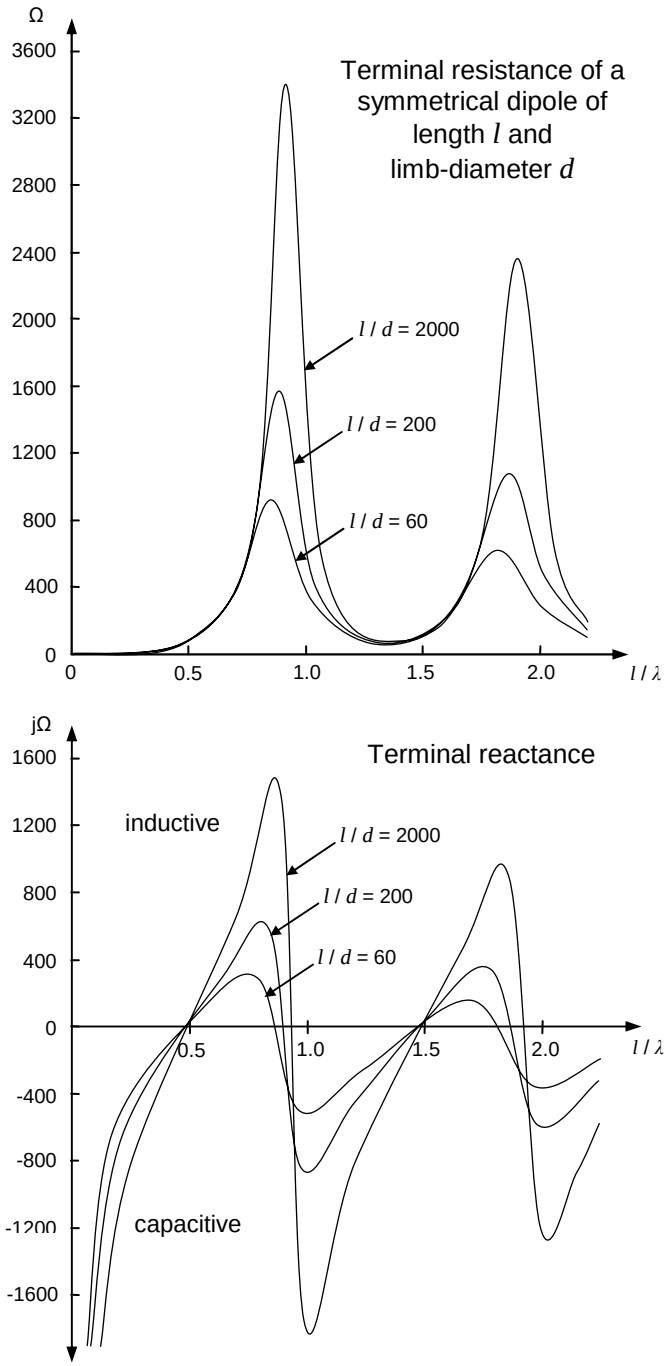
When a transmission line is carrying power, in a travelling wave that propagates between the conductors, the current in one conductor is matched by an equal current in the other, travelling in the opposite direction (*viz* two 'legs' of the same circuit). If the conductors are close to one another, some distance away the magnetic field generated by the current in one conductor is cancelled by the equal and opposite contribution from the other. Also, the electric field generated between the conductors is more-or-less confined to the space between them. Consequently, transmission lines don't radiate much. A common form of transmission line encountered in the UK for intentional use at radio frequencies is co-axial cable [31]. In this, the electric field really is confined to the dielectric between the inner and outer conductors and, being surrounded by a continuous conducting 'screen', has a value close to zero everywhere outside the cable. Another form, apparently more popular for domestic radio and television installations in the USA but becoming widespread throughout the world through its use in computer networks, is the symmetrical 'two-wire' transmission line (AKA: balanced twin feeder, figure-8 and ribbon cables, mains and telephone wiring, Cat-5 cable, etc. – essentially all non-coaxial cables). This relies on symmetry and close spacing of the conductors to constrain radiation.

Many of the authors of the standard texts on antennas, at some point in their deliberation, consider an antenna as an 'opened-out' transmission line that 'leaks' energy from the wave being propagated along the line ([32] for example). Indeed, any conductor carrying a changing current that is not *very closely coupled* to another conductor carrying an equal and opposite current is likely to give rise to a radiated wave and *can be considered an antenna.* In this context, 'close' means a separation a small fraction of λ in the medium that separates the conductors, and more-complicated balanced arrangements like three-phase power lines are not ruled out (although λ = 6,000 km in air for 50 Hz!).

11. Why 'λ/2 dipoles' and not some other length?

The terminal impedance of a dipole of arbitrary length and limb diameter can be calculated analytically, but to do this accurately the expressions involved are so long they would fill a whole chapter by themselves and wouldn't make a riveting read! Small wonder, then, that many of the well-known text books on antennas in general that go into this level of detail reproduce the results of the heroic theoretical work of Hallén [33], documented in 1938, or those of the extensive measurements by Brown and Woodward [34] from 1945. I don't have direct access to Hallén's paper but reference is made to it in [35], wherein graphs of his results are reproduced. Kraus [36] also devotes several pages to Hallén's work. Of course, with the computing power available nowadays it should be fairly straightforward to calculate this numerically using a program such as NEC [19] with some kind of batch processing.

Maintaining the time-honoured tradition (!), I have illustrated some of Hallén's results in graphs on the next page, showing the resistive and reactive components separately. In order to separate this subset of curves from [35] and to present them on horizontal axes with simple scaling, I have traced them using Bézier curves. These copies are not accurate but show the important trends. Also these results all apply to dipoles that are symmetrical about their terminals – centre-fed by a balanced source, the two 'poles' having equal length.



I have dealt with the terminal impedance of open- and short-circuited transmission-line 'stubs' in the box transmission lines as resonant stubs on the next page, and by comparison with Hallén's reactance curves it's clear that a dipole does, indeed, behave something like an open-circuit stub of similar length to each dipole limb. However, I must emphasise 'something like'; unlike a simple transmission line, the antenna radiates power, probably in a non-uniform manner along the length of its limbs (as suggested by the current distribution). Also the characteristic impedance of the equivalent transmission line, as determined by the ratio of distributed L/C, is probably not uniform along the length of the dipole limbs.

The periodic nature of the resonance is evident in the curves to the left. The terminal reactance is small when the length is near a multiple of $\lambda/2$, but the resistance is large when the length is near a multiple of λ. The maximum resistance and reactance decrease as the ratio l/d decreases – as the dipole is made more "fat" – and the rates of change of resistance and reactance with λ are also reduced. Indeed, considering the dipole either as a transmission-line stub or a lumped-element LCR resonant circuit, it follows that "fatter" dipoles offer greater bandwidth.

For the case of a stub, the slope of its terminal impedance depends on the characteristic impedance of the transmission line, $Z_0 = \sqrt{(L/C)}$, so smaller slope is achieved with greater C and/or smaller L, however non-uniform they may be.

The '-3 dB' bandwidth of a resonant circuit is inversely proportional to its Q-factor which is proportional to the ratio of energy stored by the reactance to energy lost by the resistance (dissipated or radiated) [37]. Q-factor is a real dimensionless number so the expression $Q = (1/R)\sqrt{(L/C)}$ for a general LCR circuit is readily believable, if not obvious (the reactance of an inductor is proportional to L whereas that of a capacitor is inversely proportional to C). So wider bandwidth, lower Q, is achieved with smaller reactance which, again, means greater C and/or smaller L, however non-uniform they may be.

Far apart as they are, it is more-or-less intuitive that the capacitance between the dipole limbs is increased when their diameter, and thus surface area, is increased. $C = \varepsilon A/s$ where ε is the permittivity of air, ($\varepsilon \approx 1$), A is the area of the 'plates' and s their separation – think small and large 'Leyden jars'!

Also, it can be shown [38] that the self-inductance of a single cylindrical conductor (the remainder of the circuit not contributing by being far away or comprising perpendicular conductors) is dependent on the reciprocal of d so the larger its diameter the lower its self-inductance. This fact seems to be well known in some circles [e.g. 39] and I can recall statements about earthing straps at VHF needing to be made of wide copper tape or braid, but maybe this was to reduce AC resistance on account of

the skin effect more than the inductance. However, I will suggest later that belief in this 'fact' can lead to a paradox, so I'm not completely surprised when some colleagues are incredulous about it and, intriguingly, Kraus [40] steers clear of it. For the record, the formula provided by Terman [39], suitably metricated, is $L = 0.129\,l\,[\log_e(l/d)-1]$ nH where l is the length of the conductor and d its diameter, both in metres. This formula correlates with the work of Rosa [38] and may follow from it.

For the example of a $\lambda/2$ dipole for Band II, the length would be about 1.5 m so $l/d = 60$ would correspond to a limb-diameter of 25 mm – a "fat" dipole indeed – whereas $l/d = 2000$ would correspond to a thin wire of diameter 0.75 mm: between 21 and 22 SWG in old (British) money.

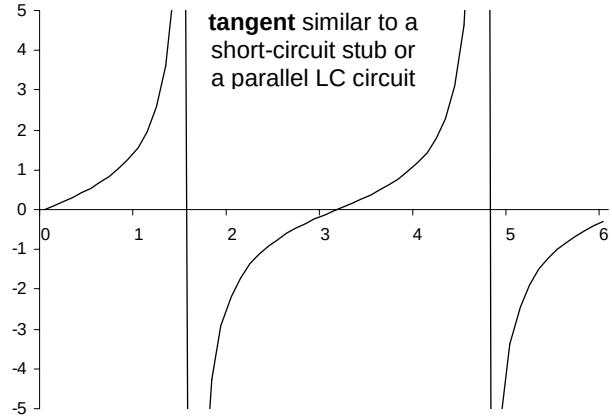
The usual objective is to design an antenna that can easily be matched to 50 Ω or 75 Ω (the reasons for the choices of these particular values are explained in my companion White Paper [31]) over a sufficiently wide bandwidth to avoid the need for adjustments or a multitude of frequency-specific variants. Penalties for not achieving a good match can include extraneous loss, uneven frequency response and incorrect operation of transmitters on account of poor VSWR – perhaps even damage to power amplifiers or transmission lines in extreme cases. The simplest and most universal solution is to make the antenna, and the transmitter or receiver to which it is connected, both present terminal impedances that are purely resistive and have the same value as the uniform characteristic impedance of the transmission line that connects them. Then maximum power transfer and 'flat' frequency response can be obtained whatever the length of the line.

However, what can be achieved in practice is often a compromise.

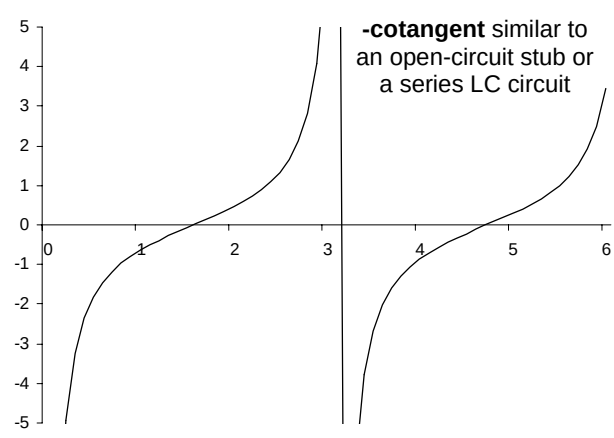
transmission lines as resonant stubs

The 'transmission line relation' [41] gives the impedance Z_x at a point on a uniform transmission line of characteristic impedance Z_0 some distance x from the load of impedance Z_L. For a lossless line:

$$Z_x = Z_0 \frac{Z_L + jZ_0 \tan \beta x}{Z_0 + jZ_L \tan \beta x} \quad \text{... where } \beta = 2\pi/\lambda.$$



If the load is made a short circuit, $Z_L = 0$, then $Z_x = j\,Z_0 \tan 2\pi x/\lambda$ which is infinite when $x = \lambda/4$, $3\lambda/4$, $5\lambda/4$, etc. If the length of the line is held constant the impedance it presents at the non-short-circuit end is purely reactive (it's preceded by j) and follows the tangent of the frequency. To the left I've plotted the tangent of an angle expressed in radians using MS Excel. Over the first cycle, this exhibits similar behaviour to the impedance of a parallel LC circuit. The reactance of the inductor (positive sign) dominates at low frequencies – it presents a low impedance to DC – the reactance of the capacitor (negative sign) dominates at high frequencies and an abrupt 'flip' from +∞ to -∞ occurs at the resonant frequency.



If the load is made an open circuit, $Z_L = \infty$, then $Z_x = -j\,Z_0 \cot 2\pi x/\lambda$ which is zero when $x = \lambda/4$, $3\lambda/4$, $5\lambda/4$, etc. In this case, if the length of the line is held constant the impedance it presents at the other end is purely reactive again (it's preceded by j) and follows minus the cotangent of the frequency. To the left I've plotted minus the cotan of an angle expressed in radians. Over the first cycle, this exhibits similar behaviour to the impedance of a series LC circuit. The reactance of the capacitor (negative sign) dominates at low frequencies – it presents a high impedance to DC – the reactance of the inductor (positive sign) dominates at high frequencies and the impedance passes through zero at the resonant frequency.

A short length of transmission line terminated in a short- or open-circuit is often referred to as a 'stub' and the slope of its terminal impedance depends on $Z_0 = \sqrt{(L/C)}$, the characteristic impedance of the transmission

Although laboratory RF instruments like signal generators and spectrum analysers are usually designed to present purely resistive terminal impedances close to 50 Ω or 75 Ω, a transmitter may be designed with a very different effective output resistance in order to provide high output power efficiently, and a receiver may be designed with a smaller effective input resistance in order to reduce the fraction of the total noise power received on account of that input resistance.

