Review of *The Forbidden Equation* by Forrest Bishop

This review has been undertaken at the request of Forrest Bishop. I'm grateful to Forrest for sending me the text, and apologise for the delay of about two years before my reply. This was the result of a number of personal circumstances which have consumed my energies for this period, and resulted in me now being off work with a PTSD-like condition. However the sick leave has given me the time to fulfill my promise to reply eventually.

The text I am working from is dated 2016 and has 14 pages. I do not propose to review this in every detail, because some of the material relates to diputes with other people, which I have not followed and in which I do not wish to get involved. However, the larger part of it relates to a document which I wrote and sent to Ivor Catt, and I will address this part of Forrest's paper. I agree with Forrest that my title, *The Catt Question*, is unspecific, and I shall refer to my document as *The Answer.*

I wish to take this opportunity of pointing out two scribal errors, one gross error, and a notational inconsistency in *The Answer*:

- On page 4, three lines up from the bottom, the text should read "However, if *we* add", where the emphasized word has been omitted. This is just the result of my brain working faster than my pen!

- At the bottom of page 5, in the context of a very crude model of conduction in the two wires of the transmission line, the following equations appear:

  $$\text{Signal: } q = e/a - e/a(1+v/c) = \frac{ev/c}{1+v/c}$$
  $$\text{Return: } q' = e/a - \frac{e}{a(1-v'/c)} = -\frac{ev'/c}{1-v'/c}$$

  In each line the first equation gives q in the form of equation (1) below, and the second equation is supposed to combine the two terms. However, there is obviously a factor of a missing from the denominator of this final term in both equations — this is just a copying error, since the correct form is used on page 6 (see below).

- In the middle of page 4, in the description of the diagram on that page, there is a gross sign error which I have read many times and never noticed. The equation in brackets should read $\partial k_z/\partial z > 0$.

- In section 1, the continuity equations are given with the x-co-ordinate along the transmission line, but in section 2 it is the z-co-ordinate that points in this direction. I shall use the latter version, and when I quote the continuity equations below I shall change the x to z.

1 The Major Problem

The 'Forbidden Equation' of Forrest's title is $i = qc$, relating the line current i to the charge per unit length on the conductor q and the speed of light c. There is nothing forbidden or controversial about the equation, and I shall refer to it as the charge-current relation. Much of what Forrest says about it is fatally undermined by one inconvenient fact:

Matter contains both positive and negative charge.

I don't believe this to be a controversial statement, but it completely alters the way in which the charge-current relation can be interpreted. This is a such an important point that I shall take a little time at the beginning to explain this in detail.

(a) No-one, I think, believes that the charge neutrality of neutral matter indicates that the matter contains *no* charge. It is accepted that in this state the densities of positive and negative charges are equal and opposite.

(b) Similarly, in the electrically-charged state there is *not* a small quantity of one sort of charge, but an imbalance between the amounts of positive and negative charge. This imbalance is always many orders of magnitude smaller than the total charge density of either positive or negative charge.

(c) This has implications for the current: by the same token the total current is the sum of the current flow of the positive charges and the current flow of the negative charges. If we treat the conductor as a one-dimensional, line-like object, in keeping with the variables used by Forrest, we have

$$q = q_+ + q_- \tag{1}$$

where q_+ is the charge per unit length of positive charges, and q_- is the charge per unit length of negative charges, and is therefore a negative quantity. Now suppose, for simplicity, that all the positive charges flow with velocity v_+ in the positive z-direction and all the negative charges flow with velocity v_- in the same direction. These are also signed quantites, and could be positive or negative. The total current i is then

$$i = q_+v_+ + q_-v_-, \tag{2}$$

a sum of two terms, each of which is a charge density multiplied by the velocity with which these charges are moving. Of course, we do not need to make the restrictive assumption that all the charges of each sign flow with the same speed. We could divide up the negative charges into fixed charges bound to the positive charges and mobile charges free to move. We could introduce a range of velocities for the mobile charges. That is, we can make equation (2) more complicated, *but we cannot make it simpler.* It must contain at least these two terms, even if one of them is in fact zero.

