Ground Waves – The End Of A Very Old Era?

by Brian Austin, G0GSF

The recent announcement that the BBC's long wave service on 198kHz may about to be discontinued will surely have rattled the tea cups in the Long Room at Lords. 'But what', trumpeted the members in their egg and bacon livery, 'will become of TMS?'.

To those not of this island, some translation may be necessary. TMS (Test Match Special) is to cricket what egg is to bacon, gin to tonic and Nick to Dave, etc. But before we all begin to panic we must assume that the BBC, in its wisdom, has some alternative arrangements afoot, or in the pipeline, and the cricket-following fraternity will not be abandoned. Or at least not yet.

However, what is very clear is that any move to close the long wave (and possibly the medium wave services too?) will bring to an end an era of broadcasting which started officially in the UK in November 1922. Even prior to that, in January 1920, the Marconi Company in Chelmsford had begun transmitting news items and music for two half-hourly periods every day on the same frequency as used by Poldhu for its telegraphic service to shipping [1].

Evidently, many reports were received from ships over 1600km away so one must assume that during those periods the ships' radio officers were not tearing their hair out, but were off duty and enjoying the entertainment. The transmissions took place on a wavelength of 2800m (or a frequency of about 107kHz), which places them in what subsequently became known as the LF or low frequency band from 30kHz to 300kHz. Common practice in those days was always to describe such things in terms of wavelength rather than frequency and so the long wave bands became synonymous with the 'wireless', and a whole new era of communicating without wires had begun.

How Radio Waves Propagate

Very little was known then about the mechanism by which those signals actually travelled from transmitter to receiver, because the science of radio propagation was naturally very much in its infancy. However, many physicists and mathematicians had been studying the problem following Marconi's remarkable achievement of spanning the Atlantic from Poldhu in Cornwall to Newfoundland in Canada only some 20 years before.

Various theoretical descriptions of the process were published and it was initially suggested that the electromagnetic waves were following the curvature of the Earth, but further scientific work began to cast doubt on this, particularly given the distance over the Poldhu to Newfoundland path, plus the rather sizeable bulge of the Earth (or more particularly the ocean) in between.

Other theories floated the idea that something in the rarefied atmosphere above the Earth was causing the rays to be reflected back beyond the bulge, thereby avoiding that particular obstacle. But what was this convenient reflector? We should remember, of course, that all this was taking place many years before Appleton provided conclusive proof, in 1924, of the existence of what is now called the 'ionosphere'. Before then everything was speculative and all mechanisms of possible propagation were up for grabs.

The idea that there might be a region of ionized gas in the upper atmosphere wasn't new. It was actually first mooted by a number of people, but it was Kennelly in the USA and Heaviside in England who were credited, quite independently in 1902, with suggesting that it could possibly cause electromagnetic waves, emanating from a transmitter at one point to be returned to Earth, after reflection, some considerable distance away thereby avoiding any obstacles in between [2].

Despite this plausible idea, it required the experiments some twenty years later of Breit and Tuve in the US and Appleton in England to positively reveal its existence. During the interim considerable research effort, mainly in Europe, went into trying to explain how radio waves actually propagated over the surface of the Earth without any intervention from above. Eventually the concept of what we now call the 'ground wave' was born.

Incidentally, it is intriguing that even today, after massive progress in the understanding of how EM waves travel through all sorts of media, there is still no generally accepted explanation as to how Marconi's spark-generated emissions travelled the almost 4000km from Poldhu to Newfoundland. Those who attended the *Institution of Electrical Engineers'* (IEE) conference in London, held in 1995 to commemorate 100 years of radio, will remember how other sparks flew when the various competing arguments were presented with much vigour by their protagonists [3].

Ground Waves And Space Waves

In 1907, the German mathematical physicist, Jonathan Zenneck (1871-1959) published the first scientific paper [4] that offered an explanation for the way electromagnetic energy in the form of radio waves interacted with the ground as they propagated in the air above. Zenneck showed that the electrical characteristics of the Earth were significant factors in this process and, also, that the frequency played a major part too.

Two years later another German by the name of Arnold Sommerfeld (1868-1951) extended this work[5] by including the transmitting antenna in his analysis,

Fig.1. Arnold Sommerfeld (left) and Jonathan Zenneck: pioneers of radio propagation

while agreeing with Zenneck that the energy travelled as a surface wave along the ground-air interface. Thus, Marconi's belief that radio waves were actually guided by the Earth was seemingly vindicated. This surface wave soon became known as the Zenneck surface wave and it has been the subject of much academic argument ever since. This article will attempt to cover some of that ground, as it were, without becoming embroiled in the mathematics, which is formidable.

Some Definitions

Before launching into the subject it's perhaps advisable, right at the outset, to define quite what is meant by a surface wave and a ground wave. Given the complexity of the subject, and the many opinionated scientists who've studied it ever since the days of Zenneck and Sommerfeld, it should come as no surprise to learn that even such definitions were fertile ground for much heated debate. Fortunately, the *Institute of Electrical and Electronics Engineers* (IEEE) in the USA – the biggest scientific body of its kind in the world – has provided the definitions that now seem pretty watertight.

