Phenomenology and logical reasoning in introductory physics courses [a)]

A. B. Arons
Department of Physics, University of Washington, Seattle, Washington 98195

(Received 15 October 1980; accepted for publication 19 March 1981)

Specific illustrations are given of questions and problems designed to lead students in introductory courses into visualizing, and reasoning qualitatively about, physical phenomena.

I. INTRODUCTION

Consider the familiar demonstration experiment in which it is shown that two parallel wires attract each other when carrying electric current in the same direction. The demonstration is associated with the name of Andre Amperè and is generally introduced in the early stages of discussion of electromagnetism. In the course of the presentation, whether in text description or lecture demonstration, the interaction between the wires is almost invariably asserted to be "magnetic" or "electromagnetic" without reference to any other *a priori* possibility. I wish to illustrate that this quick and casual assertion of the character of the interaction destroys a significant learning experience for the student and marks the loss of a valuable pedagogical opportunity.

Since, in setting up this experiment, we connect the wires to a device (a generator or a voltaic battery) that is a source of electrical charge, we should recognize that any resulting interaction might conceivably be electrostatic. Through Oersted's experiment, demonstrating the effect of a current carrying wire on a neighboring compass needle, we anticipate the possibility of a magnetic interaction. How do we eliminate one possibility in favor of the other? That this is not a trivial question, however universally it might now be ignored in conventional pedagogy, is supported by the fact that Amperè himself gives it more than a page of discussion in his original paper.[1] (He had to convince his audience that the interaction was indeed a new effect, not that of ordinary static electricity.)

Amperè notes that the effect is dynamic rather than static since it ceases, as does electrolysis, the instant that contact is broken with the battery. He points out that the wires attract each other when the adjacent ends are connected to the same terminal of the battery (if static charge were responsible for the interaction, the distributions would be identical along the two wires, and they should repel). He further points out that, if free to move, the two wires stick together "like two magnets" and do not separate after contact as would conducting objects with unlike charges. He fails to emphasize the compelling significance of the fact that three or more wires all attract each other although he is aware of this phenomenon. He also fails to note that a wire carrying current in the opposite direction repels both of two attracting wires.

Lost in the usual quick assertion that the interaction of the wire is "electromagnetic" are two important aspects of scientific thought: (1) that we should raise the question as to whether we are really confronting a new phenomenon or merely some subtle variation of one already known; (2) the opportunity to think qualitatively and phenomenologically about the two possible interactions, distinguish between them operationally, and fix them in a richer perspective in one's own mind.

My suggestion is *not* that we should present all this didactically in text and lecture in addition to what is already done but that we should refrain from the immediate assertion of a new effect and lead the students into (1) recognizing that there is a nontrivial question to be asked, and (2) visualizing and identifying the experimental differences that compel us to recognize a new kind of interaction, profoundly different from the electrostatic. This should be elicited from the students via homework and class discussion rather than by formal presentation or "explanation."

Through the preceding example, I have tried to give an illustration of what I mean by "phenomenological" thinking and reasoning as opposed to mathematical or analytical. I propose to give further examples of the importance of such reasoning and will point to other opportunities for invoking it in introductory courses. It is not my intent to diminish the value or importance of the quantitative analytical modes. These are essential and must be retained. I contend, however, that we have created a serious imbalance in which the phenomenological modes are unwisely neglected relative to the analytical. I believe that student learning, comprehension, and intellectual self-confidence could be greatly improved by rebalancing the pedagogical approach. Such rebalancing is important for the entire spectrum of our students: future teachers, scientific and engineering professionals, and nonscientists for whom a physics course is a matter of liberal education.

II. ADDITIONAL ELEMENTARY EXAMPLES

Although it is true that many textbooks contain groups of qualitative questions aimed at inducing phenomenological reasoning, much of this effort seems to fail in several ways: The questions are frequently too sophisticated at too early a stage of development; they rely on the use and application of fully formed concepts and do not render the guidance most students need at intermediate stages of development. The questions rarely guide the student into crucial and significant inquiry regarding "How do we know...? Why do we believe...? What is the evidence for...?"—that inquiry that leads most directly to synthesis and command of the new knowledge. Furthermore, the qualitative questions are usually injected peripherally and are rarely made part of the full thrust of a particular development as is done in the case illustrated in Sec. I, where one addresses the question "How do we know the interaction between current carrying wires cannot be simply electrostatic?" Finally, phenomenological questions are rarely part of testing. Students will never regard such reasoning as important unless it is required on tests and plays a role in determining their final grade.

In the remainder of this section, I shall try to give some additional examples of very basic phenomenological questions that are neither tricky nor complicated, and that can be inserted in the main thread of a development or associated with conventional quantitative problems.

