

Why Don't We Teach The Helmholtz Theorem?

Conrad Schiff*

Chesapeake Section of AAPT

Date: Oct. 25, 2014

Abstract

The traditional way we teach electricity and magnetism is to trace through the historic development by examining a host of 18th and 19th century the experiments. The usual course is to introduce integral forms of the laws and then show how these can lead to Maxwells equation (e.g. from Coulomb's law to Gauss's law). In this paper, I present a different path that starts with Maxwell's equations and the Helmholtz theorem and produces the Coulomb and Biot-Savart laws. I believe this approach, which is nearly impossible to find in the usual textbooks, offers a clearer view into the unity of electricity and magnetism at the cost of some slightly more mature vector calculus. Even this shortcoming can be turned to an advantage as the techniques can be used to present modern ideas about generalized functions and signals and to sharpen the what the fields are really doing physically far away from the regions that we are investigating.

1 Introduction

Before discussing the physics, I want to start by framing the title question within a greater context. As stated, it is neither meant to be accusatory nor Socratic. My aim for this paper is to open (or perhaps reopen) a dialog on the pedagogy of using the Helmholtz theorem. Acknowledging that I am not trained in physics education nor am I an educator, the only credentials I can offer are as a student, but as one who actually remembers the struggles learning electromagnetic field theory.

Central to the viewpoint I would like to promulgate is that the current approach, which I will refer to as the traditional approach hereafter, presented in most commonly-used textbooks [1, 2, 3, 4, 5, 6], obscures the unity of the Maxwell equations. It does so by presenting an analysis of static field configurations first, layering on time dependence only much later, and then finally culminating in Maxwell's inspired addition of the displacement current to the Ampere law. By this time, a student is likely bored, confused, or overwhelmed and can't or won't appreciate all the

*Electronic address: gravmath@yahoo.com; Corresponding author

results that follow. To such a student is lost the wonder of realizing that the fields can take on a life of their own, independent of the things that created them, and can radiate outwards; that they can reflect and refract (i.e. optics as an inherently electromagnetic phenomenon); and that they can be generated and controlled at will to form the communication network we all use on a daily basis. Considering the ubiquitousness of cell-phones, tablets, and computers all connected to the web usually by 3G or 4G wireless communications the current textbook approach seems to be asking the student to step back in time to when the larger questions concerned condensers and Voltaic piles and not download speeds and the number of bars a user has.

An alternative approach presented by Solymar [7] is to state the Maxwell equations up front and then to interrogate them mathematically to ‘discover’ a host of new phenomena. This approach has the advantage that the fields are introduced early and the equations they satisfy are complete. There is no unlearning facts later on (e.g. $\nabla \times \vec{E} \neq 0$ generally) and the structure that leads to our modern wireless world is fully available. The one serious flaw in Solymar’s implementation, however, is that there is no argument as to why knowing the divergence and curl of a vector field is all that is needed to uniquely specify the field. After all, why can’t a uniform field be added as a constant of integration? Another flaw is that the Coulomb law is ‘derived’ by assuming spherical symmetry.

In the remainder of this paper, I present a modification to the approach presented by Solymar that employs the Helmholtz theorem. This modification addresses the identified deficiencies and also brings in new pedagogical opportunities to present modern concepts. To make the advantages of this modification clearer, I start with a brief summary of the traditional approach in Sec. 2. The statement of the Helmholtz theorem and an outline of its proof are presented in Sec. 3. Sec. 4 discusses the bulk of the physics and relevant pedagogy while Sec. 5 discusses the open questions that remain in using the Helmholtz theorem and provides a critique of the method.

2 Traditional Approach

At the heart of the traditional approach is the idea of a force field as a physically real object and not just a useful mathematical construct. The prototype example is the electric field, which comes from the experimental expression for the Coulomb law stating that the force on charge q_2 due to q_1 is given by (note all equations are expressed in SI units):

$$\vec{F}_{21} = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2 (\vec{r}_2 - \vec{r}_1)}{|\vec{r}_2 - \vec{r}_1|^3}. \quad (1)$$

By allowing one of the charges, say q_2 , to become a test charge whose strength is so small that its presence does not alter the behavior of q_1 , the electric field at any point \vec{r} where the test charge may be placed is deduced as

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \frac{q_1 (\vec{r} - \vec{r}_1)}{|\vec{r} - \vec{r}_1|^3}. \quad (2)$$

A generalization of this expression to an arbitrary charge distribution within a volume V then leads to Coulomb's law

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_V d^3r' \frac{\rho(\vec{r}') (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3}. \quad (3)$$

At this point in the traditional approach, concepts of vector field divergence and curl are trotted out. Although different texts order the content differently, they all present the vanishing curl of the (static) electric field nearby discussions of the electric flux and Gauss's law for electricity. The vanishing curl of the electric field implies the existence of a conservative field, from which follows the definition of the electrostatic potential ϕ and the Laplace/Poisson equation. The introduction of electric flux and the divergence theorem lead to Gauss's law, the first of the Maxwell equations,

$$\nabla \cdot \vec{E}(\vec{r}) = \rho(\vec{r})/\epsilon_0. \quad (4)$$

Gauss's law is then used to determine the electric field of some simple (and often unmotivated) charge distributions like a spherical shell or an infinite line or sheet of charge.