Essentially, the ground wave is made up of two components: a *surface wave* and a *space wave*, with the space wave itself consisting of two components, the direct wave and the ground-reflected wave. **Figure 2** will hopefully make all this clear. It should be noted that the surface wave is attenuated as it propagates, not solely because it spreads out – as all waves do – but also because it 'leaks' into the ground where it is absorbed. As a consequence the surface wave dies out much more rapidly than the space wave. Another important feature to bear in mind is that the two components of the space wave (the direct and reflected rays) are approximately equal in amplitude, but opposite in phase when the antenna (whether vertical or horizontal) is very close to the ground. This means that the space wave is cancelled out and so, in that special case, all propagation is via the surface wave[6].

It should be noted that the scales in Fig.2 have been exaggerated in the interests of clarity. In practice the antennas are usually much closer to the ground, making the paths' lengths almost equal and the ray angle very small. In addition, the Earth is not flat so, for long distance propagation, due account must also be taken of the effects of Earth curvature.

Confusion And Controversy

Following this early theoretical work, there was a lull of a few years before the matter of how radio waves travel across the Earth's surface once again became a topic that exercised many minds – and continues to do so. Probably no other aspect of radio communication theory and practice has generated as much confusion or has led to as much debate and, certainly, few others have involved such mathematical complexity. And that made the subject so ripe for intensive study by many notable scientists and engineers.

Once again the spur for further work came from Germany with the publication, in 1919, of a paper by H. Weyl on the propagation of plane waves over a plane conductor. However, there was nothing plane about either Weyl's method or his conclusions. Twenty pages of dense calculations failed to reveal the existence of Zenneck's surface wave and that immediately caused Sommerfeld to re-immerse himself in the problem.

Meanwhile, Zenneck himself, who had been in the United States representing his government in a patent dispute and was then interned as an enemy alien when the US entered the First World War, played no further part in this expanding saga after his repatriation in 1920. Sommerfeld's next paper appeared in 1926. He now used a different method to the one he'd adopted seventeen years before and this time he confirmed Weyl's finding that the surface wave was apparently a mirage!

The result, if indeed it was noticed at all, caused few anxious moments amongst radio's growing band of practitioners for, by then, the thermionic valve had appeared and long distance radio circuits were in general use both at sea and on land. The Imperial Wireless Chain, first mooted in 1910

