A mathematical approximation in the physical sciences

JOHN D. MAHONY

Introduction

The business of making mathematical approximations in the physical sciences has a long and noble history. For example, in the earliest days of pyramid construction in ancient Egypt it was necessary to approximate lengths required in construction, especially when they involved irrational numbers. Similarly, surveyors in early Greece seeking to lay out profiles of right-angle triangles or circles on the ground invariably ended up making approximations regarding measurements of required lengths, as indeed is the case today. Practitioners have always faced the problem of having to decide when parameters in theory have been met satisfactorily in the practice of measurement. Further, before the advent of hand-held calculators, students in schools in the UK would have been very familiar with the approximation 22/7 for the transcendental number π, obtained perhaps by comparing (as this author did) the measured circumferences of many laboriously drawn circles of different sizes with their diameters. Despite the advent of sophisticated calculating devices and facilities, such as computers and spreadsheets, the practice of making approximations is still much in evidence in theoretical work in fields associated with physical phenomena. Such approximations often result in formulae that are easy to use and remember, and moreover can produce theoretical results that support directly, or otherwise, results from measurements. In this respect, the practical mathematician does not have to seek results to many decimal places when measurement facilities allow for accuracy to only a few. The purpose of this Article is to illustrate this point by discussing an example drawn from the realms of antenna theory, relating to the performance of a dipole antenna. It is not the purpose here to delve into the derivation of dipole theory, but to extract the relevant information and show how useful mathematical approximations can be employed to simplify a relationship between parameters of interest to an antenna engineer. To this end, it will first be necessary to introduce some antenna concepts that might be new to the reader.

Some antenna concepts

In this discussion an antenna is assumed to be located at the origin of a rectangular cartesian coordinate system XYZ. The usual polar angles (θ, ϕ) related to this coordinate system are understood in what follows. It is also understood that a far field ansatz applies, that is to say, the values of all quantities are those that pertain in the far field of the antenna so that the antenna, despite its obvious physical dimensions, can be viewed as a point source located at an origin. The polar angle θ is measured from the positive

https://doi.org/10.1017/mag.2022.62 Published online by Cambridge University Press

z-axis; the azimuth angle ϕ is measured counter-clockwise from the positive x-axis.

(I) *Radiation intensity*: This is the power per unit solid angle (watts/unit solid angle) radiated by the antenna. Denote it by $U(\theta, \phi)$. The element of solid angle is $d\omega = \sin\theta \, d\theta \, d\phi$.

(II) *Radiation pattern*: This is a spatial distribution of the antenna's power that is radiated to the far field. It is a measurable quantity that typically varies with angular direction and is one that, for a given antenna, can be described by a mathematical function involving the polar angles. Typically an antenna is designed to radiate maximum power in a specified direction, i.e. the 'main beam' direction, with lesser, unwanted amounts unavoidably radiated in other directions as 'side lobes'. It is customary to normalise power in all directions to that of the maximum. Hence the normalised power ratio in the maximum direction is unity.

(III) *Total radiated power*: This is obtained by integrating the radiation intensity over the entire solid angle of 4π. Accordingly, the total power radiated by the antenna, P say, is then given by

$$P = \int U \, d\omega = \int \int_{\phi = 0, \theta = 0}^{\phi = 2\pi, \theta = \pi} U \sin\theta \, d\theta \, d\phi.$$

If U is both independent of ϕ and symmetric about $\theta = \pi/2$, it follows that the total radiated power $P = 4\pi / \left(\int_0^{\pi/2} U \sin\theta \, d\theta\right)$ (power dimensions are in watts).

(IV) *Directivity*: This is the ratio of the maximum radiation intensity to the radiation intensity averaged over all directions. The maximum radiation intensity occurs at the peak of the radiating antenna's main beam, i.e. in the target direction, say. Denote this ratio by D. The radiation intensity averaged over all directions is equal to the total power radiated by the antenna divided by 4π. Then the directivity $D = 4\pi U_{\max} / P$. If now U is normalised by its maximum value to U_N, where $U_N = U / U_{\max}$, it follows from (III) above that the directivity can be expressed in the form $D = 1 / \left(\int_0^{\pi/2} U_N \sin\theta \, d\theta\right)$. Since the directivity is a power ratio it is expressed commonly in terms of decibels, as $10 \operatorname{Log}_{10}(D)$ dB. More information about directivity etc. is contained in the Appendix.