There are several problems with the above pedagogy. First, each text will be careful to point out that Gauss's law, while holding in general, can really only be applied in situations of great symmetry. The message here is that somehow Coulomb's law is more general; an impression that usually emphasized by the fact that the majority of homework problems focus determining the electric field from Eq. 3. Coulomb's law has been shown to imply Gauss's law, but the other direction is not as clear with one text even saying "Thus we have deduced Coulomb's law from Gauss's law and considerations of symmetry." [1] The second and much more important problem is that the curl and divergence of the electric field are not linked. The vanishing curl of the electric field is not married in any way to the divergence as the measure of the charge density.

As the traditional program proceeds, magnetostatics follows with an analogous discussion about the magnetic field. This time the vanishing of the divergence is used to find the vector potential and the curl is related to the current via Ampere's law

$$\nabla \times \vec{B}(\vec{r}) = \mu_0 \vec{J}(\vec{r}). \quad (5)$$

As in the electrostatic case, Ampere's law is applied only to situations with simple symmetry. The Biot-Savart law

$$\vec{B}(\vec{r}) = \frac{\mu_0}{4\pi} \int_V d^3r' \frac{\vec{J}(\vec{r}') \times (\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3}, \quad (6)$$

gives the magnetic field for a given current density within a volume V . Again it appears to be more general than Ampere's law, with the logical arrow of deduction clearly pointing from the latter to the former while converse is only tenuously seen.

By the time time-varying fields are introduced and the full Maxwell equations are on display, the linkage between the different facets of each field is highly obscured, and the basic underpinning of the theory — that the divergence and curl tells all there is to know about a field — is not to be found. The pedagogy seems to suffer from too many unconnected facts with no common framework by which to relate them.

3 Using the Helmholtz Theorem

The various shortcomings identified in the previous section can be addressed by assuming the Maxwell equations as given and by employing the Helmholtz theorem to make sense of their content. There are several slightly different discussions on the Helmholtz theorem in the literature [8, 9, 10], but in essence they say that a sufficiently well-behaved vector field can always be uniquely specified by its divergence and curl in some volume V and, in the case of a finite region, by its normal components on the boundary. For now, the region in question will be taken to be all space and the vector fields in question will be assumed to fall off fast enough that the normal components can be ignored. This point will be revisited in the Sec. 5 later. Mathematically, the decomposition takes the form [9]

$$\vec{F}(\vec{r}) = -\nabla U(\vec{r}) + \nabla \times \vec{W}(\vec{r}), \quad (7)$$

where the fields $U(\vec{r})$ and $\vec{W}(\vec{r})$ are given by

$$U(\vec{r}) = \frac{1}{4\pi} \int d^3 r' \frac{\nabla' \cdot \vec{F}(\vec{r}')}{|\vec{r} - \vec{r}'|}, \quad (8)$$

and

$$\vec{W}(\vec{r}) = \frac{1}{4\pi} \int d^3 r' \frac{\nabla' \times \vec{F}(\vec{r}')}{|\vec{r} - \vec{r}'|}, \quad (9)$$

and where ∇' denotes that the derivatives are to be taken with respect to the source points \vec{r}' .

3.1 Outline of the derivation

There are many ways to derive the Helmholtz theorem but the way that provides the most pedagogy is to start with the following representation of the delta-function in three dimensions

$$\nabla^2 \left(\frac{1}{|\vec{r} - \vec{r}'|} \right) = -4\pi \delta(\vec{r} - \vec{r}'). \quad (10)$$

Next use the identity

$$\vec{F}(\vec{r}) = \int_V d^3 r' \delta(\vec{r} - \vec{r}') \vec{F}(\vec{r}') \quad (11)$$

for an arbitrary vector field $\vec{F}(\vec{r})$ over a given volume V . Using the explicit representation of the delta-function stated above and factoring out the derivatives with respect to the field point \vec{r} yields

$$\vec{F}(\vec{r}) = \frac{-\nabla^2}{4\pi} \int_V d^3 r' \frac{\vec{F}(\vec{r}')}{|\vec{r} - \vec{r}'|}. \quad (12)$$