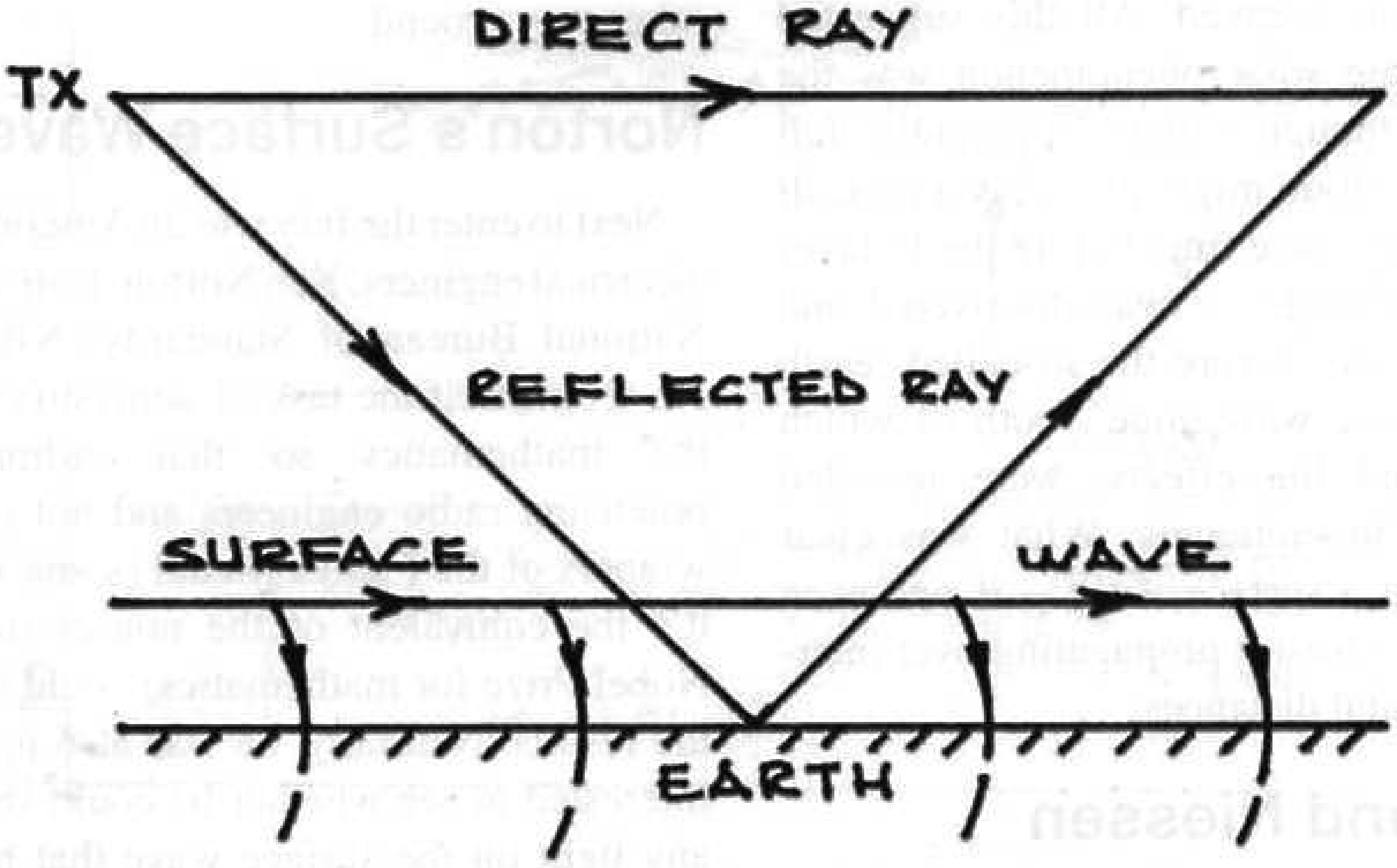

Fig.2. The ground wave and its constituent parts: the surface wave and the space wave, which itself consists of direct and reflected rays

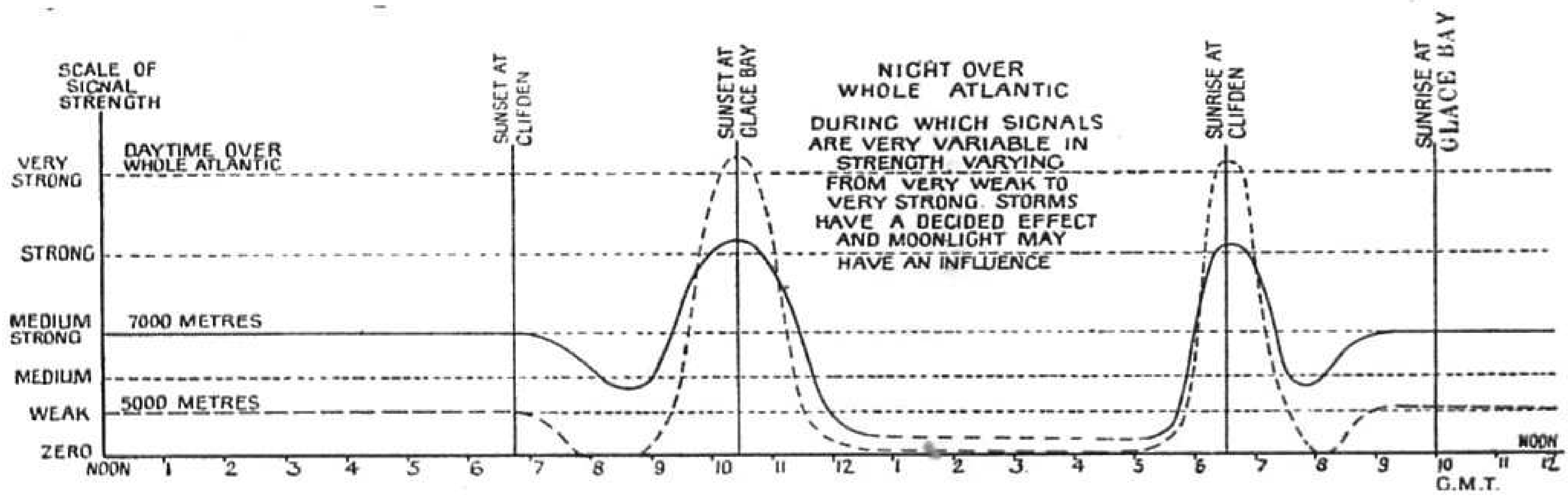

Fig.3. The daily variation in signals received in 1911 at Clifden, Ireland from Glace Bay, Nova Scotia at very long wavelengths [Ref.1, vol.2, p120]

to link Britain with much of its Empire by long wave wireless, had now given way to the short wave Beam System that brought the BBC to the world. This was another significant Marconi development that achieved the same purpose as the Imperial Wireless Chain, but it did so far more efficiently and more economically too. But the Beam System was at the mercy of the ionosphere and so understanding its vagaries soon become a major field of scientific investigation.

In addition, for some physicists, there also remained many unanswered questions about how antennas behave when close to the Earth and, also, how the signals they radiate are affected as much by the earth beneath them as by the ionosphere above. Measurements made as early as 1911 on the LF circuit between Glace Bay in Nova Scotia and Clifden in Ireland showed evidence of propagation phenomena that were decidedly puzzling.

Signals were considerably stronger during the day than at night, while at sunrise and sunset there were brief but very marked peaks in signal strength that soon decayed. All this suggested that some solar phenomenon was the cause, though others apparently felt that the moon might also play a part! It would be some time before the D-layer of the ionosphere was discovered and even longer before the so-called 'earth ionosphere waveguide', both of which explained the effects, were revealed by the theoreticians. What was clear was that a surface wave – if one even existed – wasn't propagating over inter-continental distances.

Pol and Niessen

In 1930, following a decade of further research, a Dutch electrical engineer at the Philips Research Laboratories by the name of Balth van der Pol and his colleague K.F. Niessen again attacked the problem of how EM waves might hug the surface of the Earth. They too made no concessions in their mathematical rigour and again their results confirmed Sommerfeld's findings of 1926 which concluded that the Zenneck surface wave did not exist. But the experimentalists concerned with communications over much shorter distances doggedly disagreed.