(d) Suppose, for example, a neutral wire is moving past us with velocity v. Then because it is neutral we have $q_+ = -q_-$ and $v_+ = v_- = v$. Then

$$i = (q_+ + q_-)v = 0$$

so there is no current flowing in the wire because the two terms in the current cancel. Thus the equation correctly describes zero current flow in a disconnected piece of wire thrown across a room.

(e) In the case in question, the signal conductor of a two-wire transmission line, the positive charges are stationary, $v_+ = 0$, and the negative charges have a small negative drift velocity $v_- = -v_d$, where, to be perfectly explicit, v_d is the drift speed and the negative sign gives the direction of motion. In this case

$$i = 0 + q_-v_-$$

which is a positive current obtained by negative charges flowing in the negative direction. (What a pity that the charge obtained by rubbing glass with cat's fur was defined as positive and not the charge obtained by rubbing amber with silk!)

(f) We have now arrived at an equation which expresses the current as a charge density multiplied by a velocity, but the charge density is *not* q but q_-, and the velocity is not c but v_-.

(g) So to summarize this argument, a conductor posesses three charge densities (at least): q, q_- and q_+ (or surface charge or volume charge as appropriate) at every point, and in no case is the drift velocity given by i/q, where q is the net charge.

The point in *The Answer* where the charge-current relation appears explicitly is on page 6 and is shown here.

the two lines. Similarly the net current on the two lines (from this particular line of charges) is

$$i = 0 + \frac{(-e)(-v)}{a(1+v/c)} = \frac{ev}{a(1+v/c)} = cq$$

for the signal and

$$i = 0 + \frac{(-e)(v')}{a(1-v'/c)} = \frac{-ev'}{a(1-v'/c)} = cq'$$

on the return line. The drift velocity must be such

The first part of this equation is an expression for i in the form of equation (2), where the first zero is the current of the positive charges, and in the second term is the current of the negative charge density. The negative charge density

$$q_- = \frac{-e}{a(1+v/c)},$$

was derived at the bottom of page 5, and the velocity of $v_- = -v$ was defined at the top of page 6. However the net charge density was derived on page 5

$$q = \frac{e(v/c)}{a(1+v/c)}.$$

(see note above for the scribal error in this equation, which I have here corrected). Thus the expression for i is indeed qc, as given by the last term in the equation. However the moving charge density is not q, and the velocity is not c.

In *The Answer* this section of the document concludes with the words (emphasis added)

> *This crude model* illustrates how the creation of the current as the wave passes necessarily leads to a charge density in the correct ratio.

Thus the charge-current relation appears in the context of a crude model of the conduction process, but which nevertheless captures the essential feature of a charge density appearing in every section of the line as the wave passes, even though no individual charge moves at anything like c. The reference to the 'correct ratio', namely $q = i/c$ is thus slightly puzzling since this is the first explicit appearance of this relation. However it could just as easily have been written down at a much earlier and more significant point, at the end of the general description of the fields and currents of all two-wire transmission lines carrying a voltage signal propagating in the z-direction, $V(t - z/c)$, on page 3. The final part of this description is the explicit form for the surface charge densities σ and surface current densities in the z-direction k_z:

$$\sigma = \epsilon_0 E_n = -\epsilon_0 \frac{\partial \phi}{\partial x_n} V(t - z/c)$$

$$k_z = \frac{1}{\mu_0} B_t = \frac{-1}{\mu_0 c} \frac{\partial \phi}{\partial x_n} V(t - z/c)$$

In each of the two equations the first equality gives the boundary condition at a perfect conductor relating charges to electric fields and currents to magnetic fields, and the second equality gives the expicit form by substitution from results already given. It's obvious from this second equality that

$$k_z = \frac{1}{\epsilon_0 \mu_0 c} \sigma = \sigma c.$$

This is the same charge-current relation at the level of surface charges and currents; thus it will necessarily apply also to line charges and currents which are integrals of these around the circumference of the conductor. This is the charge-current relation in its most general form (at least for the air-spaced transmission lines considered here).

This derivation from the surface boundary conditions also answers the question 'If the charge-current relation does not — indeed *cannot* — mean charge density σ (or q) travelling at the speed of light, what does it mean?' It derives from the ratio of the field strengths of the **E** and **B**-fields. Higher up page 3 we find the equations for the fields:

$$\underline{E} = -\nabla_t \phi \; V(t - z/c) \qquad \underline{B} = -\frac{1}{c} \hat{\underline{k}} \times \nabla_t \phi \; V(t - z/c)$$

Again it's not stated explicitly, but $\mathbf{E} = c\mathbf{B}$.