Now apply the vector identity

$$\nabla^2 = \nabla(\nabla \cdot) - \nabla \times (\nabla \times). \quad (13)$$

Doing so allows the expression for \vec{F} to take the form

$$\vec{F}(\vec{r}) = \frac{1}{4\pi} \nabla \times \vec{I}_v - \frac{1}{4\pi} \nabla I_s, \quad (14)$$

where the integrals

$$\vec{I}_v = \nabla \times \int_V d^3r' \frac{\vec{F}(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (15)$$

and

$$I_s = \nabla \cdot \int_V d^3r' \frac{\vec{F}(\vec{r}')}{|\vec{r} - \vec{r}'|} . \quad (16)$$

The strategy for handling these terms is to:

1. bring the derivative operator with respect to \mathbf{r} into the integral,
2. switch the derivative from \mathbf{r} to \mathbf{r}' with a cost of a minus sign,
3. integrate by parts,
4. apply the appropriate boundary conditions and boundary integral version of the divergence theorem to the total derivative piece.

Application of this strategy to the vector integral, \vec{I}_v , gives

$$\vec{I}_v = \int_V d^3r' \frac{\nabla_{\vec{r}'} \times \vec{F}(\vec{r}')}{|\vec{r} - \vec{r}'|} - \int_{\partial V} dS \frac{\hat{n} \times \vec{F}(\vec{r}')}{|\vec{r} - \vec{r}'|} , \quad (17)$$

where ∂V is the bounding surface of the volume in question and \hat{n} is the corresponding outward normal.

Likewise, the same strategy applied to the scalar integral, I_s , gives

$$I_s = \int_V d^3r' \frac{\nabla \cdot \vec{F}(\vec{r}')}{|\vec{r} - \vec{r}'|} - \int_{\partial V} dS \frac{\hat{n} \cdot \vec{F}(\vec{r}')}{|\vec{r} - \vec{r}'|} . \quad (18)$$

Allowing the bounding surfaces to recede to infinity eliminates the surface terms and Eq. 14 becomes Eq. 7 with $\vec{I}_v = 4\pi\vec{W}$ and $I_s = 4\pi U$.

4 Pedagogical Advantages

Using the Helmholtz Theorem in conjunction with a upfront statement of the Maxwell equations offers several advantages in teaching electromagnetism. I will content myself with listing three that I think are particularly useful.

The first is already manifest in the statement of the theorem itself. From the structure of Eq. 7 and the form of the functions $U(\vec{r})$ and $\vec{W}(\vec{r})$ in Eqs. 8 and 9, respectively, the student can clearly see that knowledge of the divergence and curl of a sufficiently well-behaved vector field uniquely and completely characterizes it. This observation serves to motivate why the Maxwell equations are all that are needed to solve for \vec{E} and \vec{B} . No ‘constant of integration’ vector field is needed.

The second advantage is that the derivation presented above gives further exposure to the properties of the delta-function. Not only is Eq. 10 a useful identity in its own right, it is also at the heart of Gauss’s law being a very simple and approachable form of a Green’s function. In addition, it is important to undo some of the rigidity imposed in introductory mathematics education. Students learn early that only analytic functions are valid (although they aren’t taught it explicitly that way). But as nicely

discussed by Penrose [11] in his chapter on Fourier analysis and distributions, one must be willing to allow objects like the delta-function if one wants to be able to “model signals which can transmit ‘unexpected’ (non-analytic) messages.” Such non-analytic functions form the backbone of the modern information- and message-driven economy.

The third advantage is the ‘derivation’ of the Coulomb and Biot-Savart laws from Maxwell’s equations. Start by considering the Maxwell equations, presented here in vacuum, as

$$\nabla \cdot \vec{E}(\vec{r}, t) = \rho(\vec{r}, t)/\epsilon_0, \quad (19a)$$

$$\nabla \cdot \vec{B}(\vec{r}, t) = 0, \quad (19b)$$

$$\nabla \times \vec{E}(\vec{r}, t) = -\frac{\partial \vec{B}(\vec{r}, t)}{\partial t}, \quad (19c)$$

$$\nabla \times \vec{B}(\vec{r}, t) = \mu_0 \vec{J}(\vec{r}, t) + \epsilon_0 \mu_0 \frac{\partial \vec{E}(\vec{r}, t)}{\partial t}. \quad (19d)$$