Whether it was Zenneck's or anyone else's surface wave didn't matter to them because the evidence for some sort of surface wave, of whatever name or specifics, was there for all to see. Long wave signals definitely did travel very effectively over distances up to a couple of hundred kilometres, and the only possible path between transmitter and receiver was directly over the Earth's surface and not as a result of any reflection, refraction, diffraction or any other even more exotic mode of propagation. And, of course, it was well known that if the theory didn't fit the experiment you bury the theory, not the other way around.

Norton's Surface Wave

Next to enter the fray was an American electrical engineer, Ken Norton, from the National Bureau of Standards (NBS). He set himself the task of demystifying the mathematics so that ordinary practising radio engineers and not just winners of the Field's Medal (some say it's the equivalent of the non-existent Nobel Prize for mathematics) could use the results. Naturally, he was also most interested to see whether he could shed any light on the surface wave that had so muddied the waters to date. In 1935, Norton published his initial conclusions as a brief research note in *Nature*, the world's pre-eminent scientific journal at the time. This guaranteed it maximum impact. In the note he announced that Sommerfeld, in his 1909 paper, had made a mathematical error amounting to the use of the incorrect sign in one of his equations: Zenneck's surface wave was definitely dead but in playing the hatchet man Norton then introduced a surface wave of his own!

In a masterful simplification of previously obtuse formulations he made it a simple matter for ordinary engineers to analyse, and hence to design, radio communication circuits operating over ground wave paths. Until then no one had talked of ground waves, as such. Zenneck's surface waves (and briefly those of Sommerfeld in 1909) existed only mathematically because of Zenneck's initial assumptions that

Fig.4. Ken Norton, the man who turned the mathematics of ground wave propagation into a useful engineering tool (Photo by courtesy of the NIST archives, USA)

did not include the antenna in the calculations, while Sommerfeld's of 1909 were now seen to be bedevilled by the incorrect sign (though even that is in dispute!). On the other hand, Norton showed that what he called the ground wave did indeed exist but only under circumstances that neither Zenneck nor Sommerfeld, nor indeed anyone else, had foreseen. Norton's engineering approach is discussed briefly in the accompanying box (see the next page).

So That's It Then ... ?

Well not quite. When the electrical conductivity of the ground is infinite (an impossible situation in practice, but an assumption that is often made by mathematicians to simplify problems), the surface wave disappears and all that is left is the space wave. (See Fig.2 to refresh the memory). These high-angle rays travel into space where they may be reflected by the ionosphere and so produce strong signals at great distances. We all recognise this as the well-known process of worldwide communication at HF. On the other hand, finite ground conductivity complicates the issue and leads to the existence of a surface wave which propagates directly over the Earth. Instead of decaying very slowly as Zenneck's mathematical version did, Norton's version decreased in amplitude rather more rapidly, but it was real and mighty useful too.

And there the matter rested for many years until a most mysterious wave was discovered in more recent times. The 'trapped surface wave' emerged, not surprisingly, from some more intractable mathematics. It was apparently first described by J.R. Wait whose theoretical work underpinned the US Navy's ELF communications systems with submarines (see *RB118*) and much else besides. Jim Wait published a book on all these matters in 1962[8], and though it may not actually have set the world on fire it did set the standard for research into the way in which radio waves, across a very wide spectrum, interact with the Earth and the oceans around us.

Fig.5. Jim Wait, the Canadian-born US scientist who was the sage in all matters electromagnetic. He wrote three books and almost 800 scientific papers on the subject

Wait's influence on a generation of geophysicists and engineers was profound, not only because of his mastery of his subject but also because of the way he wrote about it. Once, when I was having lunch with him in the company of many of my engineering colleagues, he was asked what he considered were the most important subjects in a modern university electrical engineering degree course. Without hesitation he answered, "English and mathematics – in that order". Needless to say those around the table were rather surprised; some even a bit scornful. But Wait's reply is worth thinking about.

Trapped Surface Waves

The trapped surface wave funnels along between layers, usually at the Earth's surface, under certain special conditions. One of these is when the sea is covered by a layer of ice as in the polar regions of the planet. And there was indeed evidence that radio waves behaved abnormally when propagating over ice-covered seas at high latitudes. But there is also an example much closer to home. In 1980, following the publication of measured data by the BBC[9], Wait suggested that the anomalous signal strengths of medium wave and long wave broadcasts across Greater London were also evidence of the same trapped surface wave phenomenon. The magnetic field component of the propagating waves was significantly stronger than its electrical counterpart.

Without going into any technical detail (for reasons that will now be clear) this effect, according to Wait, indicated that such an urban sprawl and the ice-covered sea mentioned before both favour the magnetic field of the EM wave. Since magnetic fields are produced by coils and loops carrying current, that type of terrain is said to be 'inductive' and that, according to the mathematics, is precisely the condition that will cause a surface wave to be trapped and thereby to propagate with lower attenuation than the conventional ground wave.

