What do ''voltmeters'' measure?: Faraday's law in a multiply connected region

Robert H. Romer

of antibaryons. It is plausible to assume that there is an excess of baryons over antibaryons throughout the entire universe, as discussed in

290. "Observational Tests of Antimatter Cosmologies," G. Steigman, Ann. Rev. Astron. Astrophys. **14**, 339–372 (1976). (A)

It was first pointed out in Ref. 230 that three conditions are necessary to generate a baryon excess from an initially symmetric big bang cosmology: (1) baryon number nonconserving interactions; (2) violation of C and CP (i.e., charge conjugation and charge conjugation–parity) symmetry; and (3) lack of thermal equilibrium.

Interest in generating the baryon excess was revived eleven years later by the article

*291. "Unified Gauge Theories and the Baryon Number of the Universe," M. Yoshimura, Phys. Rev. Lett. **41**, 281–284 (1978). (A)

This question is also treated in

*292. "The Cosmic Asymmetry Between Matter and Antimatter," F. Wilczek, Sci. Am. **243** (6), 82–102 (1980). (E)

293. "Baryon Number of the Universe," S. Dimopoulos and L. Susskind, Phys. Rev. D **18**, 4500–4509 (1978). (A)

294. "Matter–Antimatter Accounting, Thermodynamics, and Black-Hole Radiation," D. Toussaint, S. B. Treiman, F. Wilczek and A. Zee, Phys. Rev. D **19**, 1036–1045 (1979). (A)

295. "Cosmological Production of Baryons," S. Weinberg, Phys. Rev. Lett. **42**, 850–853 (1979). (A)

A short review survey of cosmological baryon number generation appears in

296. "Cosmological Baryon Number Generation in Grand Unified Models," J. A. Harvey, E. W. Kolb, D. B. Reiss, and S. Wolfram, Caltech preprint, CALT, 68-815 (unpublished). (A)

Some discussion of baryon asymmetry appears in most reviews on grand unified models, for example:

297. "Grand Unification and Cosmology," J. Ellis, M. K. Gaillard, and D. V. Nanopoulos, in *Unification of the Fundamental Particle Interactions 1980*, edited by S. Ferrara, J. Ellis, and P. Van Nieuwenhuizen (Plenum, New York, 1980), pp. 461–493. (A)

ACKNOWLEDGMENTS

What do "voltmeters" measure?: Faraday's law in a multiply connected region

Robert H. Romer

Physics Department, Amherst College, Amherst, Massachusetts 01002

A long solenoid carrying a varying current produces a time-dependent magnetic field and induces electric fields, even in the region exterior to the solenoid where $\partial B/\partial t$ and therefore curl E vanish. By paying attention to (a) what it is that a "voltmeter" measures and (b) the simplest properties of line integrals (e.g., under what circumstances the line integral of E is path independent), it is easy to use Faraday's law to predict the readings of voltmeters connected to various points in a circuit external to the solenoid. These predicted meter readings at first seem puzzling and paradoxical; in particular, two identical voltmeters, both connected to the same two points in the circuit, will not show identical readings. These theoretical predictions are confirmed by simple experiments.

I. INTRODUCTION

The physical problem described below is of great value in illustrating some of the most important properties of vector fields in general and of electric and magnetic fields in particular. It has a solution that is by no means obvious yet one that can be correctly derived with little more than Faraday's law,

$$\text{curl } E = -\partial B/\partial t. \tag{1}$$

Moreover, the result can readily be verified experimentally.

Consider a solenoid whose axis is perpendicular to the plane of the paper (Fig. 1). We will be concerned with currents and electric fields in or near the plane of the paper that result from a time-dependent current through the solenoid windings. Assume that the solenoid is extremely long and of very tight pitch, and that the plane of the paper is not close to either end of the solenoid. Thus the magnetic field produced is zero except within the interior of the solenoid. Suppose, for simplicity, that the current through the solenoid windings is a linear function of time. Let Φ denote the resulting flux of B through the cross section of the solenoid, and take the positive direction for calculating Φ to be into the paper. To be specific, let