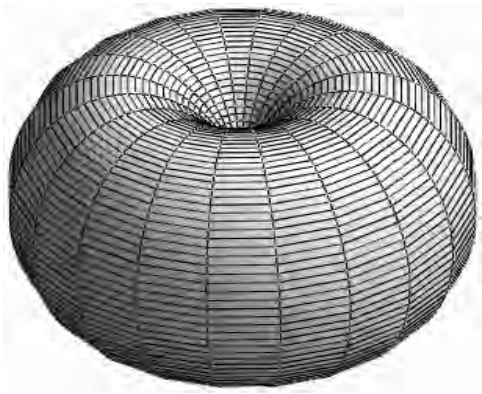
The terminal resistance of a dipole antenna can be modified a little by changing its length, or matched using some kind of transformer (I'll cover λ/4 and λ/2 transmission lines as transformers later in Section 14), and a terminal reactance can be mitigated over some frequency range by 'resonating it out' with a complementary reactance – for example, using a series inductor 'loading coil' to resonate the negative reactance, that is effective series capacitance, of a short dipole ($l/\lambda < 0.4$, or a monopole half this length). I'll mention loading again in Section 17. The main obstacle to obtaining a match over a wide bandwidth is the difficulty in realising a component or circuit that has complementary variation of reactance with frequency. In this respect, a number of observations can be made from Hallén's curves interpreting the variation with $l/\lambda = lf/c$ as the variation with frequency for fixed l:

1. the large resistances and reactances around integral multiples of λ are usually avoided because of their large rates of change with frequency (although some clever designs exist for full-wavelength dipoles involving resistance transformation and reactance compensation)
2. the "fatter" the limbs the smaller the slope of the resistance and the reactance with respect to frequency, although considerations of cost, weight and wind resistance may affect this
3. the regions around the resonances near $l/\lambda = 0.5$ and 1.5 (etc.), where the reactance passes through zero, are usually preferred because the resulting 'Q' of the matched antenna is minimised. The Q of any resonant circuit is proportional to the size of the reactances involved, as noted earlier
4. in these regions the slope of the reactance has the same sign as that of an open-circuit $\lambda/4$ stub or a series LCR circuit (i.e. the capacitor's negative-sign reactance dominates at lower frequencies and the inductor's positive-sign reactance dominates at higher frequencies)

The choice between $l/\lambda = 0.5$ or 1.5 (or greater) is explained by looking at the radiation patterns.

12. Radiation patterns of dipoles

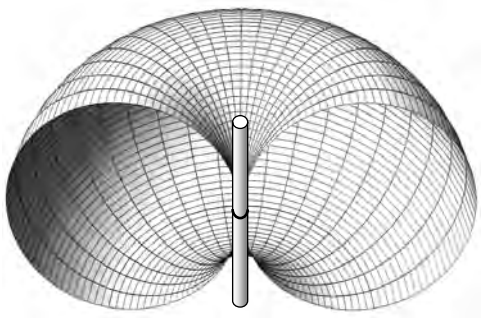
Since the distant E and H fields of a dipole are related by Z, the uniform intrinsic impedance of space, it is only necessary to calculate the variation of one of them in order to obtain the radiation pattern, which can be expressed in terms of relative field strength or relative power (density). We often present radiation patterns as the variation of E-field strength but the H-field is easier to calculate. The H-field strength at some point of inspection is related to the current in the dipole by the Biot-Savart law so, using an assumed or known current distribution, it can be found by integrating over the length of the antenna [42].



The 3-dimensional radiation pattern of a $\lambda/2$ dipole is the well-known 'doughnut' shape, or more correctly a torus surrounding the antenna with no space in the middle, as shown to the right. Symmetry dictates that the pattern is constant or 'omni-directional' in any plane perpendicular to the axis of the dipole. In any plane containing the axis, the result of the integration can be expressed in terms of the variation of electric-field strength E with varying angle θ from the axis of the dipole:

$$E = \frac{\cos[(\pi/2)\cos\theta]}{\sin\theta}$$

… at a constant radius from the centre of the antenna. The locus of this has a single annular lobe with a beamwidth of 78° as illustrated by the cut-away view of the torus to the right (the rod in the middle represents the dipole).



I used Mathcad to plot all the 3-dimensional radiation patterns in this document.

When greater gain is required, to improve the link budget or to tailor the coverage area, an array of multiple dipoles can be used with an appropriate distribution of currents (amplitude and phase) to the individual dipole elements. A well-known example is the Yagi-Uda array used for domestic television reception, which is an 'end-fire' array of dipoles coupled to a single 'driven element' (often some form of dipole). The coupling in this case is parasitic, that is, achieved by radio!

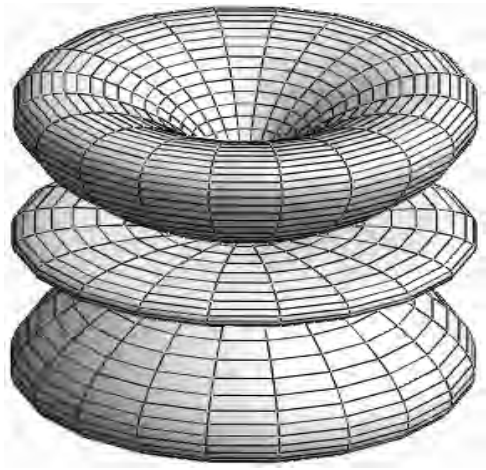
'Broadside' arrays of dipoles are used at many of the VHF/UHF radio and television transmitting stations, particularly the high-power ones, and one of the two possible forms of one-dimensional broadside array for vertical polarisation is the 'collinear' array. A balanced full-wavelength dipole is effectively a collinear array of two closely-spaced $\lambda/2$ dipoles and has a similar pattern to a single $\lambda/2$ dipole but with a smaller beamwidth of 47° and, consequently, 3 dBd gain.

This is the longest a balanced single dipole can be made whilst retaining a pattern with a single lobe. With a length greater than λ, some part of a sinusoidal current distribution must have the opposite sign [8] so, in the integration, contributions from some parts must add destructively in some directions. The result is a more complicated pattern having multiple annular lobes with nulls between them; the phase of the radiated field alternates between adjacent lobes.

Expressed as before, the pattern of a $3\lambda/2$ dipole is:

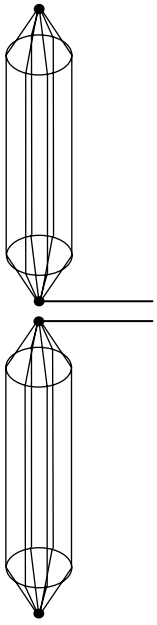
$$E = \frac{\cos\left[(3\pi/2)\cos\theta\right]}{\sin\theta}$$

Plotted 3-dimensionally this appears as shown to the right; again, the dipole sits on the axis of symmetry. This is less useful for broadcasting or communications because the nulls correspond to directions in which there would be no coverage. Longer dipoles have even greater numbers of lobes and nulls.



Consequently, the $\lambda/2$ dipole (and occasionally the λ dipole) is the preferred choice for many applications, and combining several in an array allows a great degree of control over the resulting radiation pattern.

13. The cage dipole

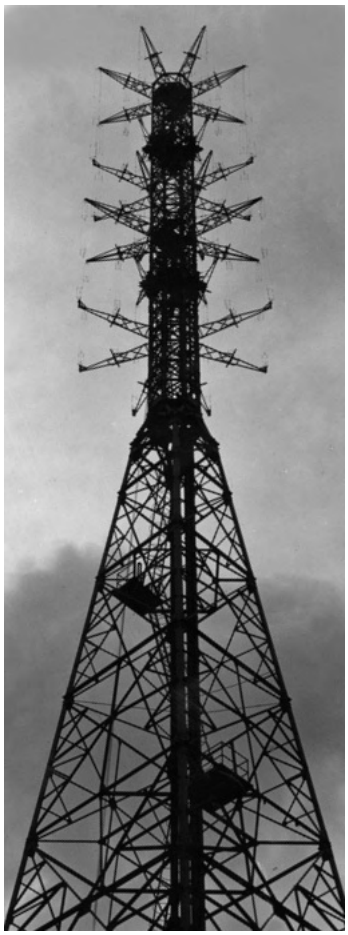


A means of achieving similar electrical properties to "fat" dipole limbs whilst constraining weight and wind resistance is to use a 'skeletal' representation of wide strips or cylinders of large diameter, made up from a number of thinner conductors separated across the width of the equivalent strip, or around the circumference of the equivalent cylinder as illustrated to the left. The conductors in each limb are connected together at least at the terminals – if they are connected at both ends the result is known as a 'cage' dipole.

Schelkunoff [43] provides some insight into 'principal waves on cage structures' and gives a general expression for the effective radius of a cage, $a_{eff} = a(na_0/a)^{1/n}$, where a_0 is the radius of each of the n conductors and a is the radius of the cage. This can result in values of a_{eff} smaller or greater than a depending on the parameter values chosen. The separation of adjacent conductors must be kept positive and much smaller than the wavelength but this is inevitable for a $\lambda/2$ cage dipole longer than it's wide – if the separation in a cylindrical cage is made about the same as a then $n = 2\pi \approx 6$.

For example, to achieve the effect of $l/d = 60$ in a $\lambda/2$ dipole for $\lambda = 3.0$ m, a design frequency of 100 MHz, would require **solid limbs** of $d = \lambda/120 =$ **25.0 mm** diameter (as before). For the equivalent cage dipole, $a_{eff} = \lambda/240$, and with $n =$ **6 conductors** having a separation about the same as a, if the conductors have $a_0 = \lambda/1200$ for example (i.e. 5 mm diameter wires), then $\lambda/240 = a(6\lambda/1200a)^{1/6}$ giving $a = 0.96\,\lambda/240$ or a cage diameter of **24.1 mm**. Alternatively, if $n =$ **4 conductors** then $a = 1.08\,\lambda/240$ giving a cage diameter (or diagonal) of **27.0 mm**. For the simplest case of two conductors, $n = 2$ and $a_{eff} = \sqrt{(a_0 s)}$ where s is the separation between the centres of the conductors so to achieve similar electrical properties using $n =$ **2 conductors** would require a separation $s =$ **62.5 mm**.

The original Band I antenna system at Alexandra Palace, for each of the separate vision and sound transmissions, consisted of eight pairs of vertical cage dipoles around the lattice tower [44] shown to the right. The limbs of these were each composed of three wires and are difficult to make out clearly in available photos (there's a clearer photo' in [45]), but the radial spars between which they were strung are obvious in historical photos like this one. Each pair of dipoles was configured as a directional end-fire array with both dipoles driven, probably to avoid interaction with the central metal tower, and all eight (most likely fed in-phase) collectively provided an omni-directional radiation pattern. Two very large cage dipoles!

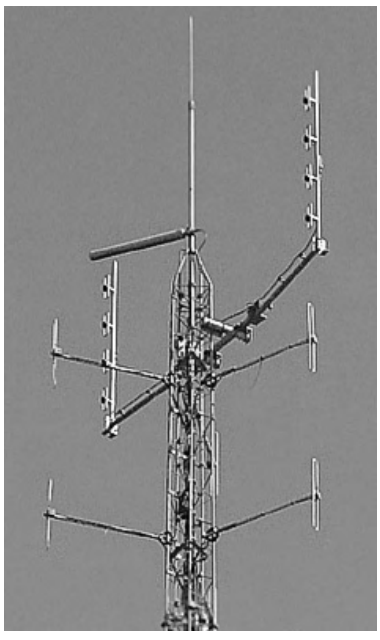


Horizontal HF cage dipoles are sometimes seen above the roofs of embassies, and large 'curtain arrays' of them are used extensively at the HF transmitting stations run by BBC World Service, and Vosper Thornycroft Communications working for them. Wire antennas used for transmitting LF and MF are often made with several parallel conductors, for example the 'T' antenna at Droitwich for BBC Radio 4 long-wave which has four parallel wires arranged as a skeletal strip. This increases the capacitance of the horizontal 'top cap' making the current distribution in the vertical, radiating, part more uniform. It may also reduce losses by distributing the earth current.

Cage dipoles have also been used for some 50 years at radio stations around the UK concerned with communication with civil aircraft. This antenna design [46] dates from before 1947 and was originally conceived to operate over the half-octave band 100 to 156 MHz used then for communication with military, as well as civil, aircraft. However, for many years the civil communication band has been limited to 118 to 137 MHz and some years ago the cage dipoles were replaced by much simpler folded dipoles, which I'll describe presently.

14. The λ/2 folded dipole

The simple λ/2 dipole (two 'prongs') gets a lot of coverage in text books because it's relatively simple to analyse – but how often have you seen one of these in use? Could it be that in most recent cases you've seen a folded dipole, as illustrated to the left? This can be described as a squashed loop, or a simple dipole with an additional continuous λ/2 conductor parallel to it and connected to its ends. There are several good reasons why the folded dipole is used much more often, as the 'driven element' (i.e. connected to the down-lead) in Yagi-Uda arrays for television reception, and for VHF/UHF communication (i.e. PMR) base stations like the one shown to the right (Banstead High Street) – there are 14 dipoles in this photograph; every one of them folded!