But, and this should surprise no one, others disagreed! A Harvard physicist by the name of R.W.P. King, a man of considerable prowess in the field, produced no evidence in a number of his scientific papers (of which[11] is but one example) of the existence of such a trapped surface wave. King and Wait immediately locked horns in the rarefied atmosphere of the academic literature. Wait stood his ground and soon he was joined by more heavyweights, most

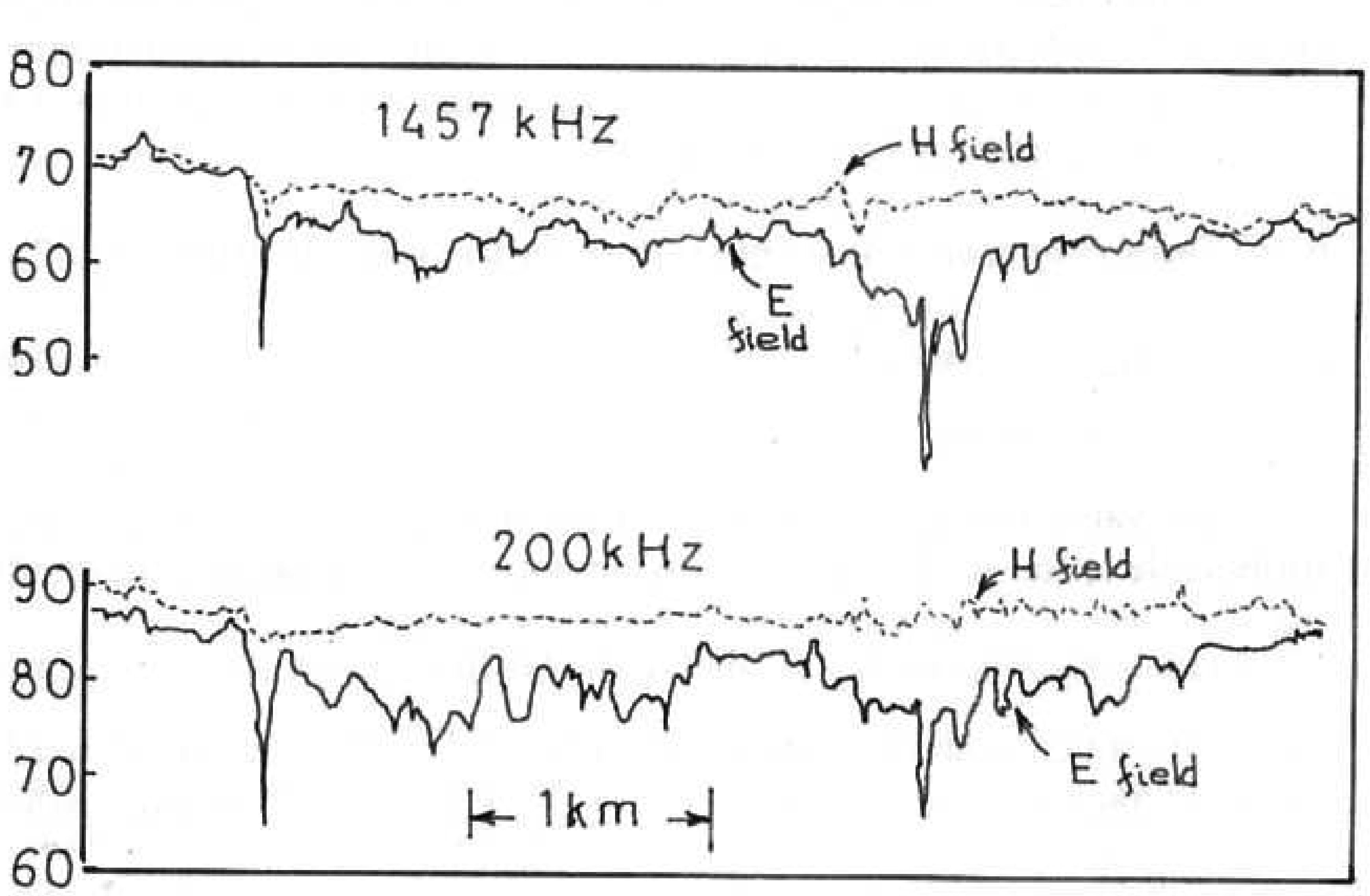

Fig.6. Measured variation in the electric and magnetic fields at medium and long wave frequencies at certain points across London. (Adapted from [9])