$$\Phi = \alpha t. \tag{2}$$

(Thus we anticipate that if $\alpha > 0$, a current will flow in a general counterclockwise sense around any conducting path which encloses the solenoid.) Around the solenoid are
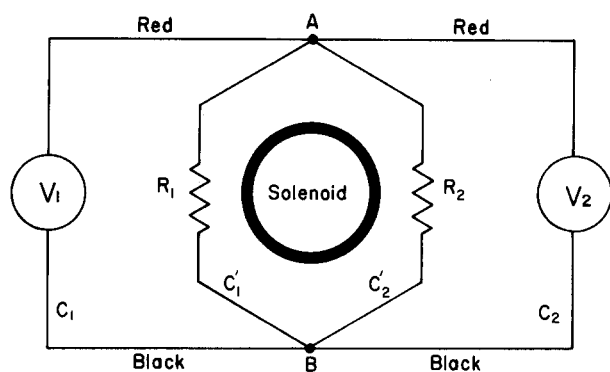
Fig. 1. Long solenoid, perpendicular to the plane of the paper, carries a current which varies linearly with time. We wish to predict the readings of the two voltmeters, V_1 and V_2. In the text, use is made of four different nonclosed paths, all of which have end points at A and B. Path C_1 passes through V_1, C_1' through R_1, C_2' through R_2, and C_2 through V_2.

located two resistors and two identical voltmeters, as shown in Fig. 1. The meters are high-resistance voltmeters which draw negligible current. Let V_1 and V_2 denote the readings of the two meters. The question is simply this: What will be the values of V_1 and V_2?

The *signs* of V_1 and V_2 constitute an especially interesting aspect of this problem; we must, therefore, be careful to distinguish between the two leads of a voltmeter. The use of labels such as " + " and " — " to make this distinction is conventional but can easily lead to confusion. Let us suppose that, in accord with common practice, the leads connected to the " + " and " — " terminals are red and black, respectively, and let us use these colors to label the terminals of the meters.

Lest the problem seem too easy, let me mention some erroneous responses which are often made when the problem is first brought up. "Since the two meters are both connected with their red leads to A and their black leads to B, they are 'both measuring the same thing,' so it is obvious that $V_1 = V_2$." A more sophisticated version is the following. "Perhaps $V_1 \neq V_2$ if $R_1 \neq R_2$, because the situation is then nonsymmetric, but surely if $R_1 = R_2$, symmetry requires that V_1 must be equal to V_2." As I will show, V_1 and V_2 are always of opposite sign and hence are never equal, though they are of equal magnitude $(V_1 = -V_2)$ if $R_1 = R_2$. This is a problem that yields quickly to clear thinking about issues such as what it is that a "voltmeter" really measures and attention to getting *signs* right; fuzzy thinking and uncritical use of concepts such as potential and voltage, or the asking of poorly defined questions such as "Where is the emf located?," are unlikely to produce the correct solution except by chance.

Klein[1] has described a simple apparatus for demonstrating the effects discussed here, but his discussion of the theory is presented in terms of the "internal resistance of the generator" and the "external load," notions which are obscure in the present context. Shadowitz[2] gives a treatment that is slightly more convincing. Unfortunately, it is preceded by a discussion of a simple dc circuit (in which $\partial \mathbf{B}/\partial t$ is of course zero), which is inconsistent with Eq. (1). (That is, he implies that curl $\mathbf{E} \neq 0$, which cannot be correct if $\partial \mathbf{B}/\partial t = 0$.) In Shadowitz's initial discussion of Faraday's law, he makes an artificial division of the electric field into conservative and nonconservative parts, a distinction not likely to be respected by real meters, which respond to the total

E whatever it may be. In his discussion of the situation discussed in this paper, he unfortunately uses the word "voltage," a term that should be carefully defined if it is to be used at all and one that would be safer not to use here, because of the obvious importance of the region inside the solenoid in which curl $\mathbf{E} \neq 0$. Moreover, the interesting question of the *signs* of the voltmeter readings is given short shrift by Shadowitz and totally ignored by Klein. Moorcroft[3] has also given an interesting discussion of this problem, but he, too, divides the field into conservative and nonconservative parts. Such an attempt to divide the electric field is unnecessary, as is the use of secondary concepts such as "voltage," "emf," or "internal resistance." Similarly, though surface charges on the various conductors contribute to the electric field, the distribution of these charges is an issue that we need not explicitly deal with if we wish simply to predict the voltmeter readings. I will discuss the problem from an extremely simple point of view and show how the solution follows in a straightforward fashion from Maxwell's equations, the geometry and topology of the physical situation, and the most elementary properties of vector fields and their line integrals.

