The Heaviside signal

An alternative view of the transverse electromagnetic wave

by I. Catt, CAM Consultants

This article proposes a different picture of electromagnetic propagation from the familiar "rolling wave" idea in which electric and magnetic fields topple over and forward, continually changing into each other as they go. The author postulates "an unchanging slab of E x H energy current" travelling forward at the speed of light, and names it "the Heaviside signal" after a concept expressed in the writings of Oliver Heaviside. This process does not rely on a causal relationship between the electric and magnetic fields, which are seen as co-existent.

MAXWELL faced up to the paradox that whereas electric circuits, in order to function properly by allowing the passage of electric current, were thought to require a complete closed circuit of conductors, electric current still seemed to flow for a time when a capacitor (which of course is an open circuit) was placed in series with the closed loop of conductors. He "cut the Gordian knot" (according to Heaviside)[1] by postulating that a new kind of current, which he called "displacement current", leapt across the plates within the capacitor. This electric current, which was uniformly distributed in the space between the capacitor plates, could even flow through a vacuum.

Maxwell followed up this daring idea by suggesting that electromagnetic waves might exist in space. Scepticism about his postulated "displacement current" was silenced in 1887 when Hertz discovered the predicted waves in space. The classic pre-Popperian requirement of a good scientific theory seemed to have been met – the prediction of further results which are later confirmed by experiment.

There are two versions of the transverse electromagnetic wave, the "rolling wave," and what we shall call here the "Heaviside signal." We shall discuss only the wide variety of views among those who believe (with the relativists) that there is no instantaneous action at a distance.

The rolling wave

The lack of action at a distance creates a fundamental difficulty for the wave in space if it is to be launched by a force in the direction of propagation. The key to the ability of a force to project a wave is that there is a pressure difference between two points along the line of propagation. However, knowledge of a difference in pressure between two points A and B which are separated by distance implies instantaneous knowledge at B of the pressure at A; that is, instantaneous action at a distance, which has been outlawed.

This dilemma seems to be overcome if it is postulated that the force which projects the wave is a lateral, shear, force. It seems a shear force can act at a point, and so not contradict relativity whereas a longitudinal force cannot.

The above kind of reasoning, combined with the postulation of displacement current, which seemed to flow at right angles to the direction of propagation, joined forces to create the notion of the rolling wave. The rolling wave contains alternating concentrations of magnetic energy $\frac{1}{2}\mu H^2$ and electric energy $\frac{1}{2}\epsilon E^2$ in the direction of propagation. It is useful to think of a road with alternate red trucks and white motor cars. The magnetic energy or flux (by Faraday's law of induction) generates electric energy and displacement current ahead of itself, which in turn (by the Biot-Savart Law) generates magnetic flux, or energy, ahead of *itself*. Each type of energy, or flux, topples over and forward, changing as it topples into the other kind of energy. It is as though in the road containing the alternate red trucks and white cars, first the red trucks reappear as white cars a little further ahead while at the same time the white cars turn into red trucks a little further ahead; then the trucks and cars change back again, moving forward a little with each metamorphosis. The analogy with the pendulum has been proposed. One can think of a long line of pendulums, alternate ones having potential energy and kinetic energy, and communicating their energy forward step by step with a change of type of energy at each one.

The Heaviside signal

Opposed to the rolling wave is what we have called the Heaviside signal. The most highly developed form of this view is that at any point in space, an electromagnetic signal always contains one kind of energy only, which is equal to the product of E and H at that point, where

$$\frac{E}{H}=\sqrt{\frac{\mu}{\epsilon}}. \qquad \text{Energy density} = \frac{E.H}{c}$$

Further, the Heaviside signal always travels forward unchanged at the speed of light, $c=1/\sqrt{\mu\epsilon}$, and never any slower. *E*, *H* and c are always mutually perpendicular.

The two men most likely to understand the "Heaviside signal" point of view and to oppose the "rolling wave" were Oliver Heaviside himself, in honour of whom it has been given its name, and Poynting, the man whose name is attached to the vector $E \times H$. However, their writings show that neither man arrived at a full understanding of the Heaviside signal described in the previous paragraph.