The folded dipole, being a variation of the simplest cage dipole, is one of the simplest 'skeletal' implementations of a "fat" antenna. Its terminal resistance can easily be deduced to a first order: see the box on the next page terminal resistance of a λ/2 folded dipole = $4 \times 73\ \Omega$ for my own effort. I must admit I find the, different, explanations of this provided by Kraus [47], Page [48] and others difficult

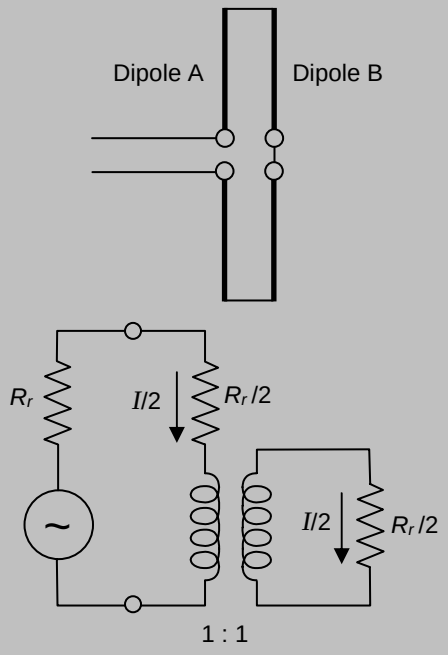
to understand and possibly incomplete, although they may be based on more-rigorous underlying treatments. For simplicity and the same reason as I gave back in Section 3, I have chosen to present the radiation resistance and the 'transformer winding' as lumped elements in series although the parallel representation used by others ([49] for example) is probably more correct for components that are distributed throughout the dipole limbs.

It may be interesting to consider what would happen if the 'short-circuits' at the ends of the limbs were not present. Then the impedance presented to the terminals for the difference-mode current would not necessarily be large at the first resonance (where $l/\lambda = 0.5$) so a more-significant part of the input current could flow that way. Some distance away, there would be little radiation from the difference-mode current because any integral would tend to cancel – the definition of an antenna I gave at the end of Section 10 would not be met. So this arrangement of conductors would make a poor transmitting or receiving antenna.

I've never tried sawing the ends off a folded dipole myself but I'm aware of someone who has! In his highly-informative web pages on antennas [50], David Jefferies, of the University of Surrey, shows network analyser traces for a folded dipole with and without the short-circuits in place – without them the terminal impedance becomes very small.

terminal resistance of a λ/2 folded dipole = 4 × 73 Ω

A folded dipole consists of a simple, non-folded λ/2 dipole ('Dipole A') the ends of which are connected to the ends of a λ/2 continuous conductor, which can be considered as a λ/2 dipole with its terminals connected together ('Dipole B'), arranged parallel to it and closely spaced. The small separation of these two dipoles, typically λ/10 or less, creates inductive coupling between them that forces their currents to be equal in magnitude and direction – like the action of a 1:1 transformer. Each has a radiation resistance and if they are made of material of the same thickness their radiation resistances must be equal. A current fed into the terminals of Dipole A will flow through its radiation resistance and also through the 'primary' of the effective transformer (distributed throughout the length of this dipole), causing an equal current to flow in the 'secondary', that is Dipole B, and through its own radiation resistance.



The radiation patterns of the combined pair of dipoles, parallel to one another, in very close proximity and carrying equal co-phased currents will be more-or-less identical to those of a simple non-folded dipole. The only way to achieve otherwise would be to reverse the direction of the current in one of the dipoles. This is almost intuitive because an integral of contributions to the H field strength from elements of current so close together is bound to show constructive addition at any angle.

Of course, these parallel conductors acting as a two-wire transmission line could carry a difference-mode current (as transmission lines do when conveying power) as well as the common-mode current. However, the 'short-circuits' between the conductors at their ends ensure that a high impedance is presented to this current mode in the vicinity of the terminals of Dipole A (see the box on the next page: λ/4 and λ/2 transmission lines as transformers) so relatively little current flows in this mode near the dipole's resonance.

Since the radiation patterns are the same, a folded dipole and a non-folded dipole (both λ/2 long and matched) must exhibit the same gain or aperture area. For an equal input power P, the resulting H field an equal distance away must have the same strength and, since this is related to the current in the dipole by the Biot-Savart law [13], the current I in the non-folded dipole must be the same as the total current in the folded dipole. These are all peak values with respect to time and location. Then the current in each of Dipole A and Dipole B must be $I/2$ and the current passing through the terminals of the folded dipole must be $I/2$.

If we say the radiation resistance of the folded dipole is R_r and that of the non-folded dipole is R_n then $P = (I/2)^2 R_r = I^2 R_n$ so $Rr = 4R_n$ and $R_n = 73\ \Omega$ so $R_r = 292\ \Omega$, and this resistance is presented to the dipole's terminals. It follows that a folded dipole with N λ/2 continuous conductors connected across the ends of the non-folded element has N^2 times the terminal resistance of a λ/2 non-folded dipole.

A different ratio results if dipoles A and B are made from material of different thickness and this can be used to tailor the terminal resistance.

λ/4 and λ/2 transmission lines as transformers

Considering again the 'transmission line relation' that gives the impedance Z_x at a point on a transmission line of characteristic impedance Z_0 some distance x from the load of impedance Z_L. For a lossless line:

$$Z_x = Z_0 \frac{Z_L + jZ_0 \tan \beta x}{Z_0 + jZ_L \tan \beta x} \quad \text{… where } \beta = 2\pi/\lambda.$$

If $x = \lambda/4$ then $Z_x = Z_0^2/Z_L$ so a quarter-wavelength transmission line transforms the load impedance to a different value. If the objective is to match the load impedance Z_L to some other, arbitrary, 'input' impedance Z_x, a solution is to connect them together with a quarter-wavelength transmission line having a characteristic impedance equal to the square-root of the product of the input and load impedances: $Z_0 = \sqrt{(Z_x.Z_L)}$. The mismatch created at each end of this line, on account of the abrupt change of impedance, cancels the effect of the mismatch at the other end and maximum power transfer is achieved so long as the length of the line is close to λ/4. This 'quarter-wave transformer' principle [51] finds use in power dividers, for example, used to connect an array of dipoles to a single 50 Ω cable.

Because of the reciprocal relationship between Z_x and Z_L, if $Z_L = 0$, a short-circuit, this is transformed to $Z_x = \infty$, an open-circuit by a λ/4 transmission line of any characteristic impedance, and *vice versa*.

If $x = \lambda/2$ then $Z_x = Z_L$ so a half-wavelength transmission line transforms the load impedance 1:1 but reverses the phase of the voltage from one end to the other. Z_0 can have any value in this case as well.

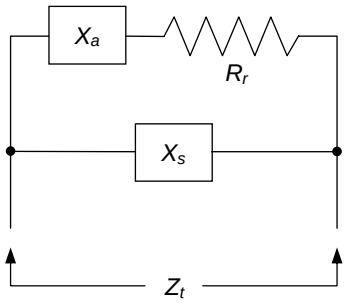
14.1 'Reactance compensation' in the folded dipole

In addition to being somewhat "fatter" than its non-folded counterpart, the folded dipole has another property that further enhances its potential bandwidth. With respect to the difference-mode current, each of the limbs behaves as a $\lambda/4$ transmission line stub with a short-circuit at its far end, so a pair of these stubs, connected in series with one another, appears in parallel with the dipole's terminals.

I covered transmission lines as resonant stubs in a box in Section 11, and we have seen the similarity between the terminal reactance of a $\lambda/2$ dipole and that of a $\lambda/4$ open-circuit stub. The terminal reactance of a $\lambda/4$ short-circuit stub is in some respects 'complementary' to this – it always has the opposite sign – so, if a short-circuit stub is connected across the dipole's terminals, the reactances tend to cancel over a range of frequencies around the dipole's first resonance. By appropriate choice of the characteristic impedance of the short-circuit stub, and its length, the resulting effective terminal reactance can be reduced to an 'S'-shaped curve either side of 0 Ω around the resonant frequency.

An 'S'-curve with small excursions can provide an acceptable match over a greater bandwidth than the monotonic slope of the reactance of the un-compensated dipole. Page alludes to this principle [52], Kraus mentions it [53], and I'll demonstrate it using some approximate modelling.

The antenna will be modelled as a lossless open-circuit stub connected in series with the radiation resistance, R_r. This stub has a pure reactance $X_a = -Z_a \cot \beta l_a$ where $\beta = 2\pi/\lambda$ and l_a is the stub length expressed in the same unit as λ. Z_a is the (average, uniform) characteristic impedance of the 'transmission line' that forms this open circuit stub; that is, the pair of dipole limbs. Of course, the radiation resistance in a real dipole is distributed throughout the length of the open-circuit stub, and the characteristic impedance of the stub is anything but uniform, so this is a greatly simplified model. There is an alternative version of the 'transmission line relation' specifically for lossy lines, but then it is assumed that the loss is uniform along the length of the line, which it most likely isn't in the case of a resonant dipole.



The effect of the difference-mode current in the two limbs will be modelled as a lossless short-circuit stub connected in parallel with the antenna terminals, having a pure reactance $X_s = Z_s \tan \beta l_s$ where l_s is the stub length expressed in the same unit as λ and $\beta = 2\pi/\lambda$ again. Z_s represents the characteristic impedance of the transmission line that forms the short circuit stub but, since there are actually two stubs connected in series, Z_s is made twice the characteristic impedance of the two-conductor transmission line from which the folded-dipole is constructed.

I should mention here that as soon as we begin to consider reactances connected in parallel, strictly, we should talk in terms of 'susceptance' the reciprocal of reactance, and 'admittance' the reciprocal of impedance. However, I intend to press on using simplified language, echoing my cop-out back in

Section 3, and I hope any purists reading this will forgive me. The result of shunting $R_r + X_a$ with X_s is a terminal impedance Z_t where:

$$\frac{1}{Z_t} = \frac{1}{\mathrm{j}X_s} + \frac{1}{R_r + \mathrm{j}X_a} = \frac{R_r + \mathrm{j}(X_a + X_s)}{X_s(-X_a + \mathrm{j}R_r)}$$

$$\therefore\ Z_t = \frac{X_s(-X_a + \mathrm{j}R_r)}{R_r + \mathrm{j}(X_a + X_s)} = \frac{X_s(-X_a + \mathrm{j}R_r)}{R_r + \mathrm{j}(X_a + X_s)} \cdot \frac{R_r - \mathrm{j}(X_a + X_s)}{R_r - \mathrm{j}(X_a + X_s)}$$

$$Z_t = \frac{R_r X_s[-X_a + (X_a + X_s)] + \mathrm{j}X_s\left[R_r{}^2 + X_a(X_a + X_s)\right]}{R_r{}^2 + (X_a + X_s)^2} = \frac{R_r X_s{}^2 + \mathrm{j}X_s\left[R_r{}^2 + X_a(X_a + X_s)\right]}{R_r{}^2 + (X_a + X_s)^2}$$

At the quarter-wave resonance of the open-circuit stub (i.e. the dipole), $l_a = \lambda/4$ so $\beta l_a = \pi/2$ and $X_a = 0$ making $Z_t = R_r X_s (X_s + \mathrm{j}R_r) / (R_r{}^2 + X_s{}^2)$ which is the parallel combination of R_r and X_s, as would be expected. It is useful to define the frequency at which this occurs as f_0, the *design centre-frequency*. Then $l_a = c/4f_0$ where c is the velocity of light, but $\beta = 2\pi f/c$ where f is the operating frequency, whatever that might be, so $\beta l_a = (\pi/2) f/f_0$ and $X_a = -Z_a \cot(\pi/2) f/f_0$.

At the quarter-wave resonance of the short-circuit stub, $l_s = \lambda/4$ so $\beta l_s = \pi/2$ and and $X_s = \infty$, whereupon terms in $X_s{}^2$ dominate and all cancel leaving $Z_t = R_r + X_a$, that is, the short-circuit stub 'disappears' as would be expected. Each of the short-circuit stubs in a conventional air-spaced folded dipole has length equal to the limb length so it is equal to the length of the open-circuit stub, $l_s = l_a$, and these two phenomena occur at the same frequency f_0. Then $X_s = Z_s \tan(\pi/2) f/f_0$. In practice, semi-circular ends to the stubs may introduce an approximation here.

The characteristic impedance of a symmetrical transmission line formed from two parallel cylindrical conductors is given by [54]:

$$Z_s = \frac{276}{\sqrt{\varepsilon_r}} \log_{10} \frac{D}{a}$$

… where a is the radius of each of the conductors, D is the separation of their centres, and ε_r is the relative permittivity of the medium between the conductors; $\varepsilon_r \approx 1$ for air, as before.

I'll take the practical example of a Band II $\lambda/2$ folded-dipole of length 1.5 metres between the extremities of its semi-circular ends. The conductors are of 19 mm diameter, 9.5 mm radius, and their centres are separated by 72 mm. In this case the result is 243 Ω but to account for the two stubs in series we'll take $Z_s = 486\ \Omega$.

This example has $l_a = l_s = 750$ mm so $f_0 = c/4l_a = 100$ MHz. Actually, the geometric mean frequency of Band II is $\sqrt{(87.5 \times 108)} = 97.2$ MHz and the reason why the limbs have been designed slightly shorter will be covered a little later (in Section 15.2).

To assign a value to Z_a we first need to consider a single cylindrical conductor which has the same electrical properties as the parallel pair of rods in each limb of the folded dipole. Using Schelkunnoff's expression for $n = 2$ from Section 13, the equivalent cylindrical dipole would have a limb diameter of $d \approx 52$ mm, and the length of the dipole $l = 2 \times 750$ mm, so $l/d \approx 30$.

Looking back at Hallén's terminal reactance curves at the beginning of Section 11, the slope around the first resonance appears to be related to the value of l/d somewhat logarithmically. For a value of $l/d \approx 30$ the slope of the reactance can be estimated with a ruler to be something like j800 Ω over a range of l/λ from 0 to 1. In this case l is the length of the dipole which is twice the length of the equivalent open-circuit stub, so over a range of $\Delta(l_a/\lambda) = 1$ or $\Delta l_a = \lambda$ the reactance excursion would be something like $\Delta X_a = 1600\ \Omega$.

Since the reactance of the open-circuit stub is $X_a = -Z_a \cot \beta l_a$, the slope of this is found by differentiating with respect to l_a.