2 Equations and their validity.

There is one other general topic I wish to discuss before attempting a review of the text of *The Forbidden Equation*, and it relates to the discussion of 'other forbidden equations' on pages 3–5 of Forrest's document. There are here eight quantities in play: in Forrest's notation they are i, v, q, ϕ_L, L_L, C_L, c and Z. (Forrest uses q_L for q in this section, but I will stick to q. He also introduces other quantities such as the impedance of free space, but these eight are sufficient for my present purpose.) Of these, four are fixed parameters characterizing the specific transmission line: L_L, C_L, c and Z. The other four are variables describing the signal as it propagates along the line.

2.1 The Fixed Parameters.

Only two of these are independent, and the other two can each be given in terms of the selected independent pair. However we can choose any pair. There are thus four relations between them, each containing three of the parameters and not the fourth.

The usual textbook choice is to take the capacitance per unit length C_L and the inductance per unit length L_L as the independent variables, and then give the other two in terms of them. (I shall give all equations with a dimensionless 1 on the right-hand side, as advocated by Forrest. I agree that this makes any algebraic manipulations particularly clear.) Thus most books give the equations

$$\boxed{c\sqrt{C_L L_L} = 1 \qquad \text{(A)}}$$

and

$$\boxed{Z\sqrt{\frac{C_L}{L_L}} = 1. \qquad \text{(B)}}$$

Note that both of these contain square roots, and thus a sign ambiguity. C_L and L_L are positive definite, but c and Z are what Forrest calls slippery variables, which can take either sign. These two cases are appropriate for signals travelling in the z-direction and the $-z$-direction: that is, if we take c to be the signal velocity not speed, and it could be negative, and when the signal propagates in the reverse direction, the current changes sign relative to the voltage, giving a negative Z. However we shall follow the usual convention and take both quantities as positive, and indicate the sign explicitly when required, just as we did above with drift velocity.

Note also that box (A) contains no Z, and box (B) no c. Thare are thus two more equations, containing no L_L and no C_L respectively. If we multiply equations (A) and (B) we obtain

$$\boxed{cZC_L = 1, \qquad \text{(C)}}$$

and if we divide (A) by (B) we obtain

$$\boxed{\frac{cL_L}{Z} = 1. \qquad \text{(D)}}$$

Forrest refers to these as forbidden, but they are just obvious consequences of the first two equations.

2.2 The Variables.

The four variables i, v, q and ϕ_L are all related to the signal propagating on the line and, in the special case of a signal propagating in the z-direction only, they are all proportional to each other. There are thus six equations relating all possible pairs of these variables, of which only three are independent (although not any three: they have to contain all four variables!). The usual textbook choices are to give the following three equations only. Firstly the definition of capacitance per unit length, the voltage-charge relation:

$$\boxed{\frac{C_L v}{q} = 1; \qquad \text{(E)}}$$

secondly the definition of inductance per unit length, the current-flux relation :

$$\boxed{\frac{L_L i}{\phi_L} = 1; \qquad \text{(F)}}$$

and the definition of characteristic impedance, the current-voltage relation :

$$\boxed{\frac{Zi}{v} = 1. \qquad \text{(G)}}$$

However by substituting from the previous equations, each of these can be given in two other forms. For example we can multiply (E) by either $(A)^2$ or $(B)^2$ to get the two alternative forms. So equations (E), (F) and (G) amount to another 9 equations.

By combining (E), (F), and (G) we can obtain the other thee relations. So (E) multiplied by (G) gives one version of the charge-current relation:

$$\boxed{\frac{ZC_L i}{q} = 1; \qquad \text{(H)}}$$

(F) divided by (G) gives the voltage-flux relation:

$$\boxed{\frac{L_L v}{Z\phi_L} = 1; \qquad \text{(I)}}$$

and (G) multiplied by $(B)^2$ and divided by (H) gives the charge-flux relation

$$\boxed{\frac{Zq}{\phi_L} = 1. \qquad \text{(J)}}$$

Again, each of these can be given in three different forms using the equations in boxes (A)–(D), so these amount to another 9 equations. Thus in total we have 22 equations. It is not surprising that textbooks give a sufficient set and no more, but all are valid and none forbidden.