II. PROPERTIES OF THE ELECTRIC FIELD OUTSIDE THE SOLENOID

We will be especially concerned with various line integrals of $\mathbf{E}$; it is important to be careful to specify in each case (a) the *direction* in which the line integral is to be calculated (to get the *sign* right) and (b) the *path* of integration, a matter of great importance when curl $\mathbf{E}$ does not vanish at all points in space. Let C denote a path joining two particular points (say A and B). The direction in which C is traversed in calculating the line integral will be indicated by the order of the limits. Thus the symbol

$$\int_{A}^{B} \mathbf{E} \cdot \mathbf{dr}$$
$$C$$

denotes the line integral of $\mathbf{E}$ along the path C, starting at A and ending at B. That is, in the infinite sum that this integral represents, each displacement vector $\mathbf{dr}$ points along C in the direction from A to B. Note, incidentally, that no matter what the curve C or the properties of the vector field, reversing the direction of integration always changes the sign of the line integral.

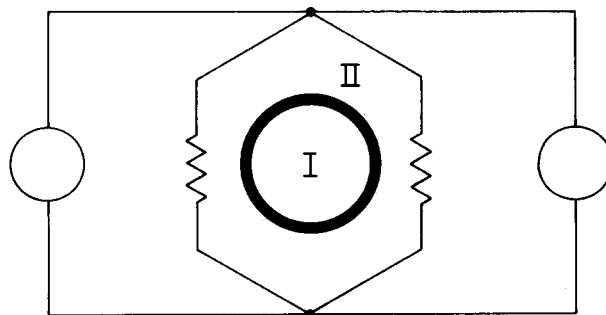We divide the plane of the paper into two regions (Fig. 2).

Fig. 2. We divide space into two regions. It is only in region I (the interior of the solenoid) that $\partial \mathbf{B}/\partial t$, and hence curl $\mathbf{E}$, are nonzero; it is in the multiply connected region II (everything exterior to the solenoid) that all observations are made.

Region I is the interior of the solenoid and region II is everything exterior to the solenoid. By virtue of the simplifying assumption expressed by Eq. (2), any induced currents and electric fields will be time independent, after an initial transient which rapidly decays in a time determined by the self-inductance and total resistance of the loop formed by R_1 and R_2. (We restrict our attention to the steady state. If the loop is superconducting, however, the "transient" never decays; in this case, the results of this paper are not applicable.) Outside the solenoid, $\mathbf{B} = 0$ except for the field produced by the time-independent current in the external loop; thus curl $\mathbf{E} = -\partial\mathbf{B}/\partial t$ vanishes throughout region II.

Because curl $\mathbf{E} = 0$ everywhere in region II, it is tempting to conclude that $\mathbf{E}$ is a conservative field in this region, that the line integral of $\mathbf{E}$ between any two points in region II is path independent. This is not true, because of the topology of region II. Even though curl $\mathbf{E}$ vanishes throughout region II, this region is not simply connected and thus $\int\mathbf{E}\cdot d\mathbf{r}$ is not necessarily path independent. [Recall that if a field $\mathbf{E}$ is irrotational (i.e., has vanishing curl) throughout a *simply* connected region, then the field is conservative in that region. That is, we can define a single-valued scalar function ϕ such that $\mathbf{E} = -\operatorname{grad}\phi$, line integrals of $\mathbf{E}$ are path independent, and any closed line integral of $\mathbf{E}$ is equal to zero.] If, in the situation shown in Fig. 1, we were to choose the "potential" to be zero at B, and then define the meaning of "the potential at A" in the usual way as

$$\phi_A = -\int_B^A \mathbf{E}\cdot d\mathbf{r},$$

the result would be ambiguous, for this line integral depends on the path. Thus it is not possible to define a single-valued scalar potential for $\mathbf{E}$, and $\mathbf{E}$ is not a conservative field. However, $\mathbf{E}$ does have properties that are just as useful for our purposes, properties that we will call "pseudo-conservative." What we *can* say about line integrals of $\mathbf{E}$ is the following. By application of Stokes's theorem and Eqs. (1) and (2), it is easy to see that $\oint\mathbf{E}\cdot d\mathbf{r}$ along a closed path following the border between regions I and II in the counterclockwise direction is just α. Elementary arguments then show that for any closed path lying entirely in region II, the line integral of $\mathbf{E}$ can have only one of a *discrete* set of values