(V) *Half power beam-width*: This is the angle between two directions in a plane containing the beam maximum, where the radiation intensity is one half of its maximum value. It is an important concept in target resolution. For example, the capability to distinguish between two sources is generally related to the antenna's half power beam-width. Moreover, if an antenna is transmitting to receivers (also antennas) that are outside this angle of interest then those receivers will secure a relatively weaker signal compared with one received by any such antenna within the half power beam-width angle.

Both directivity and half power beam-width are yardsticks used to describe and quantify performance aspects of an antenna. Engineers are usually very interested to see an easy and direct connection between the two because it can simplify the understanding of what could be an otherwise complex relationship.

Having established a framework, attention will now be turned to the exact mathematics related to the dipole, after which an approximate formulation will be considered. In the process of developing an approximate description, the dependency of directivity on half power beam width angle will become apparent.

Exact results

The requisite results from the theory relating to a dipole antenna will simply be cited in what follows. A method of obtaining the results, which involves an appreciation of the vector potential for an electric source current solution to Maxwell's equations is described in [1]. It is assumed in this reference that the current distribution across the dipole is sinusoidal, but this information is for completeness only and should not bother the reader. A schematic representation for such a dipole, coaxially fed, is as shown in Figure 1. With reference to this figure, the operating wavelength is denoted by λ and each arm of the dipole is of length $L/2$, giving an overall length of L. Energy is supplied to the dipole via a coaxial line connected to some energy source. One arm of the dipole is formed by extending the innerconductor (wire) of the coaxial cable and bending it upwards through a

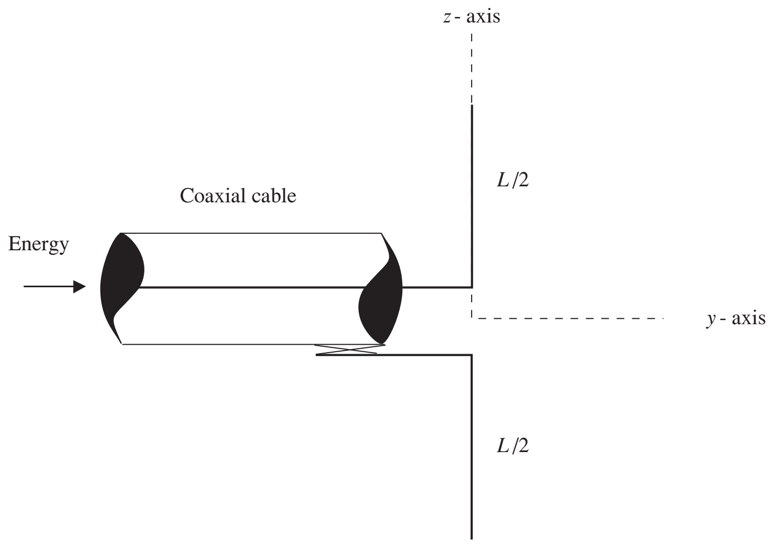

FIGURE 1: A schematic representation of a coaxially fed dipole of length L located in a cartesian reference frame

right angle, as shown in the figure. The other arm of the dipole is formed by soldering an identical wire to the outer conductor of the coaxial cable and bending it similarly downwards, as shown. Both arms are aligned to lie along the z-axis. When the source is activated, the dipole radiates energy. This method of constructing a dipole is not unique, but it does require additional physical modifications to ensure effective radiation.

A formula for the radiation intensity of the dipole can be found in [1]. When normalised to its maximum value, i.e. when $\theta = \pi/2$, it takes the form

$$U_N = \left(\frac{\cos(\alpha \cos\theta) - \cos\alpha}{(1 - \cos\alpha)\sin\theta}\right)^2, \qquad 0 \leqslant \phi \leqslant 2\pi \tag{1}$$

where $\alpha = \pi L/\lambda$ (λ denotes the operating wavelength). The quantity U_N is circularly symmetric about the z-axis, that is to say it is the same for any specified ϕ-value. This expression can be plotted on a polar diagram, for example when $\phi = 90°$ using the equations $y = U_N \sin\theta$ and $z = U_N \cos\theta$. The pattern in any other ϕ-plane cut will be identical. The result for the case, say, of the half wavelength dipole (that is when $\alpha = \pi/2$) is shown in Figure 2 below; the radius vector from the origin to any point on the curve is the value for U_N at that point.