2.3 Validity.

Of these 22 equations only some are generally true: the four equations (A)–(D) relating the fixed parameters, and equations (E) and (F) in their three equivalent forms. These apply in all possible circumstances. The remaining 12 equations are valid only for signals propagating in the z-direction. All of these depend on equation (G) of the minimal set normally given in textbooks, and, as mentioned above, this equation is only true in the forwards-propagating case. More generally, if the voltage on the transmission line contains a reflection from the far end, $V(z,t) = V_+(z-ct) + V_-(z+ct)$, the voltage-current relation becomes

$$I(z,t) = \frac{V_+(z-ct)}{Z} - \frac{V_-(z+ct)}{Z}.$$

and the ratio between the voltage and current becomes:

$$\frac{V}{I} = Z\,\frac{V_+(z-ct)+V_-(z+ct)}{V_+(z-ct)-V_-(z+ct)}.$$

This ratio can take any value depending on the particular forms of the unknown functions V_+ and V_-. Thus the voltage-current relation (G), in its three different forms, *is not generally true,* and consequently neither are equations (H), (I) and (J).

2.4 The Equation of Continuity.

Forrest refers to the charge-current relation throughout as the equation of continuity for electric current. This occurs in the opening paragraph, and the idea recurs over several pages. This is just wrong; it is not a continuity equation, and a bald assertion to the contrary does not change that. I was so puzzled by this as I read the paper that I began to wonder if Forrest meant something different by the expression 'the continuity equation'. But then on page 6 there occurs a perfectly satisfactory definition of the standard role of the continuity equation.

> The continuity equation is a staple of physics and engineering. It mathematically expresses the idea that all of the material in a given flow — a "current" — has to be accounted for at all times and places along that flow. Material cannot simply appear and disappear without any accounting.

This is exactly what the continuity equation needs to do, and $i = qc$ does not do this. Forrest also correctly notes on page 6 that the equation can be given in three different forms.

> There are three major types of density, often confused or not made clear in the literature: the linear-density, the areal-density, and the volume-density (line, area, volume). All three are used in various expositions on continuity.

These three versions are indeed correctly and clearly given in *The Answer*, on pages 1–2:

$$\frac{\partial q}{\partial t} + \frac{\partial I}{\partial x} = 0 \qquad (1)$$

$$\frac{\partial \sigma}{\partial t} + \frac{\partial K_x}{\partial x} + \frac{\partial K_y}{\partial y} = 0. \qquad (2)$$

$$\frac{\partial \rho}{\partial t} + \underline{\nabla} \cdot \underline{J} = 0. \qquad (3)$$

In the present context it is equation $\boxed{1}$ from this extract that is relevant (I'm changing the spatial co-ordinate from x to z as explained in the introduction):

$$\frac{\partial i}{\partial z} + \frac{\partial q}{\partial t} = 0.$$

This is the *only* generally valid equation relating i and q. If we substitute into this the charge-current relation under dicussion we obtain

$$c\frac{\partial q}{\partial z} + \frac{\partial q}{\partial t} = 0. \tag{3}$$

This equation goes by several names but perhaps the most heplful is the one-way wave equation. Its general solution is $q = Q(z - ct)$ where Q represents any differentiable function. Thus the condition for the validity of $i = qc$ is precisely that i and q consist solely of signals propagating in the z-direction. This is independent of the former argument, but arrives at the same conclusion. Above I showed that the assumption of a solely forwards-going signal is a pre-requisite for deriving the equation; here I show that assuming $i = qc$ and using it in a generally valid equation gives the same condition for its validity.

3 Summary

I shall end this very long preamble by a summary of the theses I have established.

(I) At all points in a conductor there are three charge densities q, q_+, q_- related by $q = q_+ + q_-$.

(II) Thus the net charge q cannot equal the mobile charge q_-, or the fixed charge would be zero which is absurd. In fact $q \ll q_-$ in all cases.

(III) Thus to identify q in $i = qc$ as the mobile charge carrying the current, and hence that charge travels at the speed of light, is erroneous.

(IV) The equation is really about the ratio of electric and magnetic fields in the propagating wave.

(V) There are many equations relating the variables in a transmission line; it is unsurprising that textbooks give a minimal set.