$$\oint \mathbf{E}\cdot d\mathbf{r} = n\alpha \quad n = 0, \pm 1, \pm 2,..., \tag{3}$$

where $n = 0$ if the path does not enclose region I, $n = 1$ if the path goes around region I once in the counterclockwise direction, $n = -2$ if the path goes twice around region I in the clockwise direction, and so on (see Fig. 3.). It follows that the line integral of $\mathbf{E}$ between two points, say A and B, is "pseudo-path-independent" in the following sense. Let C and D be any two paths with end points at A and B, both paths lying completely in region II. Then

$$\int_{C\ A}^{\ B} \mathbf{E}\cdot d\mathbf{r} = \int_{D\ A}^{\ B} \mathbf{E}\cdot d\mathbf{r}, \tag{4}$$

provided C and *D* together form a closed path that does not enclose region I.

Although it is not, strictly speaking, *necessary* to use the terms "simply connected" and "multiply connected" in describing the properties of the fields, those who have encountered these concepts in pure mathematics may find it
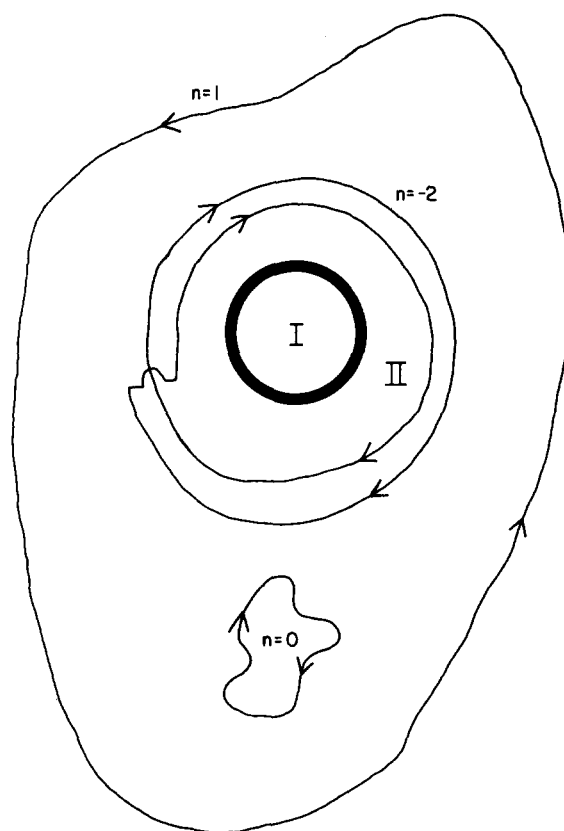

Fig. 3. For these three closed paths, the closed line integral of $\mathbf{E}$ in the indicated directions has, according to Eq. (3), the values α, -2α, and 0.

esthetically pleasing to see the possibility of using them to give an elegant description of a real physical situation.

III. WHAT DOES A "VOLTMETER" MEASURE?

Even in situations like these where "voltage" and "potential" are concepts of dubious value, a voltmeter measures *something*; we must ask exactly what it is that can be inferred from a voltmeter reading, in this application and in more ordinary ones as well. A voltmeter (whether a conventional indicating meter or an oscilloscope) is most often an ohmic device, usually of high resistance, which gives an indication (a deflection of a meter needle or of an electron beam) proportional to the (small) current that passes through it. A little thought convinces one that the voltmeter reading (call it V) is equal to the line integral of $\mathbf{E}$, $\int\mathbf{E}\cdot d\mathbf{r}$, where the path of integration passes *through the meter*, beginning at the *red* (or " $+$ ") lead and ending at the *black* (or " $-$ ") lead. (I have been unable to think of any device that is normally considered to be a voltmeter, whether a real device or an imaginary one, one with high resistance or low, or even nonohmic in character, for which it is not the case that the reading is proportional to the line integral just described. Even the imaginary miniature worker who implements the common thought-experiment definition of "potential" by carrying a tiny test charge from one point in space to another, measuring the work required, is just measuring the line integral of $\mathbf{E}$ along whatever path is chosen.)