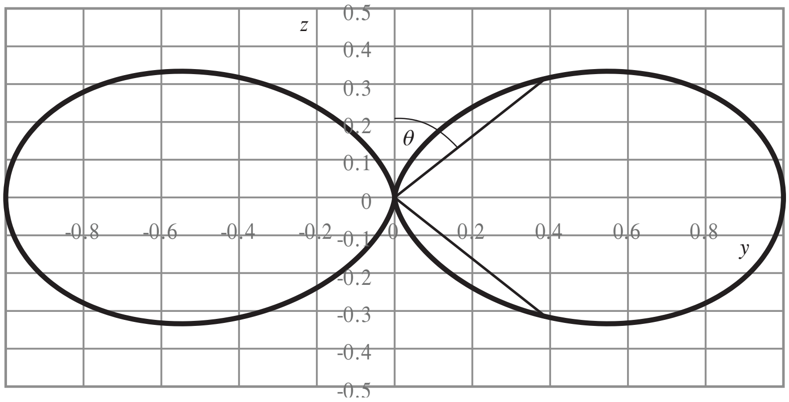

FIGURE 2: A polar plot (far field radiation pattern) of the normalised radiation intensity for a half wavelength dipole y-axis horizontal, z-axis vertical, x-axis towards the reader

A distant observer in the far field can treat the dipole as a point source with the above radiation pattern characteristics. This aspect is elaborated further in the appendix. It can be seen from the figure that the maximum value for U_N is unity (because it is normalised) and it occurs when $\theta = \pi/2$. The minimum value is zero and occurs when $\theta = 0$. Thus, in three dimensions, the radiation pattern is much like that of a doughnut with no appreciable hole. Radiation patterns of this type are said to be omnidirectional and a more precise definition of this description can be found by googling the word ‘omnidirectional’. It is possible to consider dipoles of different lengths by varying the value of the input parameter α in the above. Typically, values chosen for the dipole length here will not exceed one wavelength, so that only cases for $\alpha \leqslant \pi$ will be addressed. The interested reader might care to obtain plots for such values and see just how the radiation pattern varies. Values for $\alpha > \pi$ imply a dipole length greater than unity. In such cases, the radiation patterns show the presence of minor lobes that increase in size with increasing dipole lengths. The interested reader might care also to appreciate this by obtaining plots, in the manner that led to Figure 2, when $\alpha > \pi$.

As mentioned earlier, the quantities of interest to the antenna engineer are the half power beam-width angle and the directivity. We will treat them separately below.

Half power beam-width

It is necessary to determine first the θ-value at which the normalised radiation intensity assumes the value $\frac{1}{2}$. Then, with reference to Figure 2, the angular spread between the θ-values where this occurs is $(\pi - 2\theta)$. This is the half power beam width (hpbw) angle, which we denote by θ_{hpbw}, and it is also a quantity of interest in other antenna types such as arrays and reflectors, but such configurations are beyond the scope of this Article. This angle is also shown in the figure, sandwiched between the two radius vectors (each of length $\frac{1}{2}$) in the lobe on the right-hand side of the figure. There will of course be a similar angle on the left-hand side of the figure, but it is not shown. Mathematically speaking, the requisite value for the angle θ is determined from the equation $U_N = \frac{1}{2}$, i.e.

$$\left(\frac{\cos(\alpha \cos\theta) - \cos\alpha}{(1 - \cos\alpha)\sin\theta}\right)^2 = \frac{1}{2}. \tag{2}$$

The interested reader might care to verify that in the case of the half wavelength dipole, $\theta = 50.961141°$ and the half power beam-width value is $\theta_{\text{hpbw}} = 78.078°$. It is a relatively straightforward matter to solve (2) for a variety of dipoles of different lengths (i.e. for different values of the parameter α, here not exceeding π). A Newton-Raphson method is particularly well suited to this exercise and typical results from it, including that for the half wavelength dipole, are shown in Table 1 below.