Heaviside vacillated between the two views, the rolling wave and the Heaviside signal. He always applauded the idea of displacement current, which appears to put him on the side of the rolling wave. Further, on page 6, art. 453 of volume 3 of his "Electromagnetic Theory", when he says that the curl of *E*, not *E* itself, is the real source of the waves, he is again arguing for the rolling wave. Curliness is obviously a bid for shear, vorticular forces, a concept intrinsic to the rolling wave. However, elsewhere he seems to stand firmly for the Heaviside signal. For instance *(ibid*, art. 451, page 4), he says, "It carries all its properties with it unchanged," which is a clear statement of the Heaviside signal. In art. 452, the mention of a "slab" of signal is strongly on the side of the Heaviside signal. Heaviside mentions the slab elsewhere in his writings. One does not conceive of slabs rolling, or generating shear forces or stresses. Almost by definition, a slab, like a slab of heavy granite, moves forward unchanged at constant velocity.

Professor Poynting, who first suggested that energy was distributed in space with a density $E\times H$, also had a partial vision of the Heaviside signal. He definitely did not know that *E* is always perpendicular to *H*, and that the $\times$ in $E\times H$ means simply multiplication. (He had a term $\sin\theta$ for the angle between them.) Poynting was writing before the general agreement that light is electromagnetic, and so did not know that this Poynting energy $E\times H$ always moved forward (in the third dimension) at a constant speed, $1/\sqrt{\mu\epsilon}$, the velocity of light in the medium.

Poynting had a very good grasp of the direction of energy flow and its magnitude, but did not seem to understand the importance of reflections at a change of medium, which leads one to think of one energy current $E \times H$ flowing backwards along its previous path, passing through the next portion of forward travelling energy current. This superposition of forward and backward energy currents (implicit in the phrases "phase velocity" and "group velocity") has prevented a clear understanding of the electromagnetic wave.

For fifty years, technology did not give us the power to drive the medium with an electromagnetic signal. With the low power at our disposal, all we could do was resonate the medium with periodic (sinusoidal) excitation in the same way as we move a child on a swing. In a resonant medium, energy is necessarily flowing in both directions; most of the forward energy returns to aid the source on the next cycle.

Our inability to drive a medium except periodically insinuated itself into our group psyche, until we came to assert that nature *was* periodic (and even that it was sinusoidal). Implicit in this view were the wrong beliefs that

(1) electromagnetic energy is necessarily contrapuntal,

(2) $E/H = \sqrt{(\mu/\epsilon)}$ is not always true, (e.g. when two waves are passing through each other so that H cancels but E does not, so that $E/H = \infty$), and

(3) signals can travel slower than the speed of light $1/\sqrt{\mu\epsilon}$).

The absurdity of this third idea is easy to demonstrate if we consider a two directional highway. If all cars move at 60 m.p.h. but some (A per hour) move eastwards and some (B per hour) move westwards, no one would argue that the total passage of cars eastwards per hour past a reference point, that is, $(A-B)$, would help us to determine the velocity of cars by the formula

Flow of cars $=(A-B)$ per hour

Distance between cars $= L$

Therefore velocity of cars $=(A-B)L$ m.p.h.

However, this seems to be done, at least subconsciously, with phase velocity and group velocity. The very terms imply some such calculation.

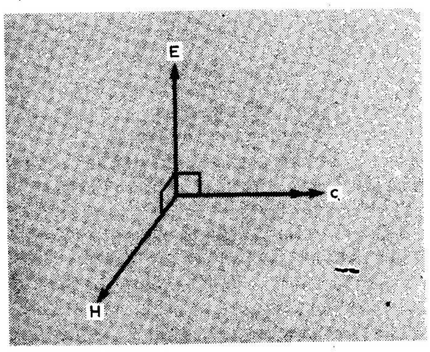
Some ten years ago the successful manufacture of high speed (1ns) logic elements capable of driving a 100 ohm load made it possible, for the first time for fifty years, to drive a medium rather than gently resonate it, as a matter of normal routine. Those driving a high speed logic step could clearly see it travelling at the speed of light for the dielectric (never any slower) and remaining unchanged on its journey. For the first time for seventy years, high speed digital engineers were privileged to see the Heaviside signal, an unchanging slab of $E \times H$ energy current guided between two conductors from one logic gate to the next. Reflections were prevented by proper termination at the destination, so that notions of phase velocity and group velocity evaporated. We saw a slab of energy launched from one point, travelling unaltered, to be absorbed by the terminating resistor at the destination.