To see how this works out in a more elementary context, consider the simple dc circuit shown in Fig. 4. The paths F and G both join the points P_1 and P_2, F passing through the
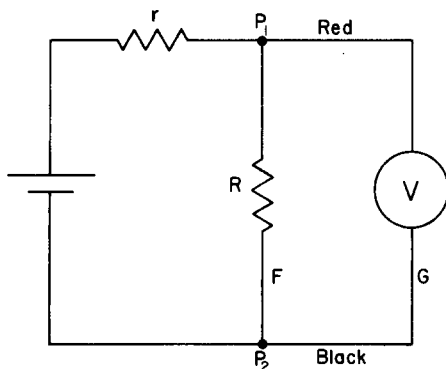
Fig. 4. Simple battery-operated circuit.

resistor R, and G passing through the voltmeter. The voltmeter reading V is equal to the line integral of $\mathbf{E}$ through the meter. In this case, $\partial \mathbf{B}/\partial t = 0$ everywhere, so curl $\mathbf{E}$ vanishes *everywhere*, and thus $\mathbf{E}$ is a genuinely conservative field, and line integrals of $\mathbf{E}$ are path independent. Therefore,

$$V = \int_{\substack{P_1 \\ G}}^{P_2} \mathbf{E} \cdot d\mathbf{r} = \int_{\substack{P_1 \\ F}}^{P_2} \mathbf{E} \cdot d\mathbf{r}, \tag{5}$$

and we conclude that in *this* circuit, V is equal to the line integral of $\mathbf{E}$ through the resistor to whose ends it is attached. I reiterate that this last statement is not obvious. A voltmeter measures the line integral of $\mathbf{E}$ through the meter itself; one must not equate this line integral with a different line integral without justification. (In this connection, see the remarks about inductors in Sec. VI.)

It follows immediately, from the discussion of what it is that a "voltmeter" measures, that in the present problem,

$$V_1 = \int_{\substack{A \\ C_1}}^{B} \mathbf{E} \cdot d\mathbf{r}; \quad V_2 = \int_{\substack{A \\ C_2}}^{B} \mathbf{E} \cdot d\mathbf{r}. \tag{6}$$

IV. SOLUTION OF THE PROBLEM

One relationship between V_1 and V_2 is obtained from Eq. (6):

$$V_1 - V_2 = \int_{\substack{A \\ C_1}}^{B} \mathbf{E} \cdot d\mathbf{r} - \int_{\substack{A \\ C_2}}^{B} \mathbf{E} \cdot d\mathbf{r}$$

$$= \int_{\substack{A \\ C_1}}^{B} \mathbf{E} \cdot d\mathbf{r} + \int_{\substack{B \\ C_2}}^{A} \mathbf{E} \cdot d\mathbf{r}. \tag{7}$$

The right-hand side of Eq. (7) is equal to the line integral of $\mathbf{E}$ along a closed path going once around the solenoid in the counterclockwise direction, and so, by Eq. (3),

$$V_1 - V_2 = \alpha. \tag{8}$$

Another relationship between V_1 and V_2 is provided by invoking Ohm's law together with the assumption that the voltmeters draw negligible currents. The latter assumption permits us to use a single symbol, I, to denote the current flowing counterclockwise in the loop formed by R_1 and R_2. By Ohm's law, the current through a resistor is equal to the line integral of $\mathbf{E}$ from one end of the resistor to the other, divided by the resistance, due attention being paid to algebraic signs. In this case,

$$I = \frac{1}{R_1} \int_{\substack{A \\ C_1'}}^{B} \mathbf{E} \cdot d\mathbf{r} = \frac{1}{R_2} \int_{\substack{B \\ C_2'}}^{A} \mathbf{E} \cdot d\mathbf{r}. \tag{9}$$