Consider the familiar mechanics problem given in virtually every introductory physics text and illustrated in Fig. 1: Cart *A* rolls without friction on a level surface. Block *B* is connected to cart *A* by a string of negligible mass running over a very light pulley *P*. This problem is usually presented in a quantitative form in which the masses, say, are given and the acceleration of the system and the tension in the string are to be calculated. Sometimes, but rather rarely, an algebraic analysis is called for, but even then the student is not led to interpret the algebraic results physically.

In recent years, I have been sporadically collecting data by asking first-year physics graduate students (who are usually working as teaching assistants in introductory courses and are helping students solve the textbook problem) the following question: "How does the force accelerating the cart compare *qualitatively* with the weight of object *B*: Is it larger, smaller, or equal to the weight of B?" In the data I have so far collected, 16 graduate students initially said "equal." When asked pointedly whether they would like to rethink their answer, three backtracked and concluded that the force accelerating *A* must be less than the weight of *B*; the remainder stuck to "equal." Only one student gave the correct answer immediately.

This observation dramatically illustrates the fact that ability to carry out the formal solution of a textbook problem gives no assurance of a real grasp of the underlying principles and physical phenomena. It illustrates, furthermore, that something has been seriously lacking in all the exposure these students had to elementary Newtonian mechanics: although they may recognize that unbalanced forces must be acting on an accelerating object in instances in which the forces are readily apparent and attention is focused on a single object, they have not registered this concept in sufficient depth to recognize its role in the two-body case of Fig. 1. Their phenomenological reasoning in this basic area is very weak. To my mind, this episode dramatizes the need of exposing students to this kind of question, not just on one isolated occasion, but in several instances, in order to help them register the concept deeply and permanently.

In order to be able to help students in this way, teachers must not only have mastered the insights themselves, they must be made aware of ways of inducing such phenomenological reasoning in the minds of their students.

A second illustration, at a similarly extremely fundamental level, actually precedes, and paves the way for, the question about the current carrying wires dealt with in Sec. I.

Many students in introductory courses have not formed a clear operational distinction between the interaction of electrically charged objects and the interaction of iron magnets with each other and with the earth, i.e., an operational distinction between electrostatic and magnetic interactions. This confusion emerges very quickly if one asks students to predict what will happen in situations that simultaneously involve charged particles and magnets. Many will, for example, expect positively charged particles to be repelled by magnetic north poles. If asked to describe what will happen in some given set of circumstances, many are likely to use the words "electric" and "magnetic" interchangeably.

An important and effective exercise in phenomenological reasoning is to ask students to prepare a list of operational similarities and differences between electrostatic and ferromagnetic interactions. An effective way is to have an entire class pool ideas, argue about them, and weed out incorrect ones. Such discussion leads to explicit awareness of the roles played by different material substances, of the fact that magnets are conductors but do not lose their interactive capacity by being handled, that both effects exhibit bipolarity but that magnetic poles are not separable, that contact between opposite magnetic poles does not result in a "discharge," and so on and so forth.

I encounter many university faculty who believe that such an exercise would be trivial. Yet I find very few elementary and secondary school teachers who can address the question competently. Dealing with such basic phenomenology was clearly not part of their educational background; they have not progressed to develop the insights and capacity on their own; and they are unaware of the confusion present in the minds of their students.

A third example illustrates the infusion of a measure of phenomenological reasoning into a quantitative problem. This problem was used on an open-book test in the second quarter of a calculus physics course during study of electrostatics and Coulomb's law. Part of the intent was to spiral back, in this new context, to the concepts of circular motion and centripetal force that the students had not had occasion to use for some time. The problem follows:

> In Fig. 2, we are looking down on two small pucks *A* and *B* resting on an air table. Each puck carries a uniformly charged sphere as indicated. Puck *A* is firmly fastened to the table. Puck *B* is free to move; friction is negligible. The total mass of puck *B* and its sphere is 125 g. Puck *B* is given a velocity of 9.45 m/sec in the direction indicated. (The point of this problem is to combine the use of concepts from this quarter's and last quarter's work. In this case, we invoke the concepts of centripetal force and centripetal acceleration. Be sure to use the *laboratory* frame of reference.)

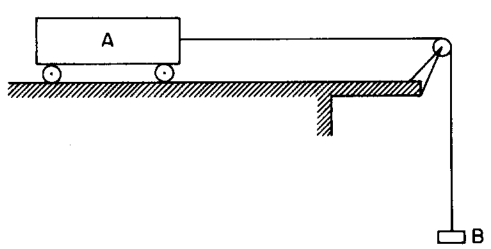

Fig. 1. Frequently used textbook problem. Cart *A* moves on level surface without friction. String and pulley have negligible inertia.