At this point we just *had* to unburden ourselves at the theoretical level of implicit contrapuntal notions. A beautiful vision resulted, now called the Heaviside signal, of a lateral strain $E \times H$ (where $E/H = \sqrt{\mu/\epsilon}$ which by definition travelled forward at velocity $1/\sqrt{\mu\epsilon}$. As it travelled forward it filled (or probed) the space ahead of it in the same way as the ripples on the surface of a pond will fill the space (surface) as they come to it. Logic designers maintained a near constant aspect ratio in the space ahead, because whenever this slab came to a change in aspect ratio (= change of characteristic impedance, better termed characteristic resistance) some of the energy current would double back on its tracks according to the well-known laws of reflection. However, this did not lead back to the old "phase velocity" and "group velocity" notions; rather, the slab of energy current split into two slabs, one to continue forward and the other to return, both slabs continuing to probe, or fill, the space presented to them on their journeys.

The Heaviside signal offers us a dramatic simplification of our view of the fundamentals of electromagnetic theory.

Definitions

First define energy current (= TEM wave = Poynting vector) as our primitive, where energy current is as follows:



Now $\sqrt{\mu/\epsilon}$ and $1/\sqrt{\mu\epsilon}$ can be independently defined. Let us define

(a) $\sqrt{\frac{\mu}{\epsilon}} = \frac{E}{H}$

which defines a constant of proportionality for the medium.

(b) $\frac{1}{\sqrt{\mu\epsilon}} =$ velocity of propagation c,

again a constant for the medium.

(c) Define $D = \epsilon E$, $B = \mu H$

Derivations

$\frac{E}{H} = \sqrt{\frac{\mu}{\epsilon}}, \quad B = \mu H$

$$\therefore \frac{E\mu}{B} = \sqrt{\frac{\mu}{\epsilon}} \quad (1)$$

$$\therefore \frac{E}{B} = \frac{1}{\sqrt{\mu\epsilon}} = c \quad (2)$$

$$\therefore E = Bc \quad (3)$$

By definition*,

$$c\frac{\partial E}{\partial x} = -\frac{\partial E}{\partial t} = -c\frac{\partial B}{\partial t} \quad (4)$$

$$\therefore \frac{\partial E}{\partial x} = -\frac{\partial B}{\partial t} \quad (5)$$

This is equation (12.5.1) in Carter (G. W. Carter, The Electromagnetic Field in its Engineering Aspects, Longmans, 1954, page 268), when he believes he is *deriving* the TEM wave, which is supposed to *result* from a causality relationship between E and B (Faraday's law of electromagnetic induction). Carter is clearly developing the rolling wave.

We see then that the equation $\partial E/\partial x = -\partial B/\partial t$ is a simple derivation from the definition of the Heaviside signal and is not based on $\partial B/\partial t$ *causing* E, as Faraday thought he had discovered.

We have shown that the passage of a TEM wave and all the mathematics that has mushroomed around it does not rely on a causality relationship (or interchange) between the electric and magnetic field. Rather, they are co-existent, co-substantial, co-eternal. The medium can only be strained in the two lateral dimensions (E and H) in fixed proportion. [In a similar way, *pressure* in a liquid in direction x does not *cause* pressure in the y (and z) direction; they co-exist.]

Faraday's great discovery in the 1830s was *not* electromagnetic induction; not a causality relationship. His great achievement was to discover that *change* was important. This started us on the road to discovering the now postulated primitive, the Heaviside signal, which can only move; it cannot stand still. Heaviside put together the main features of the new concept, but it took another century to put flesh on to the bare bones.