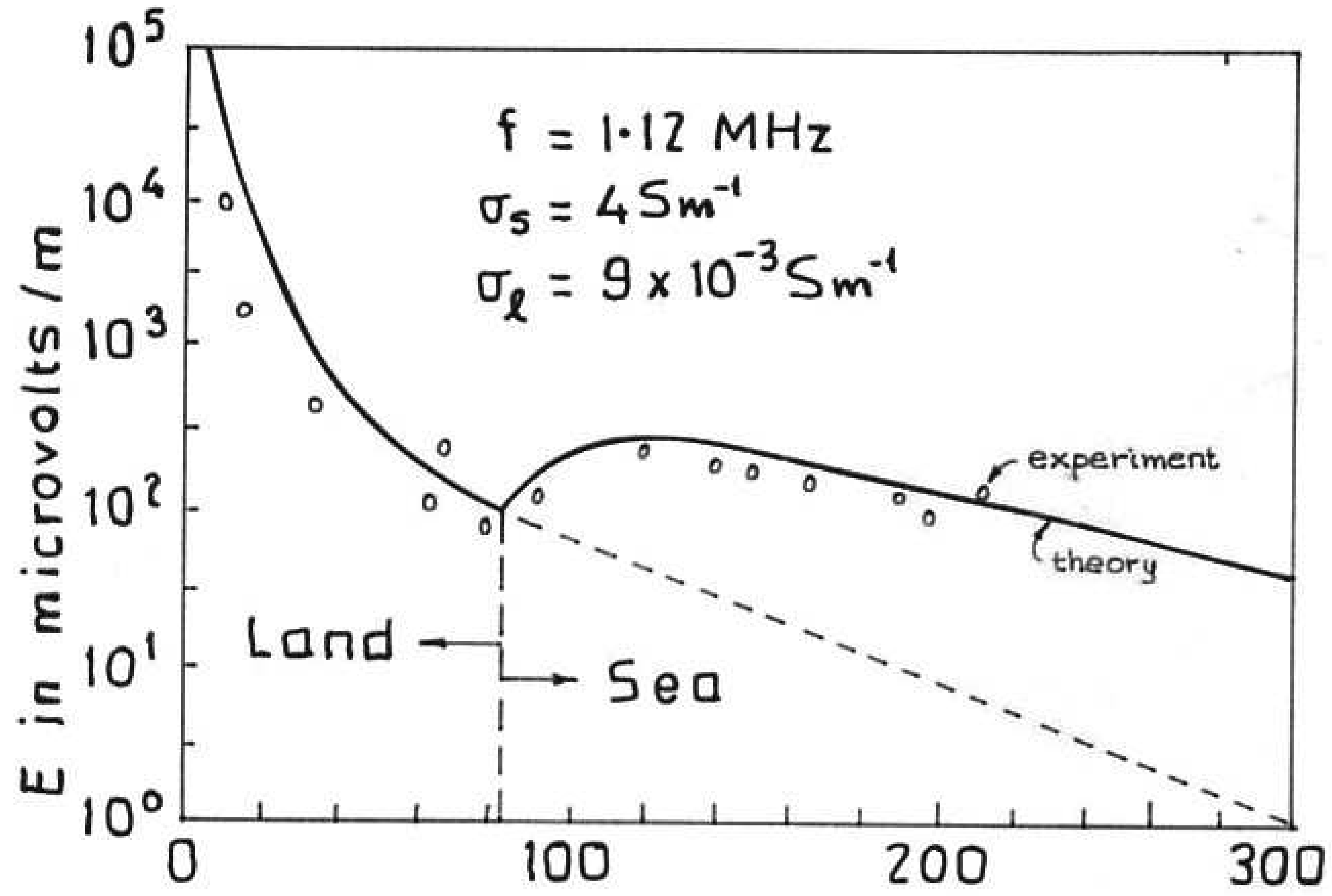

Fig.7. The ground wave 'recovery effect' over a mixed land and sea path. (Adapted from [10])

notably, Bob Collin of Case Western Reserve University, who weighed in with their arguments in support of the surface wave conjecture [12].

Clearly there's never a dull moment in the world of radio propagation, especially when it occurs close to the surface of the planet, and there is more to come!

Something For Nothing?

From the evidence already presented it's clear that the fields are directly affected by the topography and by the electrical features of the ground beneath them. Unlike the case of an all-sea path, the ground is likely to present very mixed terrain and so the propagation will vary from one place to another and even in different directions. But what is most fascinating is what happens when a ground wave signal crosses the boundary between land and sea.

This produces a most unexpected result, known as the 'recovery effect' when the direction of travel is from land to sea. **Figure 7** shows the case of a signal from a medium wave land-based transmitter, located some distance from the coastline. Initially the signal strength decays rapidly with distance as its energy is absorbed by the Earth. However, immediately on crossing the coastline, it recovers and actually increases over a certain distance before decaying again, but at a slower rate than before. Theory predicted this effect; experiments soon confirmed it, as will be noted in Fig.7.

Initially it might seem that one was getting something for nothing. Surely

Norton's Engineering Approach

Norton's great contribution to the science of ground wave propagation was to express very complicated mathematics in a form that was useful to practising radio engineers. His approach is outlined in what follows.

If the Earth's curvature is ignored, (which is reasonable when the distance, in km, between transmitter of receiver is less than $80/f_{MHz}^{1/3}$), then the ground wave field strength is given by the equation $E = E_1 A/d$, where E is the received signal, E_1 the unattenuated field strength at 1km, d is the distance to the receiver, and A the ground wave attenuation factor. If the ground is primarily a conductor (generally true at frequencies less than a few MHz), then the attenuation factor is given by $A = \dfrac{2+0.3p}{2+p+0.6p^2}$, where

$p = 0.582d(km)f^2(MHz)/\sigma(mS/m)$ is usually called the *numerical distance*, since there's rather more involved than just distance!

E_1 can either be calculated, if the radiation efficiency of the transmitting antenna is known, or it can be measured. For broadcasting applications it is usually measured and from that value, plus the attenuation factor A, the transmitter power required to provide satisfactory reception over a given distance can be determined as shown in the numerical example below.