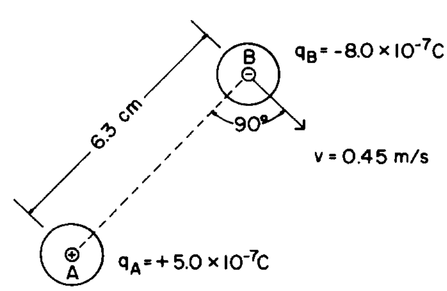

Fig. 2. Frictionless pucks subject to electrostatic interaction.

Will sphere *B* follow a circular orbit around *A* ? If not, will it deviate inward or outward? Show your numerical calculations and line of reasoning; handle units carefully. Unsupported answers will receive no credit. Be sure to show a "free-body" force diagram of sphere *B*.

The point of this way of presenting the problem was that the students were not told what to calculate. They had to *decide* what to calculate, and then they had to interpret their quantitative result. The students in this class had very little prior practice in phenomenological reasoning. They expected to be told what to prove or calculate when they were given a test question, and they did very badly in this first instance despite the essential simplicity of the problem and despite the fact that there is really very little choice as to what to calculate. The class average was 40% and only 25% of the class performed at a better than 75% level on the question. Many insisted on introducing a centrifugal force (and used the concept incorrectly) in spite of the fact that use of noninertial frames of references had been forbidden. Some of the students in this class were future secondary school teachers. To the best of my knowledge, none of these were in the upper 25% in performance. It took much repeated exposure to problems of this general variety before the majority of the class began to think about the physics of the situation under consideration instead of rushing blindly into a formal analysis. (It is worth noting that in this instance the majority of the students could successfully substitute in the formulas for Coulomb's law and for centripetal force if it were indicated that the two forces should be calculated.)

I am suggesting that the thinking required in this kind of problem is pedagogically valuable and that instruction might be significantly improved if we balanced our conventional quantitative textbook problems with more problems of this type—problems that ask what will happen in such and such simple circumstances, leaving both decision as to what to calculate and interpretation of the results up to the student. I believe it to be particularly important that we develop such capacity in future teachers.

III. BATTERIES AND BULBS

An area in which most students are very much in need of help with phenomenological reasoning is that of elementary resistive direct-current circuits. Conventional text presentations, problems, and tests channel them into exercises with the formulas for series–parallel combinations of resistors or into obtaining circuit equations by application of Kirchoff's laws but do not usually evoke any physical thinking or reasoning about what is happening in various parts of a simple circuit.

With the class of calculus physics students mentioned in the last example of Sec. II, I went through the following procedure: Using 1.5-V batteries and identical flashlight bulbs in small ceramic sockets, I set up simple series–parallel combinations in a lecture presentation, calling attention to the various levels of brightness with which the bulbs burned in any given arrangement. I called attention to the connection between bulb brightness and current and to the fact that some brightnesses changed when one bulb in a configuration was removed from its socket. I also called attention to brightness changes that occurred when additional bulbs were added in series or parallel with others already present and to brightness changes that occurred when a wire was used to short two points in a circuit. I finally called attention to the *qualitative* applicability of the concept of resistance in series and parallel combinations and to the applicability of the concept of potential difference. I then suggested that the students play with such battery-and-bulb systems for homework, either using their own equipment or borrowing kits I had available for loan. The point of the "play" was to invent their own configurations and predict what would happen when changes such as those illustrated in lecture were made. No quantitative calculations were possible or called for.

This assignment was followed by the following open-book test question:

This is an open-book test in which you are free to make use of the assigned text and of your own notes and homework. Please do not make use of any other sources such as printed study guides, etc.

In the circuit shown in Fig. 3 the battery maintains a constant potential difference between its terminals at points 1 and 2. The three identical bulbs *A*, *B*, and *C* are initially screwed into their sockets and lighted. After each of the following questions, the system is returned to this initial condition, and a new change is then made. *Indicate your reasoning briefly in all cases.*

The question "what happens to..." refers to whether the brightness or current increases, decreases, or remains unchanged.

(a) How do the brightnesses of bulbs *A*, *B*, and *C* compare with each other in the initial condition?

(b) What happens to the brightness of *each* bulb (*A*, *B*, and *C*) when bulb *A* is unscrewed? What simultaneously happens to the current in the wire at point 3?

(c) Return to the initial condition. What happens to the brightness of *each* bulb when bulb *C* is unscrewed? What simultaneously happens to the current in the wire at point 3?

(d) Return to initial condition. What happens to the brightness of *each* bulb if a wire is connected between the battery terminal at point 1 and point 4? What simultaneously happens to the current in the wire at 3? What simultaneously happens to the potential difference across bulb *B* ? What simultaneously happens to the potential difference across bulb *C* ?

(e) Return to initial condition. What happens to the brightness of *each* bulb and to the current in the wire at point 3 if a wire is connected from the battery terminal at point 2 to the socket terminal at point 5?

(f) Return to initial condition. What happens to the brightness of *each* bulb if a fourth bulb (*D*) is connected in parallel with bulb *B* (*not* in parallel with both *B* and *C*)?