Reference

1. Oliver Heaviside, Electromagnetic Theory, 1893, London, page 28 section 30.

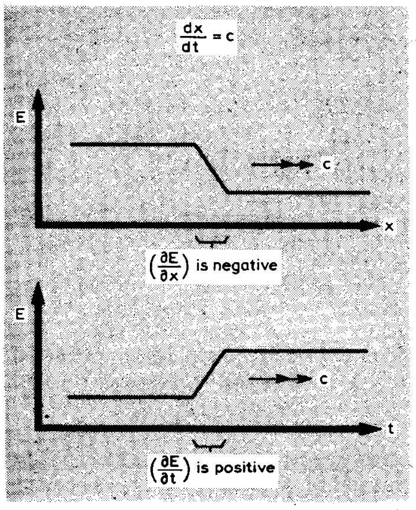
Appendix 1

By convention, if a voltage step is travelling from left to right (i.e. in a positive direction) it has a positive velocity; dx/dt is positive

$\frac{\partial E}{\partial t}$ is positive but $\frac{\partial E}{\partial x}$ is negative. This (reversal) problem is well known by anyone who has drawn out an oscilloscope trace on to paper with voltage and distance axes. This explains the minus sign in equation (4) in the article. When we travel, we *gain* distance while we *lose* time. However, we regard our velocity dx/dt as positive.

It is strange that this ambiguity in sign convention had led to a negative sign in electromagnetic theory. This in turn intro-

*See Appendix 1



duced the idea of a "Lenz's law" reluctance, or back e.m.f., in which lies nested the idea of causality,

$$i \rightarrow \int H dl \quad \text{and} \frac{dB}{dt} \rightarrow v$$

In fact, electric and magnetic fields have a positive relationship, and co-exist rather than cause each other.
Numerically,

$$c\left|\frac{\partial E}{\partial x}\right| = \left|\frac{\partial E}{\partial t}\right|$$

Therefore, since by convention $\partial E/\partial t$ is positive, $\partial E/\partial x$ is negative and c is positive, we must conclude that

$$c\frac{\partial E}{\partial x} = -\frac{\partial E}{\partial t}$$

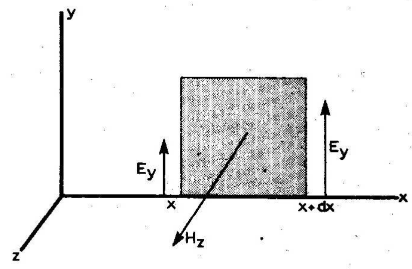
Appendix 2: the rolling wave explained

In this article, two mutually contradictory versions of the transverse electromagnetic wave have been described and compared. These were the rolling wave and the Heaviside signal. This appendix contains the first half of a very clear description of the rolling wave taken from "Fundamentals of Electricity and Magnetism" by Arthur F. Kip, Professor of Physics, University of California, Berkeley, published by McGraw-Hill, 1962, page 320. Only enough of that description is reproduced to make his approach clear.

"... Our demonstration involves the use of the first two Maxwell equations to show that such a postulated time and space variation of E gives rise to a similar time and space variation of H (but at right angles to E) and that this H variation acts back to cause the postulated variation in E. Thus, once such a wave is initiated, it is self-propagating.

"The figure below is used to show the application [of Faraday's law of induction] to the plane E wave, postulated to be moving along the x direction. A convenient closed path is drawn in the xy plane, around which we shall take the line integral of E. This is equated through [Faraday's law] to the rate of change of flux H through the plane bounded by the path of the line integral. Only the vertical parts of the line integral contribute since E is in the y direction, so that $E.\partial x = 0$. If we go around in a counter-clockwise direction, the line integral around the path chosen becomes

$$\oint E\,.\,dl = (E_y)_{x+dx}dy - (E_y)_x dy$$
$$= [(E_y)_{x+dx} - (E_y)_x]dy$$



where we are to take the values of E_y at $x+dx$ and x, respectively. The difference between these two values of E_y at the two positions is $(\partial E_y/\partial x)dx$, so we can write the line integral of Faraday's law of induction as

$$\frac{\partial E_y}{\partial x} dxdy = -\mu_o \frac{\partial H_z}{\partial t} dxdy$$

Since this relationship is true for any area $dxdy$, we may write

$$\frac{\partial E_y}{\partial x} = -\mu_o \frac{\partial H_z}{\partial t}$$

(This ends the extract from Kip. To get to the Carter equation we have to replace μH by B, of course.)

This article is taken from "Electromagnetic Theory", published by C.A.M. Publishing, 17 King Harry Lane, St Albans, Herts. The next seminar by CAM Consultants on digital electronics design will be held at St Albans on August 2-3.