L/λ	α	θ°	$\theta^\circ{}_{\text{hpbw}}$
0.0625	0.196350	45.092083	89.815834
0.125	0.392699	45.368839	89.262323
0.25	0.785398	46.482264	87.035472
0.375	1.178097	48.351153	83.297693
0.5	1.570796	50.961141	78.077719
0.625	1.963495	54.232372	71.535256
0.75	2.356194	57.996368	64.007263
0.875	2.748894	62.022252	55.955495
1	3.141593	66.082468	47.835064

TABLE 1: Half power beam-width angles for dipoles of different lengths

Directivity

The equation for the directivity is given in (IV) above, in the form

$$D = 1/\left(\int_0^{\pi/2} U_N \sin\theta \, d\theta\right), \tag{3}$$

where U_N is given by (1). Methods have been employed to reduce the integration in (3) to a series of manageable closed form steps plus an integral that must be evaluated either numerically or in look-up tables [1]. Alternatively, it can be evaluated directly using a numerical integration procedure such as, say, Simpson's rule. In such fashions it is possible to obtain the following exact directivity results for a variety of dipole lengths, as per Table 2 below.

L/λ	D eq$^{\text{n}}$(3)	D(dB)
0.0625	1.501931	1.766501
0.125	1.507772	1.783357
0.25	1.531845	1.852148
0.375	1.574661	1.971871
0.5	1.640922	2.150880
0.625	1.738782	2.402452
0.75	1.882074	2.746368
0.875	2.094060	3.209891
1	2.410998	3.821968

TABLE 2: Exact directivities for dipoles of different lengths

https://doi.org/10.1017/mag.2022.62 Published online by Cambridge University Press

The above two tables contain data showing how the directivity and the half power beam-width angles vary with different dipole lengths and one could construct a plot to display the relationship. However, it is not immediately apparent that there is a simple, explicit relationship between the directivity and the dipole's half power beamwidth angle. Fortunately, by employing suitable approximations, it is possible to extract one, as described below.

Approximate results

An approximation to the above expression for the normalised radiation intensity U_N (see (1)) can be considered in the form

$$U_N \cong \sin^{\mu}(\theta), \qquad 0 \leqslant \phi \leqslant 2\pi \tag{4}$$

where the value chosen for the parameter μ is such as to provide hopefully a good approximation to the function U_N described by (1) yielding, also hopefully, a close match to the radiation pattern of Figure 1 and a subsequent simplification in the steps required to evaluate the directivity integral of (3). To this end, the approach is first to force agreement between equations (4) and (1) at the half power beam width angle value for the dipole, of whatever length being discussed here. At such a point,

$$\mu = \frac{\ln(0.5)}{\ln\left(\cos\left(\theta_{\text{hpbw}}/2\right)\right)}. \tag{5}$$

Then, in the case of the dipole lengths shown in Table 1 it is possible to construct the following table showing values for the parameter μ that ensure a half power beam width angle match between the exact and approximate radiation intensity patterns.

L/λ	μ eq^n (5)
0.0625	2.009303
0.125	2.037605
0.25	2.156910
0.375	2.378844
0.5	2.743218
0.625	3.316473
0.75	4.204522
0.875	5.577167
1	7.720289

TABLE 3: Values for the parameter μ that ensure a match for half power beam-width angles listed in Table 1.

https://doi.org/10.1017/mag.2022.62 Published online by Cambridge University Press

As an illustration, in the case of the half wavelength dipole it is again a simple matter to plot the approximate radiation intensity pattern due to (4) and compare it with that of Figure 1. This is shown in Figure 3.