Because C_1 and C_1' together form a closed curve not enclosing the solenoid, as do C_2 and C_2', Eq. (4) allows us to replace C_1' and C_2' by C_1 and C_2, respectively, in Eq. (9). With the use of Eq. (6), Eq. (9) thus becomes

$$I = V_1/R_1 = -V_2/R_2. \tag{10}$$

Finally, from Eqs. (8) and (10), we have the result we have been seeking:

$$V_1 = \frac{R_1}{R_1 + R_2} \alpha,$$

$$V_2 = -\frac{R_2}{R_1 + R_2} \alpha. \tag{11}$$

Note that V_1 and V_2 are always of opposite sign. Even though the red leads of the two meters are connected together, as are the black leads, the meter deflections will always be in opposite directions. If, for instance, $R_2 = 2R_1$, $V_1 = (1/3)\alpha$ and $V_2 = -(2/3)\alpha$. In the "symmetric" case $(R_1 = R_2)$, V_1 and V_2 are of equal magnitude: $V_1 = \alpha/2$, $V_2 = -\alpha/2$. (The case $R_1 = R_2$ is not truly symmetric, because the fact that $\partial \mathbf{B}/\partial t$ is directed into the paper results in a distinction between the clockwise and counterclockwise directions.)

V. EXPERIMENTAL CONFIRMATION

To avoid the difficulties in comparing theory and experiment that would result from using an iron-core coil, we use an air-core solenoid; ours happens to have a length of 104 cm, a mean radius of about 5 cm, and is wound with 444 turns of 0.038-cm diameter copper wire. Its inductance is 1.8 mH and its resistance is 19 Ω. A 5-Ω resistor is placed in series with the solenoid, so that the solenoid current can be monitored. The dimensions of the solenoid are not critical, though these dimensions determine the required characteristics of other apparatus. We use a triangular-wave generator at a frequency of 300 Hz, followed by a power amplifier, to produce a solenoid current of approximately triangular form; two channels of a multiple-trace oscilloscope serve as the voltmeters V_1 and V_2. This current provides a magnetic field satisfying the conditions of Eq. (2), with α alternately positive and negative, except at the times at which the waveform reverses slope. The choice of frequency is dictated by two competing effects: the higher the frequency, the larger the induced electric fields, but if the frequency is too high, a triangular voltage applied to the solenoid will not produce an approximately triangular current. Whatever method is used to drive the solenoid, it is important to monitor the voltage across a small resistor in series with the solenoid, so that one can be sure that $\mathbf{B}$ is varying linearly while V_1 and V_2 are being observed. Figure 5 shows the results obtained with $R_1 = 1$ kΩ and $R_2 = 2$ kΩ. According to the results derived earlier, V_1 and V_2 should be time independent when the solenoid current is varying linearly. Thus, if the solenoid current is a triangular function of time, V_1 and V_2 are predicted to be "square waves," as observed. Moreover, V_1 and V_2 should be of opposite phase, and the relative amplitudes should be $|V_2/V_1| = R_2/R_1 = 2$. These predictions are consistent with the experimental results. From measurements such as those shown on the top trace in Fig. 5, it is calculated that
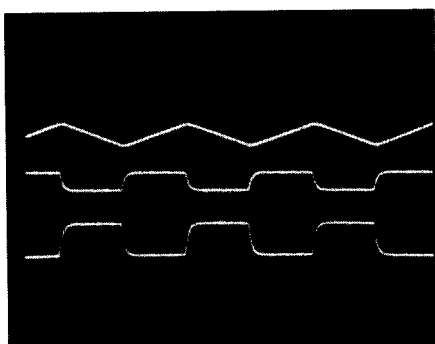
Fig. 5. Top trace is the voltage across a small resistor in series with the solenoid windings and is thus proportional to the solenoid current. The middle and lower traces show the voltages V_1 and V_2, respectively; the measured square waves have peak-to-peak voltages of approximately 4.3 and 8.4 mV, respectively.

Fig. 6. Topology of this circuit is quite different from that of Fig. 1, and for this circuit, $V_1 = V_2$.

the current through the solenoid changes at the rate of $\pm$ 1560 A/s during the two different parts of the triangular waveform. From this datum, the dimensions of the solenoid, and Eq. (11), it is predicted that V_1 and V_2 should be square waves of 4.4 and 8.8 mV peak-to-peak, respectively, a prediction in agreement with the observed oscilloscope traces to within about 5%.