(VI) The equation $i = qc$ has only limited validity.

(VII) The equation $i = qc$ is not an equation of continuity.

(VIII) The equation is not forbiddden, just not very important.

4 The Review

With some reluctance I now turn to the review of *The Forbidden Equation*. I say this because basically I am a teacher; my desire is to impart knowledge, not to criticize others. However there is much to criticize in the document and I shall try to do this as economically as possible. In the following the roman numerals refer to the theses of section 3; the abbreviation NC denotes 'no comment'; it does not indicate approval.

Reference			Text	Comment
Page	Col.	Para.		
1	1	1	equation gone missing	V
1	1	1	the continuity equation	VI, VII
1	2	3	grave heresy . . blasphemy . . .toxic	VIII

Reference Page	Col.	Para.	Text	Comment
2	1	1		NC
2	1	2	q has to be massless	III
Section 2.1				Very interesting derivation.
Section 2.2				(E)×(G)/(C) from above; the simplest
Section 2.3			Equation (22)	This is (EF/A^2)
Section 2.3			Equation (27)	This is (EG/FB^2)
Section 2.3			Equation (28)	This is (EG/FB^2)G
Section 2.3			$i/qc = 1$	((EG/FB^2)(G)(EF/A^2). Since C=AB, and the F cancels,this is obviously equivalent to (EG/C)^2; 'the scenic route'
5	2	2	fails to satisfy continuity	Complete nonsense
6	1	1		I agree
6	1	3	general continuity equation	No, this is a definition
6	1	5	does not add . . . information	False: (36) is true for a current which varies with time; (37) is not
6	2	1		This paragraph essentially says that the current does not itself define the size of the conductor, or the mobile charge density or the drift velocity. That's obviously true: what's the problem?
6	2		equation (41)	This is *not* a continuity equation; it is an erroneous attempt at my equation (2)
6	2		continuity . . c for the speed	II,III,VII
6	2		Either . . . but not both	Both are true with different interpretations
7	1	1		NC; appears to be nonsense.
7	1	3	devoted to debunking	This is a radical misreading of *The Answer*.
7	1	4	rather than derive them	As shown, they were derived on p 3.
7	2	2	does not explain	It's substituted from the previous equation!!
7	2	2	$i = i$	The two i's are different: signal and return. Since i is defined with respect to a fixed co-ordinate system the current on the return line is negative. the rest of this
7	2			column is consequentially wrong.
7	2	3	Can't move negatively	Of course it can if the motion is defined with respect to a co-ordinate system, as it usually is in physics.
7	2	4	−ve charge can't cancel −ve	This is just false.
7	2	5		I think this paragraph is complete nonsense.
8	1	2	Since $i = i$	See above
8	1	3	−ve moving W . .+ve E	This is true — they are equivalent. But in a metal only one occurs.
8	1	4	would have to accelerate	Ohms law relates a current to a force. This is because it is a terminal velocity, reached after a finite but short time ($\sim 10^{-15}$ s).

Reference			Text	Comment
Page	Col.	Para.		
Section 4.1				Multiple confusions here, see below.
Section 5				NC
Section 6				NC
Section 7				V, VI, VIII
Section 8			Theory N, H	Maxwell's equations cannot distinguish these; causality is supervenient on the equations
13	1	4	Griffiths explains	I do not have access to Griffiths; NC
13	1	5	Palmer using Theory H	I use the surface boundary condition on the on the fields to diagnose the charges and currents in the surface; I make no claim regarding causality.
13	1	7	Conspiracy theory	I assume this is irony; NC
13	2			NC
Section 9				I think this is complete nonsense
Section 10				I have no access to Morgenthaler; NC

That just leaves a small number of issues calling for longer comments than the table above permits.

- I'm particularly interested in the Catt-Bishop derivation of the charge-current relation in section 2.1, because it shows that Ivor Catt does understand the essence of the divergence-of-current argument that is the centrepiece of *The Answer*. Thus all that he seems to lack is the realisation that the mobile charge is not to be identified with the net charge, and therefore the drift velocity can be vastly less than the speed of light. The West/South dichotomy is false: the charge is already present, it just needs to stretch out or compress a bit, due to the divergence, so that it becomes out of balance with the fixed charge density.