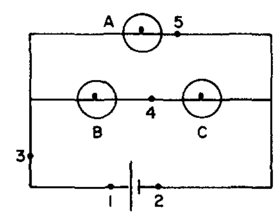

Fig. 3. Identical bulbs connected to a battery maintaining a constant potential difference across its terminals.

The class average on this question was 30% while the class average on a question calling for application of Kirchoff's laws to a much more complex circuit containing several resistors and several seats of emf was 65%. (The latter question called only for algebraic manipulation and not for phenomenological reasoning.)

Two weeks later I gave the following test question:

In the circuit shown in Fig. 4, the battery is sufficiently strong so that all the identical bulbs are visibly lighted, albeit with different degrees of brightness. The battery is new and exhibits negligible internal resistance unless it is actually short circuited.

The question, "What happens to..." refers to whether the brightness or current increases, decreases, or remains unchanged.

In answering the following questions, *indicate your reasoning briefly in all cases.*

(a) Suppose bulb *C* is removed from its socket. What happens to the brightness of each of the five bulbs? How do the final brightnesses compare with each other? What happens to the current at point 1?

(b) Return bulb *C* to its socket, restoring initial conditions. Suppose a wire is connected between points 2 and 3. What happens to the brightness of each of the five bulbs? How do the final brightnesses compare with each other?

What happens to the current at point 1?

What happens to the potential difference across bulb *C*?

What happens to the potential difference across bulb *B*?

The class average on this question was still 30%, but the greater difficulty of the question implies a slight improvement on the part of the class. I indicated that my insistence on qualitative reasoning about resistive circuits would *not* go away and that the next test would contain something of a similar nature. Now the students began to take the situation more seriously. They came in to check their reasoning on configurations and questions of their own invention. When I gave a third version of this type of question, the class average rose to 60%, still very far from where one would hope it might be on a third attempt and on such very basic and elementary physics.

This episode illustrates how much effort and repetition are necessary to get students to take questions of this kind seriously. Our conventional procedures tend to lock them into a very narrow and rigid mode. These questions led them to do far more physical thinking than any of the usual

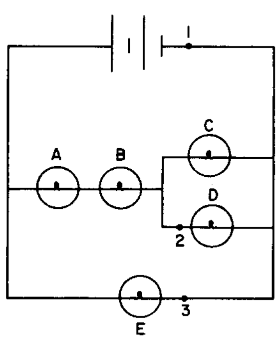

Fig. 4. Circuit consisting of identical bulbs connected to a source of constant potential difference.

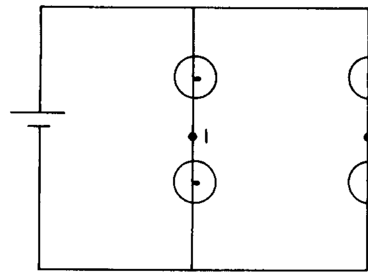

Fig. 5. Arrangement of bulbs paving way for "Wheatstone Bridge" concept.

network problems given in our standard text. Several of the better students thought about the configuration of Fig. 5 and wondered what would happen if a wire were connected between points 1 and 2; some even wondered what would happen if the bulbs were not identical. Some of these students had previously heard of the Wheatstone Bridge but had no understanding of what the device was for or how it worked. All I now had to do was hint at the idea, and they perceived the rest.

IV. THOMSON EXPERIMENT

A. Wave versus particle models

In some of the rich and important subject matter of introductory physics, we frequently neglect or overlook valuable opportunities for phenomenological thinking. I illustrate one such lost opportunity in the famous research of J. J. Thomson that marked the "discovery" of the electron.

If the "Thomson experiment" is described at all in a textbook, the description usually cites the application of crossed electric and magnetic fields to produce null deflection of the cathode beam, the elimination of the unknown velocity v from the equation for deflection with one field alone, and the calculation of e/m. Nothing is said about evidence for or against the corpuscular nature of the beam; nothing is said about other vital aspects of the investigation; nothing is said about the inferences to be drawn from the results. The entire research has been stripped of its rich phenomenological content and has been rendered pedagogically sterile. Let us examine the context more closely.

At the time of Thomson's investigation (1896–97),[2] there were two conflicting schools of thought concerning the nature of the cathode rays. The British school, following the earlier hypotheses and qualitative demonstrations of William Crookes, held the rays to be corpuscular. The Continental view, represented by Philipp Lenard, then at Bonn, held the rays to be a wave phenomenon in the ether.