Note that to the extent that the magnetic field exterior to the solenoid can be neglected, curl E is zero throughout region II, and it thus makes no difference, in principle, whether the voltmeter leads are or are not twisted or arranged to lie close to one another. In practice, however, solenoids are not infinitely long; thus R_1 and R_2 should be physically close to the solenoid and approximately midway between the ends of the solenoid. Anomalous results may be observed if the red and black leads to the oscilloscope are not physically close to one another throughout most of their length. It is crucial, though, even in the idealized experiment, that the physical configuration of the leads be such as to preserve the topological arrangement of Fig. 1 if Eq. (11) is to be valid. If, by accident or design, the leads are arranged as shown in Fig. 6, for example, V_1 *will* be equal to V_2, in both sign and magnitude, no matter what the values of R_1 and R_2. Once the apparatus is set up, the investigator will be unable to resist trying a number of arrangements of the leads, that shown in Fig. 6 and other variations suggested by Eq. (3) and Fig. 3 (wrapping one of the red leads once or twice around the solenoid before attaching it to point A, for instance) and explaining the signs and magnitudes of the resulting waveforms in terms of the simple physics described in this paper.

VI. "VOLTAGE" ACROSS AN INDUCTOR

The issues discussed in this paper are highly relevant to another disturbing question that I will briefly mention. Any introductory physics book will tell you that "the voltage across an inductor is $V_L = L\,(dI/dt)$." [Or perhaps it is $-L\,(dI/dt)$; signs *are* important, but that is not the point here.] The essence of an inductor is that it contains regions in which $\partial\mathbf{B}/\partial t \neq 0$, but that is precisely the circumstance that requires curl $\mathbf{E}\neq 0$, hence preventing us from defining a single-valued scalar potential for E and surely making the
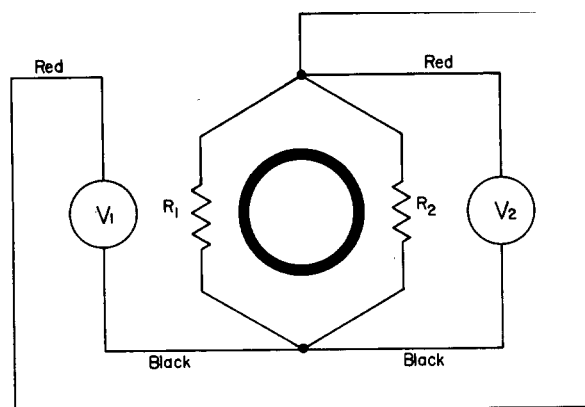
very *meaning* of "voltage" or "potential" suspect. Then what in the world can be meant by writing $V_L = L\,(dI/dt)$? Or, what can we mean by equations such as those we write to describe the transient discharge of a capacitor in a series $L - R - C$ circuit, $L\,(dI/dt) + RI + Q/C = 0$, equations which in the case of circuits *not* containing inductors are explicitly based on the proposition that E *is* a conservative field, so that any closed line integral of E vanishes? I was acutely distressed when I first asked myself these questions, and most other physicists seem to feel the same way. Most textbook authors ignore these issues, but there are a few[4,5] whose discussions of inductors will enable one to arrive at a satisfactory resolution.

VII. CONCLUSION

Of all the phenomena of physics, those associated with Faraday's law are among the most persistently fascinating and puzzling. How is it that $\partial\mathbf{B}/\partial t$ in one region demands the existence of curl E in that same region, and thus requires the existence of a nonvanishing E in other regions in which B and curl E both vanish? In a search of many books and papers on electromagnetic theory, I have been surprised to find little mention of the problem discussed in this paper, though I feel sure it must be known to many other admirers of electromagnetic theory. I myself first learned of this problem under trying circumstances, during an oral examination many years ago in which I was the examinee; I believe it was Eric Rogers and Frank Shoemaker who first tried to lead me through the beautiful physics of this situation. I am indebted to them for introducing me to the problem and for their tolerance of my initial confusion, and to J. Henry and M. Faraday for their wonderful discoveries.

[1]W. Klein, Am. J. Phys. **49**, 603 (1981).

[2]A. Shadowitz, *The Electromagnetic Field* (McGraw-Hill, New York, 1975), p. 396.

[3]D. R. Moorcroft, Am. J. Phys. **37**, 221 (1969); and **38**, 376 (1970); see also M. Phillips, Phys. Teach. **1**, 155 (1963).

[4]A. Shadowitz, Ref. 2, p. 438.

[5]R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, MA, 1964), Vol. II, p. 22-2.