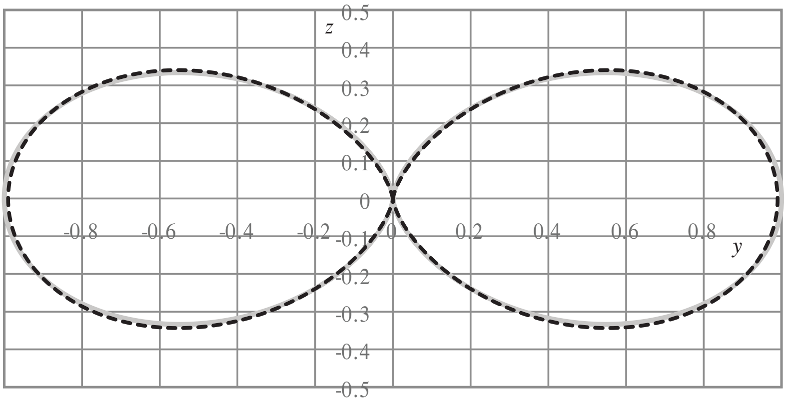

FIGURE 3: A comparison of polar plots for U_N in the case of the half wavelength dipole
exact – solid curve as in Figure 1
approximate – dashed curve when $\mu = 2.743218$

It may be appreciated from Figure 3 that the approximation with the proposed value for the parameter μ is not unreasonable, because with the naked eye it is difficult to discern the difference the exact and approximate curves. A value of $\mu = 3$ was chosen in [1] to approximate the half wavelength dipole radiation pattern of (4) and the reader might care to produce the ensuing radiation pattern to appreciate the difference between it and the one proposed here.

A similar story of encouraging plot comparisons between exact and approximate pattern representations emerges in each case of the other dipole lengths and associated μ-values cited in Table 3. It is left as an exercise for the reader to secure the associated plots to appreciate this.

Having approximated a given radiation pattern using an appropriate value for μ from (5), it remains to address the integration for the directivity, which now is given approximately by the equation

$$D \approx \frac{1}{\int_0^{\pi/2} \sin^{\mu+1}\theta \, d\theta}. \tag{6}$$

Despite appearances, the integral in the denominator can be fruitfully approximated. An approximation to it based on an asymptotic development was given first in [3] (with misprints), and later in [4]. Specifically, it was shown that such an integral could be approximated by

$$\sqrt{\frac{\pi}{2z + 1 + 0.25z^{-1}}},$$

where $z = \mu + 1$. If terms of $O(z^{-1})$ are neglected in this approximation, the leading terms in the asymptotic expansion can be retained to yield the

https://doi.org/10.1017/mag.2022.62 Published online by Cambridge University Press

simpler approximation $\sqrt{\pi(2\mu + 3)}$. Such expansions can be convergent or divergent, but the leading terms nonetheless can be of interest to the practitioner, producing results that can accord with measurement. Here, retaining the simpler terms gives

$$D \approx \sqrt{\frac{2\mu + 3}{\pi}}$$

which, for μ given by (5), may be rewritten in the form

$$D \approx \frac{1}{\sqrt{\pi}}\sqrt{\frac{2\ln(0.5)}{\ln\left(\cos\left(\frac{1}{2}\theta_{\text{hpbw}}\right)\right)} + 3}. \tag{7}$$

This is an explicit relationship between the directivity and the half power beam-width angle, albeit an approximate one.

It is appropriate now to compare values for the directivity from (7) with the exact values shown in Table 2, the associated half power beam-width angles being those displayed in Table 1. The comparison can be drawn from the results shown in Table 4 below.

L/λ	D(dB)eq$^{\text{n}}$(7)	δ(dB)
0.0625	1.745505	0.021
0.125	1.762947	0.020
0.25	1.834972	0.017
0.375	1.962912	0.009
0.5	2.157877	-0.007
0.625	2.433046	-0.031
0.75	2.800497	-0.054
0.875	3.268698	-0.059
1	3.843123	-0.021

TABLE 4: The approximate directivities from (7) and the differences between them and the exact values shown in Table 2

The differences in the Table are not large and should be weighed against the accuracies of associated gain (directivity less system losses) measurements on many field-test sites. Typically, at best these could be of the order of between about, say, 0.05 dB and 0.1 dB. On some sites it could be worse, depending on the sophistication of the measurement set up.