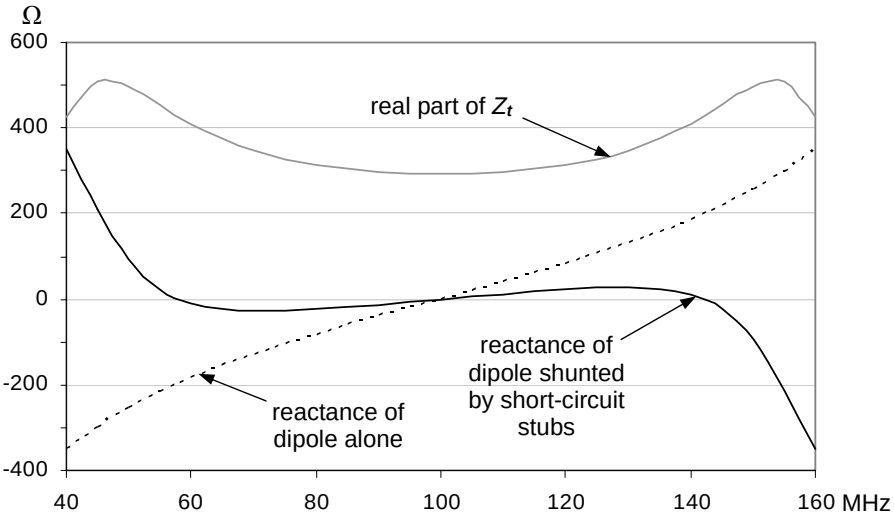
Using $\frac{\mathrm{d}}{\mathrm{d}x}\cot v = \left(-\csc^2 v\right) \cdot \left(\frac{\mathrm{d}v}{\mathrm{d}x}\right)$: $\frac{\mathrm{d}X_a}{\mathrm{d}l_a} = Z_a \beta \csc^2 \beta l_a$ and $\frac{\Delta X_a}{\Delta l_a} = \frac{1600}{\lambda}$ so $Z_a \approx \frac{1600}{\lambda} \beta \csc^2 \beta l_a$

… equating infinitesimals to incrementals. At this first resonance, $\beta l_a = \pi/2$ so $\csc^2 \beta l_a = 1$. Then $Z_a = 1600\beta/\lambda$ but $\beta = 2\pi/\lambda$ so $Z_a = 1600/2\pi = 254.7\ \Omega$.

The full set of practical values is then: f_0 = 100 MHz, Z_a = 255 Ω, Z_s = 486 Ω, l_a = l_s = 750 mm and it will be assumed that R_r = 292 Ω, as explained earlier. Variation of R_r with frequency will be neglected. Entering this lot into a spreadsheet containing expressions for the real and imaginary parts of Z_t:

$$Z_t = \frac{R_r X_s^{\,2}}{R_r^{\,2} + (X_a + X_s)^2} + \mathrm{j}\frac{X_s\left[R_r^{\,2} + X_a(X_a + X_s)\right]}{R_r^{\,2} + (X_a + X_s)^2}$$

… enables them to be calculated numerically with changing f and plotted either side of the design centre-frequency f_0, as shown below. The vertical axis applies to all the curves and for the reactances the values are multiplied by j.



Thus the 'reactance compensation' effect of the short-circuit stubs is demonstrated although, in practice, the result may not be so smoothly-changing or so symmetrical. Also, this shouldn't be interpreted as meaning such a folded dipole can be operated over 70 to 130 MHz, or more, because variation of R_r and other factors neglected in the simple model would limit the usable bandwidth.

The folded dipole, then, provides inherent 'stub compensation' of its own terminal reactance. I haven't yet been able to determine whether this fact was discovered or invented, but if the folded dipole was invented with this in mind from the outset it strikes me as a remarkable piece of lateral thinking – even reciprocal thinking! Alternatively, if it was discovered, it must surely rate as one of the more fortunate accidents of physics. Of course, the difference-mode current that needs to flow at frequencies either side of the first resonance in order to mitigate the dipole's reactance doesn't give rise to any significant additional radiation.

Where the 'S'-curve crosses j0 Ω at the two frequencies either side of f_0, $R_r^2 = -X_a(X_a + X_s)$ so $\tan(\pi/2)\, f/f_0 = \pm Z_a/\sqrt{(Z_a Z_s - R_r^2)}$ and $f = (2f_0/\pi)\arctan \pm[Z_a/\sqrt{(Z_a Z_s - R_r^2)}]$ which yields 58.2 MHz and 141.8 MHz for the above example, interpreting $\arctan(-\theta)$ as $\arctan(\pi - \theta)$. The separation of these two frequencies is maximised when $Z_a/\sqrt{(Z_a Z_s - R_r^2)}$ is minimised so, for given (assumed constant) Z_a and R_r the frequency separation is increased by increasing Z_s. However, the slope of the 'S' around f_0 also increases so the positive and negative excursions of reactance within the 'S' become greater, reducing the usable bandwidth. The optimum value for Z_s seems to be around 400 to 500 Ω for this simplified example, suggesting the 72 mm separation of the rods is a good choice.

15. Baluns

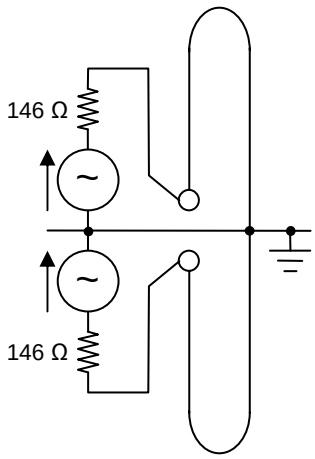
By now, you might be wondering if I've missed out an important step. A dipole antenna is most often designed to operate at its first resonance because this provides useful radiation patterns and avoids rapid changes of its terminal impedance with changing frequency. In practice, a folded dipole is often preferred because its limbs appear electrically "fatter" and it provides inherent compensation of its own terminal reactance, both favouring achievement of an acceptable match over a broad bandwidth. But a λ/2 folded dipole presents a terminal resistance of nearly 300 Ω whereas the co-axial cables used extensively in the UK for connecting antennas to transmitters and receivers have characteristic impedances of 50 Ω or 75 Ω; values close to the optima for power handling or low loss.

The missing link is the interface component that many dipole/co-axial feeder systems need to function correctly: the balanced-to-unbalanced transformer or 'balun'.

15.1 The λ/2 transmission line balun

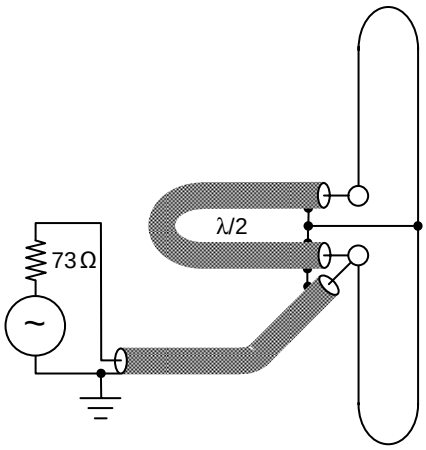
A symmetrical, centre-fed dipole is fundamentally a balanced device, so normally whatever current goes in one terminal the same amount comes out the other terminal, but in some circumstances a co-axial cable can be considered an unbalanced device. Connecting the terminals of a dipole directly to a co-axial cable can cause problems because this connects the outer co-axial conductor – the 'screen' – to one of the antenna terminals. The outer surface of the 'screen' is no less an antenna than the dipole limb so, in the transmitting sense, some of the current conveyed by the cable, which is normally carried on the inner-surface of the 'screen' (and the outer surface of the inner conductor), is diverted away from the dipole limb. Meanwhile, all the current carried by the inner conductor flows into the other dipole limb. So the current distribution in the dipole loses its symmetry, and current flows in an additional radiator, the outer surface of the screen, that can have any imaginable geometrical relationship to the dipole. Small wonder then that the radiation pattern becomes distorted – even *variable* if the far end of the down-lead terminates in your living room!

The solution is to use some kind of balun at the point of connection to the co-axial cable. Many novel designs are used in domestic and professional antennas; the use of a half wavelength transmission line seems popular (I covered λ/4 and λ/2 transmission lines as transformers in a box in Section 14).



The symmetry of a folded dipole suggests that if the equivalent generator connected directly to its terminals were divided into two generators, each of half the voltage adding in series, then the voltage at their junction would be equal to the voltage half-way across the λ/2 continuous conductor – so these two points could be connected together, and to earth, without affecting operation of the antenna. Then half the radiation resistance, 146 Ω, would appear between either of the folded dipole's terminals and earth. The resistors shown in the diagram to the left represent the internal resistances of the matched 'half-generators'. These generators could be moved away and connected to their respective dipole terminals using two equal lengths of 146 Ω co-axial cable.

With respect to earth, these two equal-voltage 'half generators' are in antiphase with one another so, with attention to matching, the same effect could be achieved using a single generator feeding two cables in parallel, one of which was λ/2 longer than the other. Indeed, a single co-axial cable can be used to feed one terminal of the dipole, with another cable, λ/2 long, joining that terminal to the other, with the screens connected together and to the centre of the λ/2 continuous conductor, as shown to the right. In this case, the half-wavelength transmission line (cable) presents an impedance of 146 Ω, with respect to earth, at the terminal to which the feeder is connected, which itself presents 146 Ω with respect to earth. The parallel impedance is half, 73 Ω, which matches well to a 75 Ω cable. The phase reversal caused by the λ/2-long transmission line guarantees that the dipole limbs are appropriately phased.



Of course, the transmission line is only λ/2-long at a specific frequency and either side of this design frequency the performance of the complete antenna will be compromised to some extent, by the introduction of spurious reactance and, possibly, asymmetry. Nevertheless, this is the basis of the baluns used in many television receiving antennas and, because the co-axial screen is connected to the point of symmetry, it works well in decoupling current flow on the outside of the screen – but it requires a *folded dipole* to achieve a match to a 75 Ω down-lead.

I have presented this argument on the basis that the antenna's terminal impedance can be divided. Interestingly, Page [55] presents it on the alternative basis that the $\lambda/2$ transmission line connected to the end of the feeder has the effect of a transformer that transforms the feeder impedance 1:4 whilst also providing the balun action. He provides a more-general result taking account of reactance in the folded dipole's terminal impedance, so what happens at frequencies either side of resonance can be deduced, but, oddly, he shows the balun feeding a *non-folded* dipole. I don't suppose anyone manufactures 18 Ω co-ax!

Two further advantages of the folded dipole are that the 'neutral' point mid-way along the continuous conductor, shown connected to earth, provides a sturdy means of mounting that doesn't interfere with the terminals, and the continuous path to earth provided by each folded limb provides a short-circuit to static build-up without need for any additional shunt choke. Small wonder, then, that folded dipoles are ubiquitous.

15.2 Variants

Since the characteristic impedance of the half-wavelength transmission line can have any value, some manufacturers of domestic Yagi-Uda antennas have used the very compact and inexpensive 'insulated wire wound around a strip of metal' type of transmission line. The wire was often co-ax inner conductor, wound bifilar to avoid introducing a large series inductance and the metal strip was connected to the metallic boom that supports the dipole elements. The screen of the down-lead was connected to this.

Others use a printed-circuit track of appropriate length, often in the shape of a horse-shoe, usually with no earth plane at all. Of course the $\lambda/2$ line can have any characteristic impedance, but one might expect it needs to be constant. Perhaps because this conductor of the transmission line is so compact it doesn't matter that the 'other conductor' is the rather remote antenna boom and down-lead screen. I'm not aware that this type of balun has a particularly bad reputation though, and the performance of antennas equipped with it probably leaves less to be desired than that of the Yagi antennas with no balun still offered by some do-it-yourself chain stores.

Concern about the absence of baluns in some domestic television antennas increased greatly following the launch of DTT when, from 1999, it became evident that reception was being impaired by impulsive interference radiated by appliances in the home with mains switches, such as thermostats. At our 2003 R&D Open Days, Spectrum Planning Group displayed domestic installations on a mock-up of a 'house' with 'good' and 'bad' antennas, with and without baluns. We also connected them with down-leads made from relatively good and bad types of cable, the former having a continuous copper-foil screen and the latter only a loosely woven wire screen. The demonstration was completed with an impulse generator that periodically switched the mains supply to an inductive load (actually a large transformer) via a length of mains cable running near the down-leads. This demonstrated the point clearly.

In Spectrum Planning Group, over the years we have bought from a well-known and respected manufacturer many folded dipoles for professional use with 50 Ω cable, some as the driven elements in Yagi-Uda arrays, operating in Band II, Band III and Band IV/V. Do these contain a transformer as well as a balun? The point of connection of the dipole terminals to the captive 50 Ω co-axial cable is moulded in polythene so the internal construction is not at all obvious. I was going to end this section here, writing '… pass me the hacksaw!', but we found a damaged one and a colleague removed the polythene moulding with a sharp knife. What this revealed was … the very compact and inexpensive 'insulated wire wound around a metal *rod*' type of transmission line. The winding is bifilar to avoid introducing a large series inductance and the metal *rod* is connected to the metallic boom that supports the dipole at its centre; the screen of the down-lead is connected to this. Later, I happened to spot a description of this in a library book [56] where it is referred to as the 'J-Beam inverse balun'.

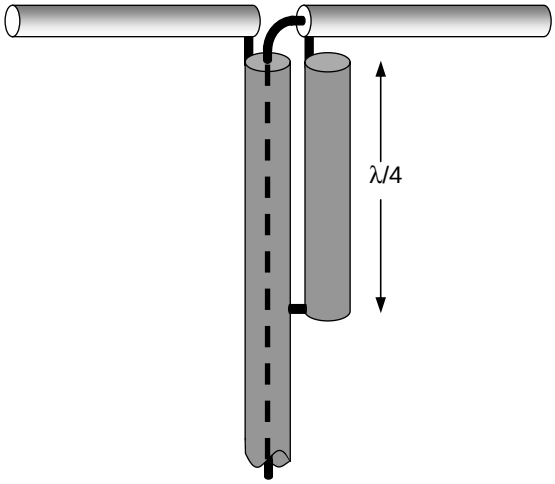
No surprises here then, but what about the impedance – there's no sign of any $\lambda/4$ transformer (one would require a characteristic impedance of 61 Ω)?