BBC *Third Programme* transmission from Daventry to Hull (c.1957).

d = 183km; f = 647kHz; P_t = 150kW; E_1 (meas.) = 4500mV/m; σ = 12mS/m. Hence, on substituting these into the expressions above, p = 3.71 and A = 0.223. Then, $E = \dfrac{4500 \times 0.223}{183} = 5.5$mV/m is the field strength at Hull. The value actually measured by the BBC was 'about 7mV/m'. The difference of around 2dB is probably attributable to the single value of conductivity used to represent the conductivity of the total path.

To calculate E_1, we use the fact that in the far field (i.e. when $d > \lambda/2\pi$) then $E = \dfrac{\sqrt{30G_tP_t}}{d}$, where G_t is the gain of the antenna and P_t is the power delivered to it by the transmitter. Antenna gain and directivity are related by the efficiency η_t, thus $G_t = \eta_t D_t$; while for a short monopole above perfect ground D_t=3, therefore $E = \dfrac{\sqrt{90\eta_t P_t}}{d}$. For the prescribed distance of 1km, the value of E_1 in mV/m is $\sqrt{90\eta_t P_t}$. Clearly, the radiation efficiency of the antenna, or its radiated power, are needed in order to calculate E_1.

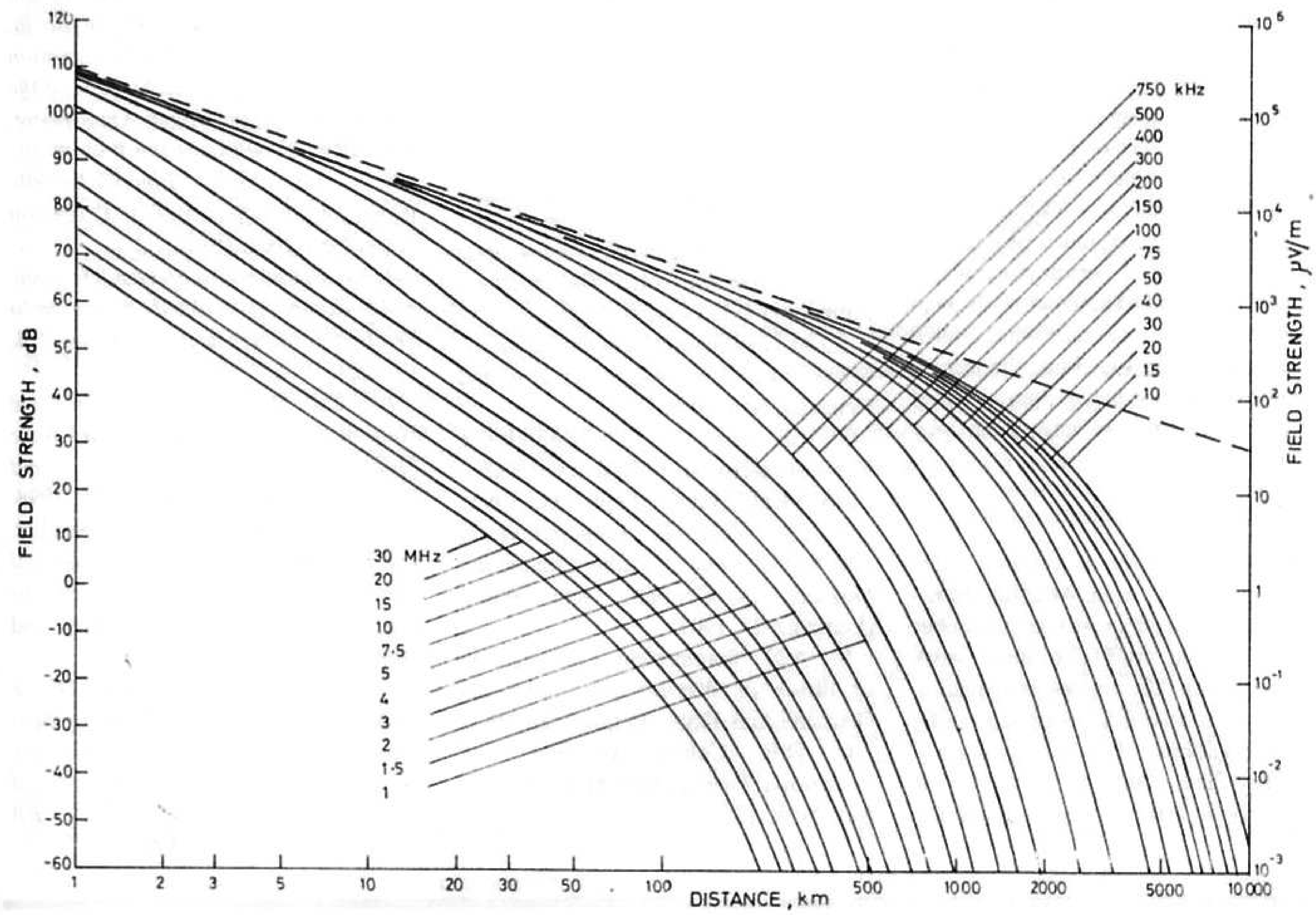

Fig.8. A family of ground wave propagation curves at various frequencies for ground conductivity of 3mS/m and relative permittivity of 4. The dashed line shows the perfect ground case for a short vertical antenna radiating 1kW

Table 1: Transmission Distance At Arnhem

Frequency (MHz)	WS68P: d(max) km	WS 22: d(max) km
2	7.2	9.9
3	5.1	7.0
4	4.3	5.8

it's impossible for a signal suddenly to increase in strength? However, energy considerations reveal that no fundamental rules are being broken. Rather, the explanation lies in the distribution of that energy over the two regions. Over the poorly conducting land it is distributed high above the surface whereas over the sea, a far better conductor, it is much more tightly bound to that surface. At the boundary, re-distribution has to occur in order to restore the energy balance.