It is possible to carry the approximation a step further by employing in (7) the expansion for the "ln cos" function given in [2]. In a somewhat tedious but straightforward manner it can be shown, using a small argument approximation, that this reduces the equation to the following simpler,

https://doi.org/10.1017/mag.2022.62 Published online by Cambridge University Press

approximate form

$$D \approx \frac{2\sqrt{(4 \ln 2)/\pi}}{\theta_{\text{hpbw}}}\left\{1 + A\theta_{\text{hpbw}}^2 + O\left(\theta_{\text{hpbw}}^4\right)\right\} \tag{8}$$

where $$A = \frac{1}{16}\left(\frac{3}{2 \ln 2} - \frac{1}{3}\right) = 0.114419327.$$

In (8), the half power beam-width angle is assumed to be in radians, but it is customary to appreciate the angle in degrees because that is what is seen in radiation pattern plots on a field site. Thus, if the half power beam-width angle in degrees is denoted by say, ω, it will suffice to replace θ_{hpbw} in the above by $\omega\pi/180$, so that the equation above can be approximated in the form

$$D \approx \frac{107.651582}{\omega}\left\{1 + (\kappa \times 10^{-5})\,\omega^2\right\} \tag{9}$$

where $$\kappa = 3.485412.$$

Not surprisingly, results from (9) are not as accurate as those in Table 4. This is evidenced below in Table 5.

L/λ	D(dB)eq^n(9)	δ(dB)
0.0625	1.862723	-0.096
0.125	1.877843	-0.094
0.25	1.940801	-0.089
0.375	2.054515	-0.083
0.5	2.231662	-0.081
0.625	2.487778	-0.085
0.75	2.837594	-0.091
0.875	3.291597	-0.082
1	3.855985	-0.034

TABLE 5: The approximate directivities from (9) and the differences between them and the exact values shown in Table 2

Such differences notwithstanding, it is possible to improve this agreement by adjusting heuristically the value for κ in (9). The interested reader might care to do this by varying interactively on a spreadsheet the output from (9) as κ undergoes small changes. In this manner, the author found that it was possible to reduce the δ-dB error when compared with exact results to less than about 0.03 dB everywhere, with a value of $\kappa \approx 3.04$. Alternatively, it is conceivable that a best fit curve of the form of (9) with higher order terms could produce more accurate results. Other authors have also considered the problem of finding an explicit but

https://doi.org/10.1017/mag.2022.62 Published online by Cambridge University Press

approximate relationship between the directivity and the half power beam width angle in omnidirectional antennas and the results of some of their efforts can be found in [1, Chapter 2].

Conclusion

An example drawn from the realms of physics (antenna theory) has been used to demonstrate how approximations can be employed in a theory to produce expressions that facilitate an ease in calculation and comprehension related to physical quantities that are of interest to a practitioner and can be related to measurable quantities. Despite the reliance nowadays on immediately available software packages to obtain exact results from theoretical formulations (that may themselves have limitations regarding the accurate representation of physical phenomena) the use of approximations should not be discouraged because they can allow for an immediate appreciation and interpretation of physical processes. Moreover, they can yield fruitful, back-of-an-envelope results, results that can be more or less in accordance with what might be expected either from measurement or theory. Whilst the theory associated with antennas of this type will lie beyond what is usually encountered at undergraduate levels, the associated mathematics and its manipulations are most certainly not. The above has involved only integrations (numerical and approximate) and other simple approximations, and all of the ensuing processes have involved functions no more difficult to comprehend than the commonly understood terms of simple trigonometry.

Appendix

Further radiation pattern concepts

Detailed information concerning the theory and results pertaining to a radiating dipole supporting a sinusoidal current distribution on its arms can be found in [1], and in other related texts. It can be found also on websites by googling the appropriate words. A typical schematic representation for such a dipole was shown in Figure 1. Establishing parameters of interest from theory is beyond the scope of this note, and the reader must accept that they are as described here and in the literature. Broadly speaking, if the dipole is being used to transmit a radio frequency (rf) signal, a current is induced in the arms of the dipole when the input power source is activated. This current radiates an electromagnetic field according to a set of principles, Maxwell's equations. Typically, these equations can be met by students at an advanced undergraduate level and at a postgraduate level when set within the framework of antenna theory. The concept of the radiation pattern due to a dipole can be appreciated from the schematic of Figure 4. The dipole is located physically at the centre of the figure and the radiation pattern due to it (see Figure 1) is shown also superimposed. This is a plane figure that, in the case of the dipole, is typical of all other ϕ-plane cuts. In addition, an observation point on the arc of a large circle of radius R

(the far field of the antenna) is shown with a radius vector connecting it to the central point.