This Band II folded dipole has a total length of 1500 mm, half an air-wavelength at 100 MHz whereas the geometric mean frequency of Band II is 97.2 MHz, as noted before, which would require the length to be 1544 mm. Reference to Hallén's curves in Section 11 shows the terminal resistance can be reduced by shortening the dipole, at the expense of introducing a capacitive series reactance. It would appear, then, that this dipole has been shortened by about 3%, which may be enough to

make the terminal resistance 200 Ω, to match a 50 Ω cable via the λ/2 transmission-line balun. The additional reactance is probably 'resonated out' by a bit of additional series inductance provided by the (non-bifilar) ends of the balun 'winding'.

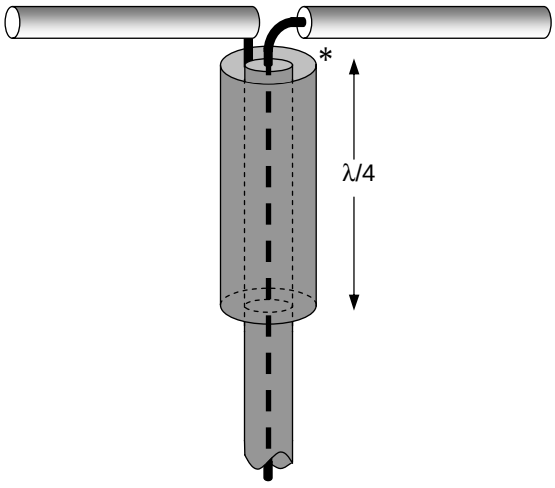
15.3 Some other forms of balun

This is but one of the many ingenious forms of balun that have been thought up and I'll illustrate a few more here.



The 'Pawsey stub', shown to the right, is a refinement of the theme 'two can play at that game'! Since one dipole limb is connected to the co-axial screen, if a dummy cable screen of exactly the same dimensions, laid-out alongside the real down-lead, were connected to the other limb then symmetry would be restored. A saving is achieved by shortening this to λ/4 length and connecting its far end to the outside of the screen of the real down-lead. The resulting symmetrical λ/4 short-circuit stub presents a high impedance across the dipole terminals, at resonance, and reactance compensation at frequencies nearby. There is no impedance-transformation effect so this type of balun is ideal for interfacing a non-folded dipole to a 75 Ω co-axial cable.

In Spectrum Planning Group we have some cylindrical dipoles terminated in Pawsey stubs built in-house many years ago for calibrating the gains of receiving antennas used for UHF field-strength measurement. The family of three different sizes, shown to the right, are for operation at different frequencies in Band IV/V corresponding to the earliest-used three receiving antenna groups: Group A, Group B and Group C/D – see [57]. In each case the stub is made from two pieces of aluminium alloy 'U' channel material separated by a (black) polythene slab, and the 50 Ω co-axial cable is connected part-way down the stub in order to provide the requisite impedance transformation by auto-transformer action.

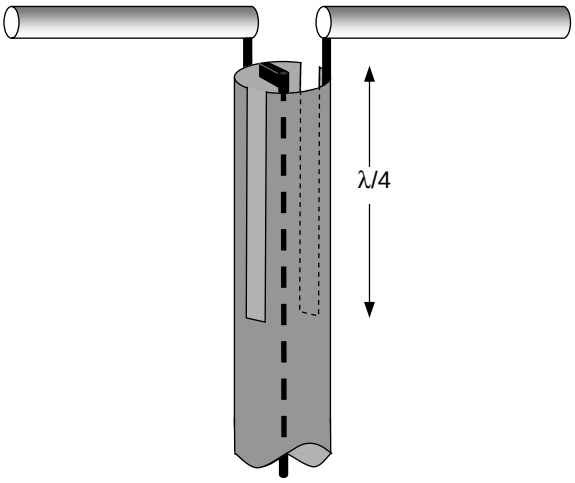


The λ/4 sleeve choke or 'bazooka' is another λ/4 short-circuit stub, co-axial this time, that presents a high impedance to the current that would otherwise flow on the outside of the down-lead screen. A cylindrical conducting sleeve surrounds the down-lead; it is open at the top and closed the bottom end where it is connected to the outside of the co-axial screen, as shown to the right. The top of the sleeve needs to be close to the dipole to avoid exposing part of the screen on which unwanted current could flow. The ratio of diameters of the conductors making up the co-axial stub affects its characteristic impedance, as does the presence of any dielectric between them (e.g. the cable sheath), and this impedance has some effect on the performance of the 'choke' with changing frequency (as mentioned in Section 11).

In this case there is no reactance compensation because the stub is not connected across the dipole

terminals. I'd expect that if the terminal of the dipole limb connected to the co-axial inner conductor were also connected to the sleeve where I've shown an asterisk this would form a co-axial Pawsey stub, but I've never seen this applied in practice.

A form of co-axial Pawsey stub used extensively in VHF/UHF transmitting equipment has a narrow $\lambda/4$-long slot cut on each side of the co-axial screen as shown here. The separated halves of the outer conductor at the end of the line provide the balanced terminals, which I've shown connected to the dipole. The inner co-axial conductor is connected to one of these terminals. Internally this still operates as a co-axial line, but current flowing on the outside surface of the outer conductor is subjected to the short-circuit $\lambda/4$ stub. This is sometimes shown surrounded by a $\lambda/4$ sleeve choke as in the previous example.

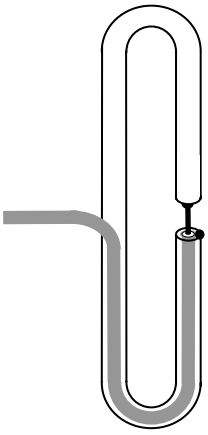


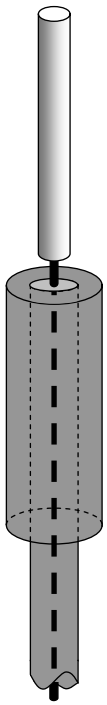
Many other forms of balun are to be found in the literature ([59] for example), some more complicated but offering further possibilities for *in situ* impedance transformation.

Another way to choke current on the outside of a co-axial screen is to wind part of the down-lead into a coil, perhaps with a magnetic core of suitable composition to reduce its size (e.g. ferrite [58] or 'iron powder' in the form of a rod, ring or clamp). This forms a 'common-mode' choke, of which there are varieties using many other styles of 1:1 transformer, transmission line and lumped component.

15.4 No balun required

A further variation is to surround the end of the co-axial cable by a $\lambda/4$ sleeve, closed and connected to the co-axial screen at the top and open at the bottom, and to connect a $\lambda/4$-long dipole limb co-axially to the end of the inner conductor. The result is then a sleeve dipole, as shown to the right, in which the sleeve acts simultaneously as a choke and a "fat" lower dipole limb. I will mention some other forms of co-axial antenna later but, really, this is an extensive topic in its own right and I don't have space to do it justice here.



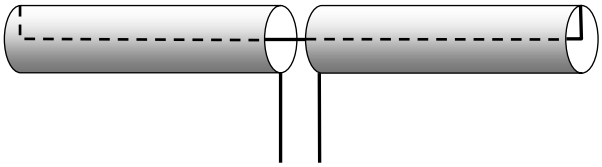


There is a way to feed a folded dipole co-axially which avoids having the co-axial screen connected externally to one of the dipole's terminals; this can be found in some of the popular text books. The cable is passed through one half of the tubular folded dipole, from a hole the 'neutral' mid-point to the terminals, to which its two conductors are then connected as shown to the left. This achieves the same effect as a balun and there is probably no need to connect the outside of the screen to the neutral mid-point. Unfortunately, this arrangement would require 292 Ω co-axial cable to achieve a match directly. Alternatively, to match the dipole to a 75 Ω down-lead using a single quarter-wave transformer would require a length of 148 Ω cable. Both these characteristic impedances are uncommon.

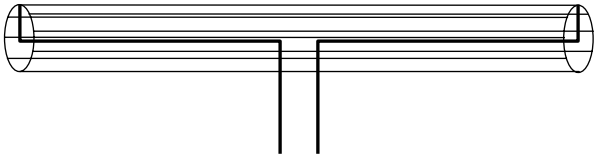
Of course, interaction between an antenna and its down-lead cannot always be completely overcome using a balun or a choke, or one of these co-axial arrangements. Any conductor in the vicinity of a transmitting antenna is likely to have currents induced in its surface by the field surrounding the antenna, giving rise to re-radiation. It was shown in Section 12 that minimum field strength occurs on the axis of a dipole so the down-lead from a co-axial antenna needs to remain aligned co-axially for some distance (e.g. a few wavelengths) in order to preserve symmetry. For a symmetrical dipole, like the one shown above to the left, the down-lead needs to be kept in the plane of symmetry for a similar distance.

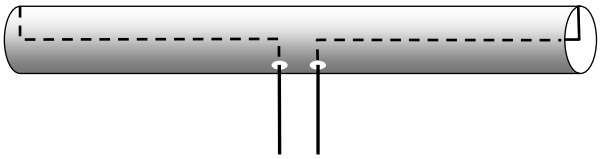
16. Some other types of dipole

The reactance compensation principle of Section 14.1 can usefully be applied to any other type of dipole or monopole antenna by using an external stub or, by clever design, building the stub into part of the antenna. Kraus [53] describes stub compensation of a non-folded dipole using an external open-circuit $\lambda/2$ transmission line. Page [59] illustrates how a non-folded dipole made from two open cylinders, each $\lambda/4$ long, can be given the stub compensation of a folded dipole by passing a wire through the axis and connecting it to the ends of both limbs, as shown to the right. There is no multiplication of the terminal resistance in this case because the wire forms the inner conductor of a co-axial line and cannot radiate. This arrangement is sometimes known as the 'double bazooka' antenna.

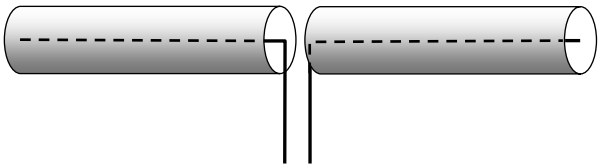


The *dual* of this arrangement would correspond to a folded dipole with an infinite number of $\lambda/2$ continuous conductors connected across its ends. Increasing the number of such discrete conductors would progressively increase the terminal resistance so, in the limit, a continuous conductive cylinder surrounding the non-folded dipole, as shown here, would make the terminal resistance infinite (at the $\lambda/2$ resonance). Considered as a pair of short-circuit $\lambda/4$ co-axial stubs connected in series, this is obvious. This would only work as an antenna if somehow current could be encouraged to flow on the outside surface of the cylinder but, even then, it might be difficult to match it to a sensible impedance.



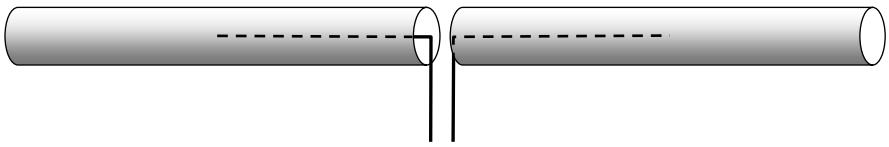


One way to force current onto the outside surface would be to cut the tube in two, reverting to two $\lambda/4$-long cylinders. If the ends of the axial wires were then disconnected from the ends of the tubes the result would be a pair of open-circuit $\lambda/4$ co-axial stubs.

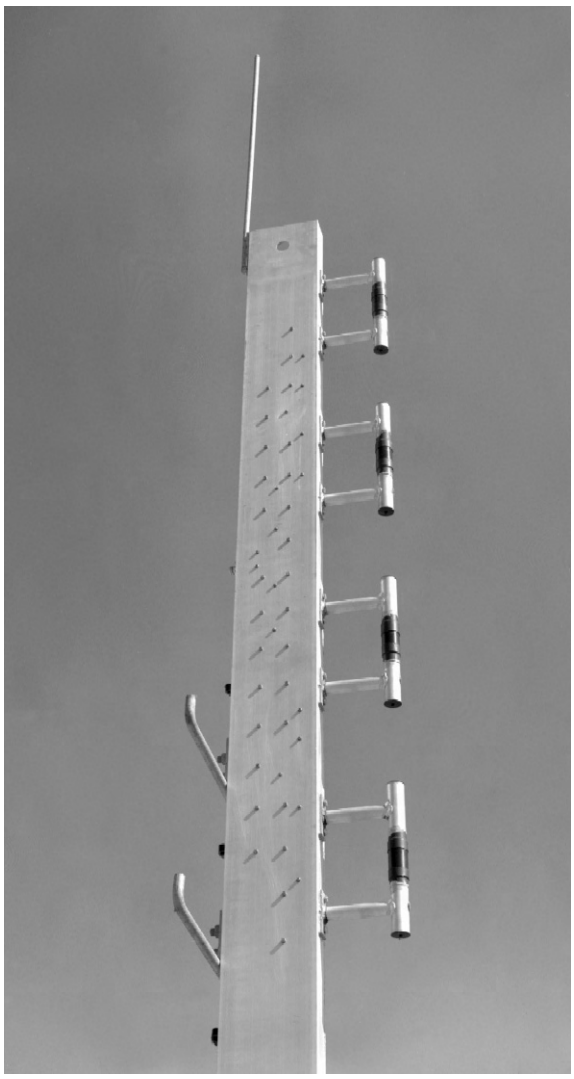


16.1 A full-wave dipole and the 'four-lambda cardioid' antenna

Interestingly, the slope of the terminal reactance of such a stub is 'complementary' to that of a dipole of length approaching λ; the full wavelength. So if the length of each tube is then increased towards $\lambda/2$ but the wires kept at $\lambda/4$ length, as illustrated below, the resulting approximately 'full-wave' dipole can be arranged to present a predominantly resistive termination over some bandwidth. Another 'S'-curve, but by a series combination in this case.