Since the Coulomb and Biot-Savart laws are in the domain of the electro- and magnetostatics, all terms in Eqs. 19 involving time derivatives are set equal to zero to give

$$\nabla \cdot \vec{E}(\vec{r}) = \rho(\vec{r})/\epsilon_0, \quad (20a)$$

$$\nabla \cdot \vec{B}(\vec{r}) = 0, \quad (20b)$$

$$\nabla \times \vec{E}(\vec{r}) = 0, \quad (20c)$$

$$\nabla \times \vec{B}(\vec{r}) = \mu_0 \vec{J}(\vec{r}). \quad (20d)$$

Now substituting Eqs. 20a and 20c into Eqs. 8 and Eqs. 9 yields

$$U(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_V d^3 r' \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (21)$$

and

$$\vec{W}(\vec{r}) = \vec{0}. \quad (22)$$

Eq. 21 is clearly identified as the electrostatic potential and upon substitution into Eq. 7 yields the familiar form of the Coulomb law found in Eq. 3. In a similar fashion, substitution of Eqs. 20b and 20d into $U(\vec{r})$ and $\vec{W}(\vec{r})$ and then those expressions into Eq. 7 yields the Biot-Savart law found in Eq. 6.

5 Critique

The approach given above is not without difficulties, none of which are insurmountable but are worth discussion.

The first difficulty is the treatment of the boundary values of the fields at infinity. There is an implicit idea of the field being well-behaved at infinity that is used to set the surface integrals in Eqs. 15 and 16 to zero. The exact details of what this means are found in [9] and boil down to the conditions that both the divergence and curl of the field vanish faster than r^{-1} . This statement has obvious physical implications on the

sources of the field but discussing this in a classroom setting can be tricky since integrals tend to be intimidating and the discussion of convergence properties is usually glossed over in physics classes. Refinement on this point is vital as no discussion of the Maxwell equations is complete without specification of the behavior of the fields at a boundary.

The second difficulty is more pronounced. It turns out the form of the Helmholtz equation as discussed above is valid only for time-invariant fields. A more complicated form, involving integrals with retarded time can be derived [12, 13] which allows the recovery of the laws of Faraday and Ampere in addition to those of Coulomb and Biot-Savart. It isn't clear that such a one-stop-shop approach will be meaningful for any except the most advance students since the basic concept of retarded time is subtle and thus is usually relegated to the very end of a traditional program or skipped entirely. One the other hand, not discussing it at all presents a situation in which the student has to unlearn something at a latter point. A reasonable compromise is pointed out by Heras [13] in which the student can be told that the divergence and curl uniquely specify a vector field via the Helmholtz theorem, that there exists time-independent and time-varying forms, and then apply the time-independent form as above. How students may react to such a program remains to be seen.

6 Conclusion

An alternative to the traditional teaching of electromagnetism has been presented in which the Maxwell equations, in their entirety, are used as a starting point. The Helmholtz theorem, which demonstrates how a vector field can be uniquely constructed once its divergence and curl are known, provides a unified framework for understanding why the Maxwell equations related charge and current sources to derivatives of the electric and magnetic fields. It also allows the more traditional equations of Coulomb and Biot-Savart to be derived thus making a connection with experiment. Although not without its short-comings, this approach seems to be superior in its pedagogy to the traditional approach and I hope that this paper helps to contribute to a dialog on its use.

References

- [1] D. Halliday and R. Resnick, *Physics, Parts 1 & 2 Combined, 3rd Edition*, John Wiley & Sons, New York, 1978
- [2] J. R. Reitz, F. J. Milford, and R. W. Christy, *Foundations of Electromagnetic Theory, 4th Edition*, Addison-Wesley, Reading, 1993
- [3] J. D. Jackson, *Classical Electrodynamics, 2nd Edition*, John Wiley & Sons, New York, 1975
- [4] D. R. Frankl, *Electromagnetic Theory*, Prentice-Hall, Englewood Cliffs, New Jersey, 1986
- [5] J. C. Slater and N. H. Frank, *Electromagnetism*, Dover Publications, New York, 1969

- [6] D. J. Griffiths, *Introduction to Electrodynamics, 4th Edition*, Addison-Wesley, New York, 2012
- [7] L. Solymar, *Lectures on Electromagnetic Theory*, Oxford University Press, Oxford, 1986
- [8] G. Arfken, *Mathematical Methods for Physicists, 3rd Edition*, Academic Press, Orlando 1985
- [9] B. P. Miller, “Interpretations from Helmholtz’s theorem in classical electromagnetism”, *Am. J. Phys.* 52 , 948 (1984)
- [10] R. Mirman, “Maxwell’s Equations and Helmholtz’s Theorem”, *Am. J. Phys.* 33, 503 (1965)
- [11] R. Penrose, *The Road to Reality*, Alfred A. Knopf, New York, 2005
- [12] A. M. Davis, “A generalized Helmholtz theorem for time-varying vector fields”, *Am. J. Phys.* 74, 72 (2006)
- [13] J. A. Heras, “Comment on A generalized Helmholtz theorem for time-varying vector fields, by Artice M. Davis [*Am. J. Phys.*74, 7276 (2006)]”, *Am. J. Phys.* 74, 743 (2006)