Practical Applications

So far, much of this discussion has been about the theories, the ideas and even the controversies about how radio waves behave when close to the Earth. In addition, the parts played over the last 100 years by leading radio scientists in unravelling the intricacies of it all have added much light, some colour and also, it must be admitted, not a little heat as well, at times. But ultimately it is the applications of those medium and long electromagnetic waves that are of real interest. Probably foremost amongst them has been broadcasting, for without ground waves there would have been no broadcasting in those early days and, in all likelihood therefore, radio as we know it today might have been a very different beast.

Broadcasters required effective tools to plan their services and, as we've seen, so much is owed to Ken Norton and his simplified methods of calculating field strengths for a given set of practical situations. Nowadays, such methods have been turned into computer programs, but for many years graphical solutions were the norm and **Figure 8** shows the form of graph used to determine the variation of ground wave field strength with distance above a curved Earth at frequencies from VLF to HF.

Ground waves also propagate effectively at HF as that graph shows. Until the 1960s when the British Army adopted VHF for much of its tactical communications, 'HF ground wave' was the dominant mode in use. As the equipment was either carried on a man's back, or was mounted in a vehicle, the antennas were, *per force*, always short whips. This usually made them very inefficient, especially at the lower HF frequencies required for maximum possible ground wave range. In addition, the transmitters of that time only produced just a few watts, often much less. As a result, tactical communications presented some problems and never was this more serious than during the Battle of Arnhem in late September 1944.

Many military radio enthusiasts (both amateur and professional) have tried to explain quite why the British army's radio communications were essentially non-existent for long periods throughout that battle. My calculations using the techniques described in this article produced the results shown in **Table 1**. It shows the maximum achievable distance, over the frequency range of interest, by the two types of radio then in use by the army at Arnhem: the WS 68P with an output of 250mW, and the WS 22 capable of 6dB more power. Both sets were assumed to be using the standard 3.4m whip antenna, either on a man's back or mounted on a vehicle, such as a Jeep. The ground conductivity was assumed to be 3mS/m (typical of much of Arnhem) and the criterion of performance was a signal-to-noise ratio (SNR) of 10dB, with the receiver being externally noise limited by atmospheric noise. Since the critical distances during the battle often exceeded 10km one might now understand why there was a problem.

In Conclusion

Today, well over a century since Marconi's first trans-Atlantic transmission, we stand in awe of a number of remarkable men for their work in explaining the radio propagation processes involved in everyday radio communications. Many of those early pioneers died within the first half of the previous century, but some of the more recent giants left us much more recently. Jim Wait died in 1998, Ronold King passed away at the age of 100 in 2005, while Bob Collin died as recently as 2010. Is that the end of an era? Perhaps not.

References

1. H.M. Dowsett, *Wireless Telegraphy and Broadcasting* vol.1, The Gresham Publishing Co., London, 1923.

2. J.M. Goodman, *HF Communications: Science and Technology*, Van Nostrand Reinhold, New York, 1992.

3. IEE (now the IET) Conference Publication No. 411: *100 Years of Radio*, Savoy Place, London, September 1995.

4. J.A. Zenneck, *Propagation of EM Waves along a Plane Conducting Surface*, (in German), Ann. Phys., (Leipzig), 23, 1907.

5. A.N Sommerfeld, *Propagation of Waves in Wireless Telegraphy*, (in German), Ann. Phys., (Leipzig), 28, 1909.

6. R.E. Collin, *Antennas and Radiowave Propagation*, McGraw-Hill, New York, 1985.

7. K.A. Norton, *The Propagation of Radio Waves over a Plane Earth*, Nature, 135, 1935 and *The Propagation of Radio Waves over the Surface of the Earth and in the Upper Atmosphere*, Proc. IRE, 24, 1936.

8. J.R. Wait, *Electromagnetic Waves in Stratified Media*, Pergamon Press, New York, 1962.

9. J.H.Causebrook, *Electric/Magnetic Field Ratios of Ground Waves in a Realistic Terrain*, *Electronics Letters*, 14, 19, Sept. 1978.

10. J.R Wait, *The Ancient and Modern History of EM Ground-Wave Propagation*, *IEEE Antennas and Propagation Magazine*, 40, 5, Oct. 1998.

11. R.W.P. King and S.S. Sandler, *The Electromagnetic Field of a Vertical Electric Dipole over the Earth or Sea*, IEEE Trans. on Antennas and Propagation, 42, 3, 1994.

12. R.E. Collin, *Hertzian Dipole Radiating over a Lossy Earth or Sea: Some Early and Late 20th-Century Controversies*, *IEEE Antennas and Propagation Magazine*, ~~40~~ 46, 2, 2004. ***RB***