Notwithstanding all I've said about the ubiquity of $\lambda/2$ dipoles, this principle is employed in a design of collinear array originally supplied to the BBC by EMI Sound & Vision Equipment Ltd. and used at more than 100 of the low-power ($\leq$50 W ERP) television relay stations around the UK. This uses 4 full-wave dipoles, which probably achieves a significant simplification of the feeding system relative to using 8 half-wave dipoles, and a reduction in its loss. Different versions were made for Band IV and Band V.



Shown to the left, each of the four elements is such a full-wave dipole, comprising a pair of ~λ/2-long tubes, 25.4 mm in diameter. These tubes are attached to the supporting 'cantilever' – a length of aluminium alloy 'U'-channel – by a pair of metal support legs (of length somewhere between λ/4 and λ/2), connected to the tubes about half way along their length (but not exactly, in the Band V version pictured). It's not obvious to me how this arrangement works but, conceivably, the legs and cantilever between them could form a short-circuit stub connected across part of the dipole, that might provide further reactance compensation.

Each dipole is fed by a short length of co-axial cable passing through the lower leg to the mid-point of the dipole which is protected by a black plastic sleeve. Documentation to hand doesn't show clearly how the two co-axial conductors are terminated but a text description indicates clearly enough that the inner conductor passes into an open-circuit stub inside the upper dipole limb. The outer co-axial conductor is probably connected to the end of the lower dipole limb at the mid-point; it would be difficult to achieve an open-circuit stub on this side. The requisite resistance transformation is most likely built into the four-way power divider, a printed circuit assembly encased within the cantilever.

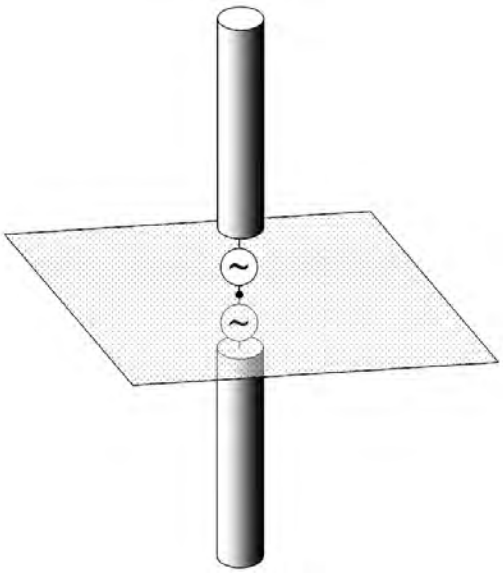
For the Band IV version, the radiation pattern in the vertical plane (VRP – Vertical Radiation Pattern) has a main lobe with a beamwidth of only 14° and sidelobes, resulting from the product of the element's VRP and an 'array factor' [61] accounting for the number of elements, their physical arrangement including their separation, and the currents fed to them. This is tailored to 'beam tilt' the main lobe downwards by about 6° by applying different amounts of phase shift (i.e. line length) to the four currents. Also, the null that would otherwise exist between the main lobe and the first sidelobe has been 'filled' to -20 dB (relative to the main-lobe peak) by manipulating the relative amplitudes of the currents. The pattern in the horizontal plane, a single nominally 'cardioid' (heart-shaped) lobe, has a beamwidth of 180° resulting from interaction between the dipoles and the cantilever. The peak gain of this array is stated to be about 9 dBd.

The 'input reflection coefficient' is claimed to be less than 10% over the whole of Band IV or Band V for each of the two versions. When a load of impedance Z_L is connected to an otherwise matched system with a characteristic impedance Z_0 the voltage reflection coefficient is given by $\rho_v = (Z_L - Z_0)/(Z_L + Z_0)$. The 'return loss', that is the proportion of the power fed to the load that is reflected back, is $RL = 20\log_{10}|\rho_v|$ (dB) which in this case has a value >20 dB. Values of this order are considered adequate for transmitting antennas in many practical applications.

17. Monopole antennas

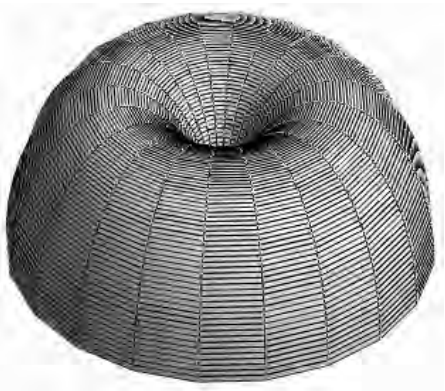
Another common type of antenna is the λ/4 monopole, usually operated above some sort of 'ground plane' (or 'earth plane').

A centre-fed λ/2 dipole is a symmetrical structure with two identical limbs so an equatorial plane can be considered that bisects the structure, with a whole limb on each side. If expressions are derived for the fields surrounding a dipole [62] it is found that the electric field in the vicinity of this plane is always aligned perpendicularly to it – there *are* radial components of E-field, but they depend on the cosine of the angle from the axis of the dipole and vanish at 90°, that is, at the equatorial plane.



It follows that if a large, thin, perfectly-conducting sheet were placed exactly in this plane, cutting the dipole and the generator feeding it in two, then nothing would be different – there would be no electrical or electromagnetic effect on the operation of the dipole. If the sheet had infinite size, dividing space in two and separate 'half-generators' were provided in each half, the radiation in one half-space would be completely independent of what was going on in the other. Indeed, one half-dipole and its half-generator could be removed without affecting radiation by the other.

What would then remain is a $\lambda/4$ monopole perpendicular to a large conducting plane. Its radiation pattern in 3 dimensions is half that of the $\lambda/2$ dipole – like the top half of a doughnut sliced through its 'equator' as shown below. Neglecting losses, if the whole of the input power were fed to this monopole it would be radiated through half the solid angle and the Poynting vector would have twice the value, so this antenna is said to have 3 dB gain over a $\lambda/2$ dipole. In practice, the limited extent of the ground plane and other imperfections like its finite conductivity usually result in 'undercutting' of the VRP at angles close to the ground plane; that is, reduction of the transmitted field strength, effectively creating an upward-tilting lobe.



An alternative explanation is the 'method of images' [63], in which the conducting sheet is considered as a 'mirror' in which a reflected image of the monopole contributes to the resulting radiation.

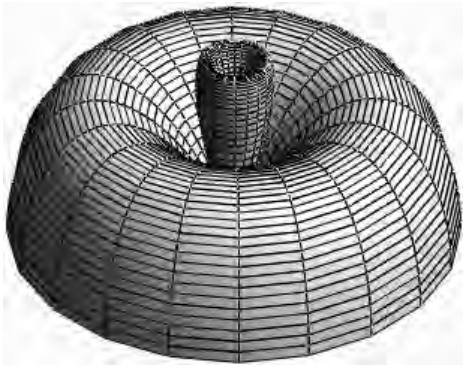
For a thin $\lambda/4$ monopole over a large ground plane, the radiation resistance is half that of the thin $\lambda/2$ dipole: 36.5 Ω, although some books give it as 37.5 Ω (probably having already rounded-up the dipole's value to 75 Ω). This is almost intuitive because the integral of the Poynting vector for the case of the $\lambda/2$ dipole, as outlined in Section 5, would take equal contributions from both halves. The expression $P = (I_0/\sqrt{2})^2 Rr$ shows that $Rr \propto P$ so half the radiation resistance accounts for half the total power radiated, that is the radiation from each half of the dipole. Therefore, a $\lambda/4$ monopole has fundamentally half the radiation resistance of a $\lambda/2$ dipole. In practice, again, the characteristics of the real ground plane can affect this.

Two well-known practical implementations are the VHF/UHF 'ground plane' antenna, shown to the left using skeletal radial rods ($\geq \lambda/4$ long), and the medium-wave transmitting antenna using a guyed lattice mast and wire 'radials' buried in the earth. Sometimes the monopole element is folded to increase its radiation resistance, which may simplify matching.

A monopole can also be designed $\lambda/2$ long; sometimes referred to as an 'end-fed dipole'. This requires an impedance transformer at the base of the element (lumped component or transmission line stub) but a ground-plane may be unnecessary since the current distribution on the element is a complete half-sinusoid and the current at the feed point should be very small (and the voltage correspondingly large). However, some kind of choke may be needed to prevent current being passed to the outside of the screen if a co-axial feed cable is used and the 'secondary' of the transformer is effectively connected between the element and the co-axial screen – if a multiple of $\lambda/2$ long, the outside of the screen could present an impedance similar to that of the element. Clever

designs combine the choke function within the transformer. An early design of this type is the 'Zeppelin' antenna [64], originally designed for HF use, to be trailed on the end of a transmission line some distance behind or beneath a Zeppelin balloon keeping the high voltages well away from the hydrogen contained therein. This type of antenna is also referred to as a 'J antenna' [65] or a 'J-pole antenna' (but is not quite the same design as what I'll describe in the next section).

A further variation is to make the element $5/8\,\lambda$ long. The additional length increases the aperture area and affords a little gain over a $\lambda/4$ element, squashing the main lobe and giving rise to an upward-pointing sidelobe, as shown to the right. The terminal impedance has a resistive part close to $50\,\Omega$, and there is a capacitive series reactance that is easily mitigated using a series inductor 'loading coil'.

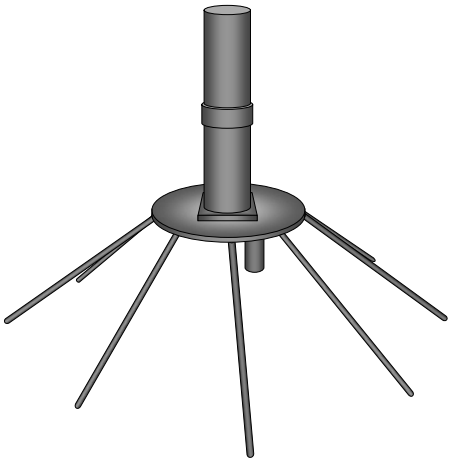


There are also numerous techniques for increasing the radiation efficiency of a monopole of limited length such as adding a capacitance "hat" to the top of the element or lumped series inductance part way up it; another extensive topic in its own right (e.g. see [66]). Loading can also be applied in a distributed fashion, as in the normal-mode helix antenna [67] in which inductance is added continuously along the length of a short monopole to help mitigate its capacitive terminal reactance.

In some practical cases, notably devices using normal-mode helix antennas or whip antennas such as hand-held transceivers and domestic portable receivers, or receivers that use the headphone cable as an antenna, the presence of something approximating to a large conducting ground plane is not evident. However, there is usually some piece of metalwork connected to the other terminal of the equipment and capacitance between this and the body of the person using the equipment, or connection to some other cable (e.g. for mains power) can provide a so-called 'counterpoise', acting as a skeletal ground plane or the other limb of the 'parent' dipole. Ineffectiveness of the counterpoise can make this kind of antenna system very inefficient.

Of course, nowadays extremely compact antennas are commonplace in mobile telephone handsets, partly by virtue of the short wavelengths involved. It seems to be generally held that such antennas have gains considerably less than an isotropic source, and there is some evidence supporting figures like -10 dBi. The physical length of the antenna is a small faction of λ (e.g. $\lambda/8$ = 42 mm at 900 MHz) so this is unlikely to be on account of lobes in the radiation pattern pointing the wrong way, although there may be scope for nulls in one or two directions on account of symmetry as happens in the case of an infinitesimal doublet, as noted in Section 8. Apart from the short length, the principal cause is probably what I mentioned in Section 14, and follows from Section 10: that when a conductor intended to work as an antenna is located close to other conductors, strong coupling between them can favour difference-mode (transmission line) currents making the effective radiation resistance very small, leading to low radiation efficiency.

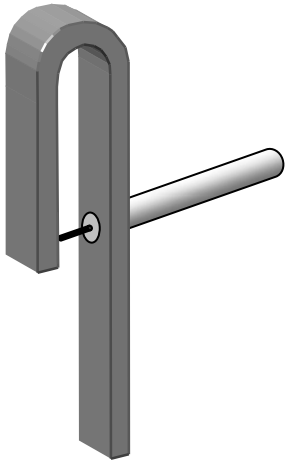
In some designs of VHF/UHF ground-plane antenna, the radials are sloped downwards [68]. This has the effect of increasing the terminal resistance and reducing the angle of elevation of the VRP lobe. A particularly interesting example is the AS-390/SRC military UHF communications antenna [69], [70], illustrated to the left, which also has the down-lead connected part way up the monopole via a short-circuit stub.



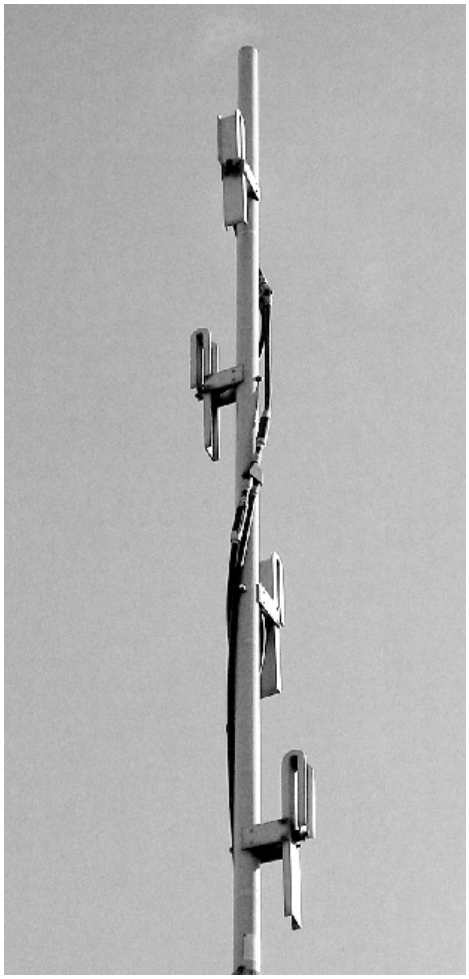
The ultimate case of the radials tilted downwards, by 90°, is effectively reversion to the 'parent' dipole. Sloping radials can also be considered a skeletal representation of a cone and really this type of antenna is a hybrid: one half conventional dipole and one half biconical dipole. I will discuss dipoles with conical elements in just a moment, in Section 19, but first I'd like to mention another interesting hybrid that can be seen above the roofs of countless tall buildings around the UK.