Lenard's view was based on his numerous careful experiments with what were, for a long time, referred to as "Lenard rays." He studied an effect originally discovered by Hertz: The transmission of cathode rays through very thin metal foil "windows" in the end of the cathode-ray tube and the penetration of the rays through different gases after passing through the foil. Finding that the rays penetrated the foil and continued on for another centimeter or two, still in straight lines, Lenard initially became convinced that the cathode rays could not be corpuscular or material in character, but must be wave disturbances in the ether. He could not, at that time, conceive material charged particles penetrating a substance as dense as the metal foil without being deflected from straight-line paths. (He changed his view shortly thereafter.)

The opposing points of view of Crookes and Lenard exemplify the sharp distinction between corpuscular and wave phenomena that had emerged in nineteenth-century scientific thought. It was believed that these two kinds of behavior were mutually exclusive, that any given phenomenon must be of either one class or the other, that no manifestation could exhibit both corpuscular *and* wavelike aspects. (This complete dichotomy is something worth having students thinking about, and examining the validity of, as a prelude to subsequent introduction of our modern views of wave-particle duality).

Another experimental fact that at that time stood in the way of the corpuscular hypothesis was failure to achieve electrostatic deflection of the cathode beam by passing it between charged capacitor plates built into the tube. Crookes had shown the beam to be deflected in the direction expected for moving negatively charged particles on passage through a magnetic field, but attempts to produce electrostatic deflection yielded null results.

Thomson's paper of 1897, a classic of modern physics, made a concerted attack on the question of the nature of the cathode rays and proved to be a decisive treatment of the problem. The elegant and simple steps of physical reasoning threading this paper have tremendous pedagogical value.

B. Proof that the beam carries negative charge

Referring to the conflicting corpuscular and wave hypotheses, Thomson revealed some of the factors that had moulded his thought:

"The electrified particle theory has, for purposes of research, a great advantage over the aetherial theory, since it is definite and its consequences can be predicted; with the aetherial theory it is impossible to predict what will happen under any given circumstances, as on this theory we are dealing with hitherto unobserved phenomena in the aether, of whose laws we are ignorant."

"The following experiments were made to test some of the consequences of the electrified-particle theory."

Thomson first repeated and extended a very fundamental experiment carried out by the French physicist Perrin two years earlier. Perrin had inserted an electrometer cup (a metal cup, acting as a Faraday "ice pail," connected to an electroscope through the wall of the tube) into the tube opposite the cathode. When the tube was turned on, the cathode beam entered the cup, and the electroscope registered the collection of negative charge. When the beam was deflected magnetically so that it did not fall into the cup, no further charge was collected by the electrometer. Thomson extended this experiment by putting the electrometer cup at the *side* of the tube instead of opposite the cathode (Fig. 6). When the tube was turned on, the electrometer showed no charge; when the cathode beam was deflected magnetically so that it entered the cup, the electrometer indicated collection of negative charge. "This experiment shows," wrote Thomson, "that however we twist and deflect the cathode rays by magnetic forces, the negative electrification follows the same path as the rays, and that this negative electrification is indissolubly connected with the cathode rays."

Although the magnetic-deflection experiments of

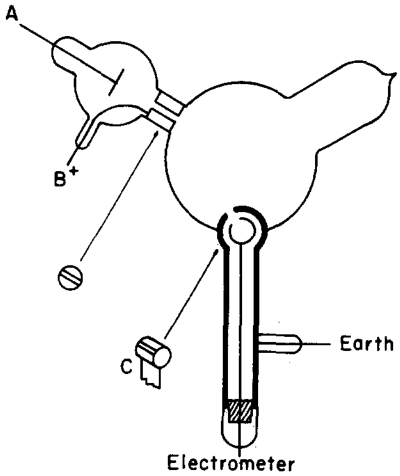

Fig. 6. Thomson's tube for demonstrating that cathode rays continue to transport negative charge even when deflected from their original path.

Crookes and others had supplied indirect evidence that cathode rays were associated with moving negative charge, the experiments of Perrin and Thomson provided the first *direct* confirmation. (Students should be led to think and argue about the motivation for these experiments and the interpretation of the results.)

C. Achieving electrostatic deflection

Thomson then attacked another crucial problem:

"An objection very generally urged against the view that the cathode rays are negatively electrified particles is that hitherto no deflection of the rays has been observed under a small electrostatic force.... Hertz made the rays travel between two parallel plates of metal placed inside the discharge tube, but found that they were not deflected when the plates were connected with a battery of storage cells; on repeating this experiment I at first got the same result, but subsequent experiments showed that the absence of deflection is due to the conductivity conferred on the rarefied gas by the cathode rays. On measuring this conductivity it was found that it diminished very rapidly as the exhaustion increased; it seemed then that on trying Hertz's experiment at very high exhaustions there might be a chance of detecting the deflection of the cathode rays by an electrostatic force."