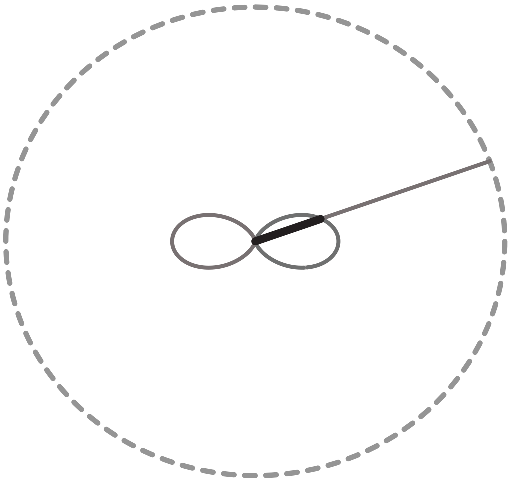

FIGURE 4: Dipole pattern and far field observer at distance r from the origin (axes as described in Figure 1 of the main text)

When the dipole is radiating, the observer on this far field circle can point a receiving device (typically a horn) at it to record a received power level (a signal strength) that is proportional to the length of the heavy black line. This signal strength can be compared (via a ratio) with that obtained when the observer is on the horizontal axis (the y-axis in this instance), that is at the point of maximum signal strength. This point is also referred to as the 'beam peak'. In principle, the observer can move around the circle and thus log, at different angles, a complete record of received power levels referenced to this maximum. This log of records is the radiation pattern which, for this antenna type, is the same in all ϕ-plane cuts. However, the reality on an antenna field site is that the observer's position is occupied by a stationary transmitter pointing at the dipole, which in turn rotates on a turntable about its z-axis to receive the transmitted energy, according to its radiation pattern (this is a reciprocity principle which allows in principle for the far field observer and the antenna under test to be interchanged).

The concept of directivity is important in antenna theory. It is a directional performance indicator that shows just how well energy is radiated by the antenna in a direction of maximum signal strength (the peak of the beam), compared with the total amount of energy that it radiates. It is computed or calculated from the antenna's radiation pattern, which is measurable and predictable. It is closely related to the concept of antenna gain, G say, via an overall multiplicative efficiency factor, which should ideally be small. The efficiency factor due to system losses is deducted from the directivity figure to produce a value for the antenna gain. It is a quantity which is close in value to the directivity, but it is beyond the scope of this exercise to address it more fully. Suffice it to say, it is one that can be measured on a field site, for example by comparing the antenna's maximum power received from a distant radiating source with that similarly received by a gain standard, which typically is a rectangular horn whose radiation characteristics in respect of directivity and gain are well known. For instance, if G_T and G_S respectively denote the gains of the test antenna and the standard, and if P_T and P_S similarly denote their powers received in the test configuration then the gain of the antenna under test is determined from the formula $G_T / G_S = P_T / P_S$. On field sites where gain is measured, it is generally done to no more than two decimal places on a dB scale. Thus the determination of directivity to no more than the same number of decimal places on such a scale is not necessary. Knowing that the antenna gain is close to the directivity, it is possible to anticipate roughly a value for it by first determining the half power beam width from a radiation pattern measurement. Then the approximate relationship of the type proposed here between this and the directivity allows for an immediate feel for the antenna gain to see if a likely performance is on track. If it is not, there is a cause for concern. Such checks and balances are an essential part of any design, manufacturing and test process.

References

1. Constantine A. Balanis, *Antenna theory analysis and design*, (3rd edn), Wiley (2005).
2. Milton Abramovitz and Irene Stegun, *Handbook of Mathematical Functions*, Dover (1964).
3. John D. Mahony, Omnidirectional pattern directivity in the absence of minor lobes revisited, *IEEE Antennas and Propagation Magazine*, Vol. 40, No. 4 (August 1998) pp. 76-77.
4. John Mahony, An improved approximation to a well-known integral, *Math. Gaz.* **92** (March 2008) pp. 106-110.

10.1017/mag.2022.62

JOHN D. MAHONY
5 Bluewater View, Mount Pleasant,
Christchurch 8081, New Zealand
e-mail: *johndmahony@gmail.com*