18. The 'J-pole'

The 400 to 470 MHz band immediately below Band IV has been used for PMR (Private Mobile Radio) since the 1960s and the development of miniature personal two-way radios like the original Pye 'Pocketfone' outfit (separate transmitter and receiver) overcame the need for police telephone kiosks like the 'Tardis' and smaller street pillars. According to Team Simoco [71] the Pocketfone was launched in 1964 and 70,000 were in use by 1975, 1000 of them at London Airport – clearly a case of providing a facility that had not existed before, which we now take for granted. The standard companion base-station installation included an antenna system consisting of four dipole elements arranged as a vertical collinear array [72] as shown below. In this case the elements were the *Centaur* of the dipole fraternity, each having one limb of a folded-dipole and the other a simple, non-folded limb, as illustrated to the left. This was fed directly by a co-axial cable, without apparent need for a balun; the outer co-axial conductor was connected to the inner ends of both limbs (actually a continuous metal strip, and mechanical support was provided here) and the inner conductor was connected to the other end of the folded limb. To a first order, one might surmise this arrangement presents the terminal impedance of a folded monopole (i.e. half a folded dipole), 146 Ω, to the cable; the non-folded limb acting as a $\lambda/4$ counterpoise without affecting the impedance. If the four cables feeding these elements were each cut to an integral number of half-wavelengths (paying attention to the cable's velocity factor, but not to its characteristic impedance), when connected in parallel they would present a quarter of this impedance, 36.5 Ω, which would provide a fairly good match to a 50 Ω down-lead: a return loss of about 16 dB. Making the dipoles marginally longer would increase the terminal resistance a little, at the expense of introducing some inductive reactance.



An example of such a collinear array can be seen above the roof of one of our buildings at Kingswood Warren (some of the facilities staff used to carry UHF PMR walkie-talkies) and this is the one pictured to the right. On closer inspection, as well as 'T-pieces' there are larger diameter sections in the cables connecting the elements together; $\lambda/4$ transformers perhaps. The rotation of successive elements around the supporting pole suggests an effort to overcome unwanted interaction between the elements and the pole distorting their radiation patterns. I understand this array was offered with a variety of different radiation patterns achieved by different element orientations (see [73]). Perhaps a case of making a virtue out of necessity?



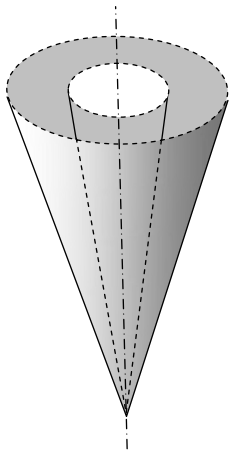
The practical implementations of this antenna appear always to have used a rectangular 'U'-channel material rather than the thick strip in my drawing above; perhaps a means of achieving "fatness" and rigidity without excessive weight?

The fact that more-recent PMR installations, like the one shown at the beginning of Section 14, use conventional 'full' folded dipoles suggests this design had some shortcomings. Modern TETRA [74] base stations necessarily involve potential simultaneous use of many different frequencies and, I understand from the manufacturer [72], it is generation of intermodulation products that has rendered obsolete this bolted-together design. The modern, full, folded-dipole equivalent has all-welded construction.

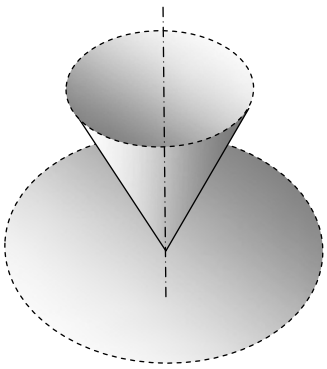
Nevertheless, I doff my hat to those that designed, installed and have maintained these antenna systems over the past 41 years.

19. Dipoles with conical elements

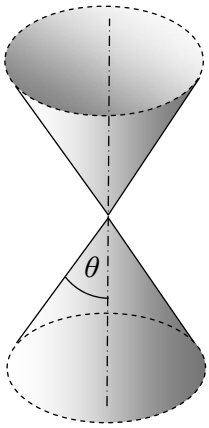
Resonant dipoles are fine for moderate-bandwidth applications, like the 21% fractional bandwidth of Band II, but resonant devices are not well suited to more extreme requirements approaching an octave (100% fractional bandwidth); they need a different approach based on fundamentally-non-resonant types of antenna. I mentioned in Section 11 that a uniform, constant-impedance, transmission line can provide the simplest solution to a broadband match, so if such a line can be designed in which the conductors are not very well coupled to each other it will also operate as an antenna. One of the simplest implementations of this is the biconical dipole.



For a co-axial line the characteristic impedance is given by $Z_0 = (60/\sqrt{\varepsilon_r})\log_e(b/a)$ where a and b are the diameters of the inner and outer conductors, respectively, and ε_r is the dielectric constant of the medium separating them [75]. If the absolute diameters of both conductors are increased progressively along the length of the line in a linear fashion, but the ratio of diameters is held constant, the impedance will be uniform. This leads to the notion of a co-axial pair of conical conductors, of infinite length, behaving as a constant-impedance line. At any distance from its apex, the diameter of the circular cross-section of each cone is proportional to the tangent of its apex angle. To the right, I've drawn a truncated section of this near the common apex. The terminals of this 'near end' of the transmission line would be at the almost-coincident apices of the cones; the other end of the line would be an infinite distance above.

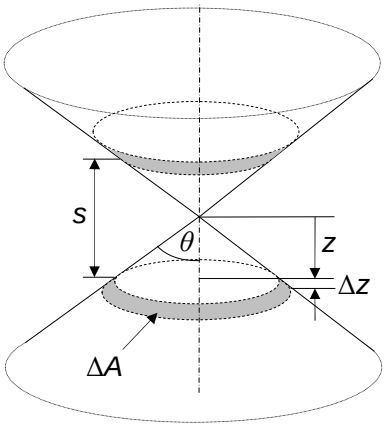


A bit of imaginative manipulation of the two apex angles leads to the idea of an infinite 'discone' where the apex angle of the outer cone has been increased to 180° making this conductor an infinite disc. I've drawn a truncated section of this to the left. Of course, the 'ratio of diameters' becomes meaningless in this case.



Taking the process even further can yield a symmetrical biconical transmission line, part of which I've drawn to the right. Again, the terminals of the 'near end' of this line would be at the near-coincident apices of the cones. With symmetrical conical limbs of infinite extent, this transmission line has uniform characteristic impedance that depends, geometrically, only on the apex half-angle θ. The principal mode of propagation is a spherical wave in the maximised space between the cones, some of the power in which is radiated – hence the biconical dipole antenna [76]. Progressively increasing the diameters in a non-linear fashion, but keeping their ratio constant, would yield exotic shapes like [77] Kraus's 'volcano smoke' antenna!

The characteristic impedance of an infinite biconical dipole, $Z = 120 \log_e(\cot\theta/2)$, which is real, is also its radiation resistance [78] and $\theta = 67°$ to make $Z = 50\ \Omega$.



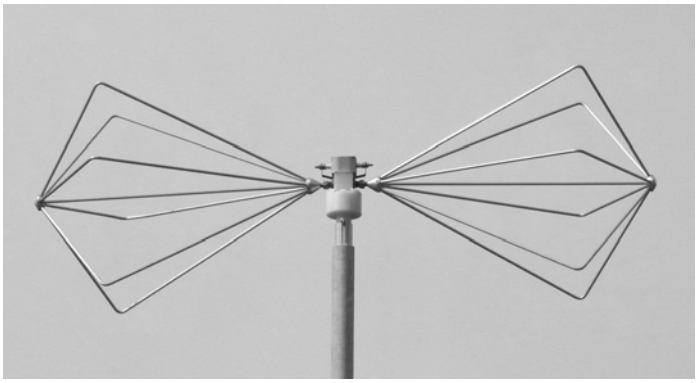
Uniform characteristic impedance implies the L/C ratio must not change along the length of the line and to a rough approximation it is easily shown that the capacitance per unit length is constant. As noted in Section 11, $C = A/s$ for two parallel plates, where A is the area of the plates and s their separation in air. The incremental area ΔA of an annular strip on the surface of a cone is proportional to the radius from the axis of the cone which depends on the axial distance z from the apex: $\Delta A = 2\pi z \tan\theta \Delta z$. The separation of two such strips, on the two cones symmetrical about the common apex, $s = 2z$ so $\Delta C = \Delta A/2z = \pi\tan\theta\Delta z$ which is constant per unit length (i.e. for constant Δz). This is not accurate because the 'plates' are not parallel, though they would be nearly so for large θ, but it demonstrates the point.

A more accurate formulation, derived from [79], yields $\Delta C/\Delta z = 2\pi/[\log_e \cot^2(\theta/2)]$.

However, treating the incremental surfaces of the cones as cylindrical conductors and applying the Rosa/Terman formula I mentioned in Section 11 would imply progressively increasing inductance per unit length with radius from the axis which would lead to a paradox because then the characteristic impedance couldn't be uniform – but Schelkunoff [79] and Kraus [76], to name but two, demonstrate fairly rigorously by other means that it is! Evidently, treating parts of cones as cylinders in this case is stretching the approximation too far.

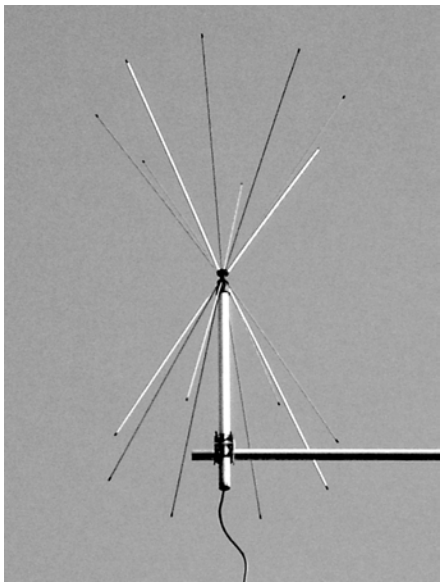
19.1 Biconical dipoles

There is no limit to the bandwidth of an infinite constant-impedance transmission line but any practical implementation of the biconical dipole has limbs of limited extent forming an open-circuit stub in the same manner as a resonant dipole. In the transmitting case, radiation from the biconical transmission line makes it lossy so the wave reflected by the open-circuit end is somewhat attenuated and if the conical surfaces were long enough their far ends would be rendered electrically 'invisible' at the terminals. In practice, the surfaces are usually made at least $\lambda/4$ long at the lower band edge so, at worst, the antenna behaves as a rather "fat" $\lambda/2$ resonant dipole. At higher frequencies its behaviour tends towards that of a true biconical transmission line, the upper limit depending principally on the accuracy of the practical implementation of the 'near-coincident apices'. Between these two extremes an acceptable return loss may be achieved over an octave, or more, depending on what constitutes 'acceptable' for the intended application (e.g. 10 dB). Notwithstanding these limitations, this is still one of the simplest truly 'wideband' antennas.



We have a cage biconical dipole in Spectrum Planning Group, shown to the left, built in-house for early EMC measurements. The terminals of the dipole limbs are connected to a balun (the small box in the middle) by rather long copper straps and its performance is not as good as would be hoped. The converging cones at the ends of the limbs may be a convenient way to prevent this item of laboratory equipment potentially "poking one's eye out" rather than having significant electrical properties.

A skeletal biconical dipole, shown to the right, can be seen above the roof of another of the buildings at Kingswood Warren and was installed for monitoring FM and DAB transmissions. This is operated as a 'sleeve dipole'; the supporting pole and down-lead passing through the centre of the lower conical limb form a co-axial transmission line having a large characteristic impedance everywhere except near the apex. As recalled earlier, the impedance of a co-axial line depends on the ratio of the diameters of the two conductors and, in principle, the much-greater impedance of this part of the system chokes currents from flowing on the pole or the outside of the down-lead screen, avoiding the need for a balun.



Variants of the biconical dipole have been employed over the years in several designs of television transmitting antenna, some using half a biconical dipole (i.e. the result of bisecting it through its axis). Other designs have used the 'bow-tie' element – a dipole made from a pair of triangular conductors [80] which is effectively a 2-dimensional projection of a biconical dipole. A further variation on this theme is the 'bat-wing' element [81] which combines a slot radiator with a flat conductor in the shape of two triangles (apex to apex).

19.2 Discones

In the same way that the symmetrical $\lambda/2$ dipole is related to the $\lambda/4$ 'ground plane' monopole antenna, so the practical biconical dipole is related to the 'discone' antenna, popular with VHF/UHF scanner enthusiasts (very/ultra short-wave listeners?). This is probably the best known form of wideband antenna and a skeletal example can be seen above the roof of the offices of Spectrum Planning Group at Kingswood Warren, as shown to the left. We installed this discone to monitor DAB transmissions, and anything else vertically-polarised we might be working on at VHF/UHF.