Great discoveries have more than once hinged on basic insights like the one modestly advanced by Thomson. The insights may appear obvious or almost trivial in retrospect, but at the critical moment they were far from apparent to other individuals working in the same field. As a result of his preceding year and a half of experimenting and thinking about conductivity induced in gases by x rays, Thomson was very sensitive to the possible role of this phenomenon. He realized that cathode rays as well as x rays induce conductivity in gases, and he was well prepared to visualize the possible consequence. As it happened, newly developed vacuum techniques made it possible for him to test his ideas by achieving sufficiently high vacuum to suppress the conductivity:

"The rays from the cathode C (Fig. 7) pass through a slit in the anode A, which is a metal plug fitting tightly into the tube and connected with earth; after passing through a second slit in another earth-connected metal

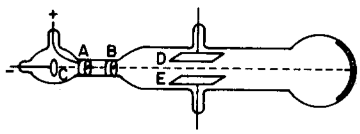

Fig. 7. Thomson's tube with capacitor plates D and E for producing electrical deflection of the cathode beam.

plug B, they travel between two parallel aluminum plates about 5 cm long and 2 broad and at a distance of 1.5 cm apart; they then fall on the end of the tube and produce a narrow well-defined fluorescent patch. A scale pasted on the outside of the tube serves to measure the deflection of this patch."

"At high exhaustions the rays were deflected when the two aluminum plates were connected with a battery of small storage cells; and rays were depressed when the upper plate was connected with the negative pole of the battery, the lower with the positive, and raised [when connections were reversed]. The deflection was proportional to the difference of potential between the plates, and I could detect the deflection when the potential difference was as small as two volts."

"It was only when the vacuum was a good one that the deflection took place, but that the absence of deflection is due to the conductivity of the medium is shown by what takes place when the vacuum has just arrived at the stage at which the deflection begins. At this stage there is a deflection of the rays when the plates are first connected with the terminals of the battery, but if this connection is maintained the patch of fluorescence gradually creeps back to its undeflected position. This is just what would happen if the space between the plates were a conductor, though a very bad one, for then the positive and negative ions between the plates would slowly diffuse until the positive plate became coated with negative ions, the negative plate with positive ones: thus the electric intensity between the plates would vanish and the cathode rays be free from electrostatic force...."

"As the cathode rays carry a charge of negative electricity, are deflected by an electrostatic force as if they were negatively electrified, and are acted on by a magnetic force in just the way in which this force would act on a negatively electrified body moving along the path of these rays. I can see no escape from the conclusion that they are charges of electricity carried by particles of matter."

Several years ago I put the following question on a qualifying examination for physics graduate students who had completed their course work and were about to begin thesis research: They were given the facts about the observation of magnetic deflection of the cathode beam and the failure to obtain electrostatic deflection. They were told about Thomson's hypothesis that ionization and resulting conductivity of the residual gas had something to do with this failure. They were told that attaining higher evacuation of the tube resulted in observation of the electrostatic deflection. They were asked to account for the experimental observations: what must have been happening at the lower vacua to prevent electrostatic deflection of the beam?

In the results of the examination, it became obvious that very few of the graduate students had had practice in dealing with such phenomenological reasoning. Only two out of twelve gave more or less satisfactory responses to the question. Should they not have had the opportunity to think about and visualize ion migration and space charge phenomena at much earlier stages of instruction?

D. Measurement of e/m

Thomson proceeds with:

"The question next arises. What are these particles? Are they atoms, or molecules, or matter in a still finer state of subdivision? To throw some light on this point, I have made a series of measurements of the ratio of the mass of these particles to the charge carried by [them]."

He then outlines the classic experiment with the crossed electric and magnetic fields, leading to the numerical value of e/m for the hypothetical cathode particles. Since this part is well known, I shall try to bring out only the phenomenological aspects that are lost in transmission in the majority of texts.

Analysis of the trajectory of the charged particle, both in the electrical field alone and in the magnetic field alone, provides an opportunity for repetition and review of previously studied material. The former carries back to concepts of projectile motion and emphasizes the mathematical identity of the gravitational and electrical cases while the latter injects review of circular motion, centripetal force, and the cross-product rule for direction of force on a moving charged particle in a B field. These analyses also lead the student to establish a connection between the deflection observed on the screen and the parameters of tube dimensions and field extent.

Such review is *not* a waste of time. It is essential for developing mastery and understanding; a serious fault of much modern course work manifests itself in the breathless volume and pace of coverage that exclude the possibility of such significant review and repetition.

A phenomenological aspect on which Thomson does not dwell but which provides particularly fine pedagogical opportunity resides in interpretation of the experimental fact that the spot on the screen retains its coherence, not smearing out under deflection with either electric and magnetic fields. Coupled with the algebraic description of the trajectory of a hypothetical particle, what does this experimental observation imply about the homogeneity, character, and origin of the beam? It is an observed fact that a positive ion beam (formed from the ionized gas in the cathode ray tube) does *not* retain its coherence, and a spot it forms on a screen *does* smear out under electric or magnetic deflection. How is this difference between the two types of beam to be interpreted? (Most first-year graduate students have great difficulty with this question.)