- Forrest claims the the first three pages of *The Answer* are devoted to dubunking the Southerner theory. This is a radical misreading, and leads him to ignore the central results on which my answer depends. So I here give a précis. Changes in charge density are the result of current divergence (p.1). Hence the answer to the Catt question depends on knowing the currents in the conductors (middle of p.2). These can be derived from the fields which satisfy Maxwell's equations and the boundary conditions (bottom of p.2). I then give the fields for a two-wire transmission line of arbitrary geometry (top of p. 3), and use them to diagnose the currents and charges using the boundary conditions (bottom of p.3). *In passing* I then note that since the currents all flow in the z-direction this approach has categorically disproved the Southerner theory. Although the charge-current relation was not written down explicitly on page 3, I was obviously aware of it as I then immediately go on to show that no charge travels at the speed of light, but the divergent current creates the unbalanced charge density locally, and that this is the answer to the Catt Question. Thus at this point I have completed the task I set on p. 1, and the document could have ended at this point:

The advancing wave causes currents to flow in a previously uncharged section of line. This spatially inhomogeneous current, which on the return line illustrated is <u>divergent</u> ($\partial k_z/\partial z < 0$), necessarily creates negative charge. Thus the Charge on the return conductor is created by separation of positive and negative charges already present on the surface of the conductor. This is the answer to the Catt Question.

(bottom of p. 4). The text then continues (and I discuss this more fully below) 'This is as far as electromagnetism, as a macroscopic continuum theory can take us'. However I then go on to give two explicit examples of density arising locally from a divergent current: the 'crude model' of conduction already discussed, and a line of stationary traffic starting to move.

- There is an underlying problem in *The Forbidden Equation* of ignoring the theoretical hierarchy. This is addressed explicitly in *The Answer*: on page 2 I note that I am using a specfific approximation (perfect conductors) because it seemed to me that this was the framework within which the Catt question had been posed. Within this approximation it is simply meaningless to discuss the force needed to drive the current (we have explicitly assumed that it is zero), and even more so to discuss the acceleration of the electrons. Even we treated the problem more exactly using Ohm's law, this is still a force balance equation which does not take into account the inertia of the electrons. In order to do that we would have to move beyond Ohm's law to a constitutive relation of the form

$$\tau\rho\frac{\mathrm{d}\mathbf{J}}{\mathrm{d}t} + \rho\mathbf{J} = \mathbf{E}$$

where τ is a time depending on the charge-to-mass ratio of the electron, on the order of 1 fs. For all times significantly longer than this the effect of this change is effectively zero, and therefore we continue to use Ohm's law even though this ignores the inertial properties of the electrons.

Another theoretical choice is the macroscopic continuum version of Maxwell's Equations. There is a microscopic form, in which individual charges, and currents due to their motion, enter directly. Because the number of charges is so large in bulk matter these equations are very difficult to use. In fact I have only ever seen them used in two contexts: atomic physics, where they are the equations used to describe the forces between the atomic particles, and in the derivation of the macroscopic equations by an averaging procedure. This derivation is given at great length and scholarly detail in the book by de Groot, and the later volume by de Groot and Suttorp, and in a truncated form in Jackson's *Classical Electrodynamics*. Once we have made this approximation any discussion of the motion of individual charges is meaningless — we are treating them all as a continuum. I am quite explicit about this in *The Answer*: that is the significance of the remark quoted above that this is as far as classical electromagnetism can take us. In order to discuss what the actual charges in a metal are doing I have to go beyond Maxwell's equations by adding a fact about the microscopic constitution of the metallic conductor. Within the framework of Maxwell's equations *any* specific model of charge motion that reproduces the actual average charge densities and currents given by the equations is equally valid or invalid. This includes the implied model behind *The Forbidden Equation* in which the neutral conductor is like an empty pipe and any charge has to come from somewhere else, like water in a pipe, but at the speed of light. The reason that this is *not* what actually happens isn't anything to do with Maxwell's equations, but with these auxiliary assumptions, which are simply false. So in fact the Catt question isn't even about electromagnetism at all: the details of the solution to the equations are agreed on all sides. The discussion is entirely about the non-electromagnetic ideas underlying different ways in which these charges and currents might occur in an actual metallic conductor.

Christopher W P Palmer
20 August 2020