Again, the length of the skeletal conical surface is $\lambda/4$ at the lowest operating frequency but, evidently in this case, the 'large, thin, perfectly-conducting sheet' has been greatly truncated! The rods in the skeletal disc were probably shortened at the design stage until the return loss became unacceptable slightly below the lowest intended operating frequency.

The characteristic impedance (or radiation resistance) of an infinite discone is half that of the 'parent' biconical dipole, $Z = 60 \log_e(\cot \theta/2)$ and $\theta = 47°$ to make $Z = 50\ \Omega$.

We also have a collection of UHF discone antennas that were originally designed and built in-house for vertically-polarised television transmitter site tests using a helium-filled balloon. Although requirements for such site tests are uncommon nowadays, and mechanical 'aerial' platforms such as Simon hoists can be hired in, these antennas have found use when we have needed a UHF dipole that can easily be mounted co-axially on top of a mast.

Oddly for a wideband type of antenna, a family of three different sizes was built for frequencies of 522, 666 and 756 MHz (mid-Channel 27, mid-Channel 45 and three-quarters of the way up Channel 56!); representative of three of the widely-used television receiving 'aerial groups' [57]. According to the Research Department Report about these antennas [82] 'radiation from the supporting tube was found to cause excessive irregularity of the vertical radiation pattern' so a co-axial sleeve choke was added to the metallic supporting tube through which the feeder passes (hidden within a larger-diameter glass-fibre tube in each case). Unfortunately, the resonant nature of this choke defeated the wideband nature of the antenna! So why persevere with the, relatively-complicated, discone? Could it be significant that the centre of the disc presents a relatively low impedance point to which a string can be anchored to suspend this antenna from the balloon … especially when the string is wet?

Even more oddly, in its intended original application the 'small transmitter' was connected directly to the C-type plug at the end of the 'supporting tube' so one might imagine there would have been scope for doing away with this tube altogether and minimising the length of feeder that could cause spurious radiation. In the limit, by making the discone large enough, the entire transmitter could have been built within the cone thereby exposing no feeder at all!

Perhaps this mystery is related in some way to range measurements of the VRP (with the antenna mounted on its side, of course) using a much-longer feeder.

20. Conclusions

I have presented here my 'take' on the topic of dipole antennas: a bundle of facts and inferences that, presented together in this order, may help to explain the bases for the kinds of link budget often encountered in spectrum planning work.

In 'PART 1: characteristics of isotropic sources and $\lambda/2$ dipole antennas', I've drawn heavily on the work of J. D. Kraus, concentrating on how we can account for these characteristics in link budgets for communication and broadcasting systems. Along the way, I've noted ways in which some of these characteristics, like effective aperture area and radiation resistance, can be calculated for simple antenna designs and hypothetical references. A glance through the references listed below demonstrates the point I made at the outset, that a logically-ordered explanation for this purpose can take a highly-convoluted path through one of the standard text books on antennas (which, for its own intended purpose, is ordered entirely logically).

In 'PART 2: electrical engineering of dipole antennas' I have delved a little deeper into the basis for the operation of several practical implementations of the dipole antenna, explaining the significance of the $\lambda/2$ length and why we most often see folded dipoles nowadays. I have touched on baluns and cage dipoles, and I've mentioned co-axial dipoles and antennas that use conical elements, as well as a couple of time-honoured, unusual designs and some BBC in-house specials.

I originally embarked on this endeavour in response to questions from some colleagues which, I'm told, I had answered when this document in an early draft had only four or five pages! I was motivated to keep going because of the old adage "the more you think you know, the more you realise you don't" which, in my experience, seems to apply to antennas more than any other topic in electronic engineering. I realise I've no more than scratched the surface here but I hope this White Paper proves useful to one or two.

I am indebted to all those colleagues who have assisted me by proof reading, in particular Dave Darlington, John Green, Peter Shelswell, Ranulph Poole, and Jonathan Stott who also helped me with some of the arithmetic. I also wish to thank John Barrett and Kevin Claydon for some of the photographs.

21. References

1. A short history of BBC R&D,
   available in March 2006 from www.bbc.co.uk/rd/about/history/history.shtml
2. GANDY, C. An introduction to the Eureka DAB System and a guide to how it works.
   BBC R&D White Paper WHP 061, June 2003 (re-printed from 1994), available in March 2006 from
   www.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP061.pdf
3. KRAUS, J.D. Antennas. Second edition. McGraw-Hill, Inc., New York, 1988, ISBN 0-07-035422-7
4. *Ibid.* Page 517
5. *Ibid.* Pages 175 to 179
6. *Ibid.* Page 36, Footnote 1
7. *Ibid.* Page 44
8. *Ibid.* Page 219
9. *Ibid.* Page 30
10. *Ibid.* Pages 224 to 227
11. *Ibid.* Page 206
12. *Ibid.* Page 215
13. KRAUS, J.D. Electromagnetics. Fourth edition. McGraw-Hill, Inc., New York, 1992,
    ISBN 0-07-035621-1. Page 238.
    Note that Kraus refers to this expression as 'Ampère's law' in his book 'Antennas', Page 211, but it is a development of the Biot-Savart law that leads to Ampère's (circuital) law by integration

14. KRAUS. Antennas. *Ibid.* Pages 224 to 227

15. *Ibid.* Pages 413 to 422

16. JORDAN, E.C. and BALMAIN, K.G. Electromagnetic waves and radiating systems. Second edition. Prentice-Hall, Inc., New Jersey, 1968. Pages 540 to 545

17. *Ibid.* Page 548

18. KRAUS. Electromagnetics. *Ibid.* Page 757

19. BURKE, G.J. and POGGIO, A.J. Numerical Electromagnetics Code (NEC) - Method of Moments. Lawrence Livermore National Laboratory Report UCID-18834, 1981, available in March 2006 (with the wrong cover page) from www.nec2.org/other/nec2prt1.pdf

20. KRAUS. Antennas. *Ibid.* Page 394

21. Information about a commercial software tool named 'FEKO' that uses the method of moments together with other EM modelling techniques, available in March 2006 from www.feko.info/solutions.html

22. SCHELKUNOFF, S.A. and FRIIS, H.T. Antennas, theory and practice. John Wiley & Sons, Inc., New York, 1952. Page 337

23. KRAUS. Antennas. *Ibid.* Page 32

24. *Ibid.* Page 34

25. *Ibid.* Page 38

26. *Ibid.* Page 39

27. *Ibid.* Page 40

28. *Ibid.* Page 49

29. *Ibid.* Page 48

30. *Ibid.* Page 47

31. GANDY, C. Co-axial cables. BBC R&D White Paper WHP 126, September 2005, available in March 2006 from www.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP126.pdf

32. JORDAN and BALMAIN. Electromagnetic waves and radiating systems. *Ibid.* Pages 388 to 393

33. HALLÉN, E. Theoretical investigations into the transmitting and receiving qualities of antennae. *Nova Acta Regiae Soc. Sci. Upsaliensis*, Series IV, Vol. 11, No. 4, 1938. Pages 1 to 44

34. BROWN, G.H. and WOODWARD, O.M.Jr. Experimentally determined impedance characteristics of cylindrical antennas. *Proc. IRE*, Vol. 33, April 1945. Pages 257 to 262

35. JORDAN and BALMAIN. Electromagnetic waves and radiating systems. *Ibid.* Pages 566 and 567

36. KRAUS. Antennas. *Ibid.* Pages 360 to 376

37. *Ibid.* Page 713

38. ROSA, E.B. The self and mutual inductances of linear conductors. *Bulletin of the Bureau of Standards*, Vol. 4, No. 2, 1908. Pages 301 to 305

39. TERMAN, F.E. Radio engineers' handbook. First edition. McGraw-Hill, Inc., New York, 1943. Pages 48 and 49

40. KRAUS. Electromagnetics. *Ibid.* Pages 235 to 240

41. *Ibid.* Page 503

42. KRAUS. Antennas. *Ibid.* Page 220

43. SCHELKUNOFF and FRIIS. Antennas, theory and practice. *Ibid.* Pages 110 and 111 (wherein reference is made to another of Schelkunoff's books: Electromagnetic waves. Van Nostrand, New York, 1943. Pages 292 and 293)

44. Information about the history of Alexandra Palace,
available in March 2006 from www.transdiffusion.org/emc/baird/ally_pally

45. WILLIAMS, H.P. Antenna theory and design: Volume 2. Pitman & Sons Ltd., London, 1950. Pages 126 to 127

46. BROWN, A.H. and STANIER, B.A. Recent developments on VHF ground-communication aerials for short distances. *JIEE*, Vol. 94, Part IIIA, No. 14, 1947. Pages 637 and 638

47. KRAUS. Antennas. *Ibid.* Pages 511 and 512

48. PAGE, H. An introduction to the theory of aerials. BBC Research Department Report BBC RD 1963/27 (E-085). Pages 67 to 69.
Available in March 2006 from www.bbc.co.uk/rd/pubs/reports/1964-27.pdf

49. JASIK, H. Antenna engineering handbook. First edition. McGraw-Hill, Inc., New York, 1961. Pages 3-13 and 3-14

50. David Jefferies' article on 'Dipole impedances', available in March 2006 from www.ee.surrey.ac.uk/Personal/D.Jefferies/rodres.html

51. KRAUS. Electromagnetics. *Ibid.* Pages 522 to 532

52. PAGE. An introduction to the theory of aerials. *Ibid.* Page 69

53. KRAUS. Antennas. *Ibid.* Pages 739 to 741

54. KRAUS. Electromagnetics. *Ibid.* Page 499

55. PAGE. An introduction to the theory of aerials. *Ibid.* Page 73

56. KING, G.J. The practical aerial handbook. Second edition. Newnes-Butterworth, London, 1970 (reprinted in 1976). Pages 58 and 59

57. BBC Reception Advice fact sheet on 'TV aerials and reception', available in March 2006 from www.bbc.co.uk/reception/factsheets/pdfs/tv_aerials_factsheet.pdf

58. Information on the magnetic properties of ferrite materials from the manufacturer of 'Fair-rite' products, available in March 2006 from
www.elnamagnetics.com/library/catalogs/fairrite/14th_rev2/Properties_of_Materials.pdf

59. JASIK. Antenna engineering handbook. *Ibid.* Page 31-23

60. PAGE. An introduction to the theory of aerials. *Ibid.* Page 69

61. KRAUS. Antennas. *Ibid.* Pages 137 to 141

62. *Ibid.* Pages 218 to 221

63. *Ibid.* Page 462

64. Description of the Zeppelin end-fed $\lambda/2$ antenna
(also known as a 'J antenna' or a 'J-pole' antenna) available in March 2006 from
www.radio-electronics.com/info/antennas/j_antenna/j_pole_antenna.php

65. KING. The practical aerial handbook. *Ibid.* Pages 53 and 54

66. JOHNSON, R.C. and JASIK, H. Antenna applications reference guide. Second edition. McGraw-Hill Inc., New York, 1984, ISBN 0-07-032284-8. Pages 2-8 and 2-9

67. KRAUS. Antennas. *Ibid.* Pages 334 to 338

68. JOHNSON and JASIK. *Ibid.* Page 4-13

69. Description of the US military AS-390/SRC antenna, available in March 2006 from
www.combatindex.com/hardware/detail/sensor/comms/as-390.html

70. United States Navy. Electronics Technician nonresident training course. Volume 7 – antennas and wave propagation, NAVEDTRA 14092. Pages 2-16 to 2-17, Fig. 2-28 View B, available in March 2006 from www.globalsecurity.org/military/library/policy/navy/nrtc/14092.pdf

71. Contact details for Team Simoco, the modern evolution of the radiotelephone parts of Pye, available in March 2006 from www.teamsimoco.com/support/technical.html

72. Contact details for Sigma Wireless, the modern evolution of the company that supplied the collinear array antennas to Pye, available in March 2006 from www.sigmawireless.ie/default.htm. Also see www.sigmawireless.ie/documents/Antenna%20System%20DesignV5.pdf, their own Technical White Paper on 'Antenna system design (400 MHz) – Version 5'. The antenna appears in diagrams on Pages 23, 25, 26 and 28

73. JOHNSON and JASIK. Antenna applications reference guide. *Ibid.* Page 4-15

74. TETRA World Congress 2005, an Introduction to TETRA presented by members of The TETRA MoU Association, available in March 2006 from www.tetramou.com/tetramou.aspx?id=2671

75. KRAUS. Electromagnetics. *Ibid.* Page 498

76. KRAUS. Antennas. *Ibid.* Pages 340 to 358

77. *Ibid.* Pages 693 and 694

78. WOLFF, E.A. Antenna analysis. Artech House, 1988, ISBN 0-89006-286-2. Page 95

79. SCHELKUNOFF and FRIIS. Antennas, theory and practice. *Ibid.* Pages 104 to 106

80. JOHNSON and JASIK. Antenna applications reference guide. *Ibid.* Pages 6-12 to 6-13.

81. *Ibid.* Page 5-14

82. HARTLEY, W. Vertically-polarized light-weight U.H.F aerials for site testing. BBC Research Department Report BBC RD 1969/32 (RA-44)

83. KITCHEN, F.A. The design of an omni-directional aerial system for the frequency range 225-400 Mc/s. *Proc. IEE*, Vol. 98, Part III (radio and communication engineering), No. 56, November 1951. Pages 409 to 415

84. Information about one type of aeronautical communication transmitter, available in March 2006 from www.rockwellcollins.com/ecat/gs/Communication_Systems.html?smenu=101 (see 'Transmitter Specifications' two-thirds of the way down the page)

85. Contact details for Racal Antennas, available in March 2006 from www.racalantennas.com/5-2-ContactInfo.htm

86. Information on Qinetiq sites, available in March 2006 from www.qinetiq.com/home/aboutqq/useful_links/locations.html

All the text books and most of the journal articles listed here were found in the BBC R&D Library.