The elegant idea behind the null deflection in the crossed-fields technique is that the unknown velocity v of the hypothetical particles is eliminated via the relation $evB = eE$. Text presentations make the algebraic elimination, and obtain the expression for e/m with no further reference to v or its order of magnitude. This is not what Thomson does. He calculates v and shows it to be of the order of 10^7 m/sec, i.e., the order of one tenth of the velocity of light. With this numerical result, he questions the plausibility of the view that cathode rays are a wavelike disturbance in the ether. We should have the students calculate the value of v from representative data and argue about possible interpretations.

Few textbooks bother to assert that the effect of gravity on the deflection of the cathode particles is negligible; they leave it as an implication without even calling the question to the attention of the students. Yet here is a valuable opportunity for comparing sizes of physical effects and also applying earlier concepts in a new, richer context. I have had graduate students agree with the statement that gravity is negligible in this instance because the particles have such small mass, completely losing sight of their supposed "knowledge" from earlier work that g is the same for all masses. It is necessary to make students review the gravitational trajectory and recognize the role of the very high value of v in making the gravitational deflection immeasurably small. Surely we should make them begin such thinking about sizes of effects on the atomic–molecular scale while they are undergraduates.

E. Thomson's observations

Using the method referred to in Sec. IV D, Thomson made measurements of the charge-to-mass ratio of the cathode beam in tubes with electrodes of different metals (aluminum, platinum, iron) and with different gases (air, hydrogen, carbon dioxide) initially present and remaining in small amounts after evacuation.

In his paper Thomson remarks on certain systematic errors that he believed made his values of e/m somewhat lower than the true ones. At this juncture, however, he was not striving for high accuracy. He was pioneering new experimental techniques, and quite a few years were to elapse before their accuracy became such that they were reliable to within a few percent. Rather, he was interested in orders of magnitude, and he was trying to establish whether the charge-to-mass ratio associated with cathode rays varied over a wide range of values, as it was known to do with different ions in electrolysis and in conducting gases or whether it was essentially constant.

The results of all his different measurements fell between 0.67 and 0.9×10^{11} coul/kg. This being a very much narrower range of variation than that observed for different ions in electrolysis, and also being within the range of uncertainty of his experimental measurements.

Students should be led to confront the following questions: What was the point of using different electrode materials and different residual gases? What might be inferred from the results despite the large uncertainty?

It should be noted that, on the basis of the accumulated evidence, Thomson felt that the particle nature of cathode rays was fully demonstrated and any possibility of a wave interpretation excluded. This view was in fact quickly accepted by the scientific community.

F. Interpretation of the order of magnitude of e/m

It must be repeatedly emphasized to undergraduates that no physical quantity is large or small standing by itself. It is large or small only in comparison with some other relevant quantity, and, when a newly measured quantity is compared with one already known, a new or deeper physical insight frequently emerges.

In his paper, Thomson compares the charge to mass ratio associated with the hydrogen ion in electrolysis, ($q_H/m_H = 96\,500/0.001 = 9.6\times 10^7$ coul/kg) with his value for the cathode particle. Students should be led to make this comparison and argue about its implications with respect to the nature of the cathode particle and its possible role in the structure of matter. (A graduate student, who had been led through the preceding story and was then asked on an examination what might be inferred from the comparison of the charge-to-mass ratios, responded, "The e/m of electrons is much larger, as it *should* be.") Thomson's remark is the following:

> "Thus for the carriers of electricity in the cathode rays [e/m is very large] compared to its value in electrolysis. The [size of e/m] may be due to the smallness of m or the largeness of e, or to a combination of these two. That the carriers of the charges in cathode rays are small compared with ordinary molecules is shown, I think, by Lenard's results as to the rate at which the brightness of the (fluorescence) produced by these rays diminishes with the length of path traveled by the ray."

Thus Thomson uses this result to account for the penetrating power of the cathode particles and further to undermine the contention that cathode rays are a disturbance in the ether; at the same time he proceeds to argue that the cathode particle might be a *universally* occurring subatomic entity.

V. CONCLUSION

It is not my intention to argue that the Thomson experiment is itself unique or indispensible for achieving the educational experience I have outlined. Any comparable episode, similarly fundamental and with similar richness of context, would serve the purpose. The point is to use it in such depth and at such a pace as to give the students the opportunity to engage in the sophisticated, difficult, but exceedingly valuable phenomenological thinking inherent in such situations. It seems to me that pauses of this kind are far more conducive to student learning and intellectual growth than are rapid "coverage" of more advanced, more topical, or more mathematical material.

The great majority of students respond favorably and enthusiastically to the opportunity for thinking through something like the Thomson experiment. Never having had to bring together in one context as much prior knowledge as they do in such an instance, they naively marvel at the unity of physics and the beauty and richness of the perspective that emerges. They value being able to deal with the "how do we know...? why do we believe...?" questions at so basic a level; they sense how deeply this defines "understanding" and how profoundly it affects their view of their own knowledge and intellect.

It should be noted that the interpretations and inferences made in the Thomson experiment are not a matter of unequivocal proof and inescapable conclusion. They are, in many respects, conjectures, based on plausible guesswork concerning order and simplicity in nature (e.g., that it is corpuscle of charge and not the mass that has the same magnitude in cathode particles and in hydrogen ions). Students almost never have the opportunity to confront guesswork of this kind, and they labor under the delusion that all of scientific investigation involves "answers" as pat and unequivocal as those to be obtained in their narrow and artificial end-of-chapter problems.

Not only do questions and perspectives of the type illustrated above give students in introductory courses some sense of the real nature of physical thought, they also provide the repeated encounter in increasingly sophisticated context that, to use the Piagetian terms, is essential for enhancing their intellectual processes of accommodation, equilibration, and self-regulation.

[a]Prepared for presentation at the International Conference on Education for Physics Teaching, Triesk, Italy 1–6 September 1980.

[1]A. Ampère, Ann. Chim. Phys. Ser. II. **15**, 59 (1820).

[2]J. J. Thomson, Philos. Mag. **44**, 5 (1897).

Momentum propagation by traveling waves on a string

Reuben Benumof
The College of Staten Island, Staten Island, New York 10301

(Received 6 October 1980; accepted for publication 21 January 1981)

The one-dimensional wave equation for traveling transverse waves on a string implies three interesting consequences. (1) A traveling wave propagates both transverse and longitudinal momentum. (2) The time-rate of change of longitudinal momentum density is equal to the space-rate of decrease of stored energy density. (3) The longitudinal momentum density wave travels with the same speed as the displacement wave but may have a different form. As a result of the transmission of longitudinal momentum, longitudinal forces are exerted on both absorbers and reflectors.

I. PROBLEM OF MOMENTUM PROPAGATION ON A STRING

We are all accustomed to the idea that a transverse traveling wave carries transverse momentum in the direction of propagation. In the case of a string, for example, if y is the transverse displacement, then $\partial y/\partial t$ is the transverse velocity at a given point, and $\mu\ \partial y/\partial t$ is the transverse momentum per unit length, where μ is the mass per unit length. Thus we expect the transverse momentum density to be propagated with the speed of the displacement wave and, in effect, to be simply one of the characteristics of the wave itself.

On the other hand, we know intuitively that a transverse traveling wave can exert longitudinal forces when it impinges on an absorber or a reflector. Certainly, we are not surprised to learn that an electromagnetic wave in free space, which is definitely transverse, exerts pressure when it strikes a material surface. In the case of a string, which is presumably familiar, it is nevertheless not at all obvious how or why longitudinal momentum is propagated by transverse traveling waves. Accordingly, this paper has two purposes. First, we wish to explain quantitatively how longitudinal momentum is propagated by transverse traveling waves on a stretched string. Second, when the waves are either absorbed or reflected, we want to show how the longitudinal forces exerted by the string may be calculated. This matter was treated briefly in an earlier paper.[1] Here we shall use different methods and considerably amplify the earlier treatment.

This topic is important because transverse waves on a string serve as a prototype for wave motion in general. We can easily visualize such waves, and for this reason wave motion in other media is often clarified by alluding to waves on a string. Thus if the string is to serve as a model, we must understand how transverse traveling waves propagate longitudinal momentum and exert longitudinal forces.

II. ORIGIN OF A TRAVELING WAVE

To discuss traveling waves realistically, we may use the following situation as a model. The left end of a very long string is connected to a source of energy, such as a vibrator, and the right end, suitably supported, is attached to a hanging weight. Before the vibrator is turned on, the string is horizontal. When the vibrator is activated, a wave begins to travel along the string from left to right. Energy and momentum supplied by the vibrator are carried by the wave. The wave cannot be established immediately along the entire length of the string. Otherwise, an infinite amount of power would be required. It is important to recognize that the string thus has two sections. The left-hand section vibrates and contains the traveling wave. This section continually increases in length. The right-hand section has not yet been activated and is therefore quiescent. This section acts as a perfect absorber for the traveling wave.

III. TENSION IN A PERFECTLY FLEXIBLE STRING

We now wish to discuss the tension in the string. Figure 1 shows a differential element of the string of initial length dx. When this element is displaced transversely from its equilibrium position at some time t, its new length is given by

$$d\mathbf{r} = \mathbf{i}(dx + dz) + \mathbf{j}\,dy = \mathbf{i}\left(1 + \frac{\partial z}{\partial x}\right)dx + \mathbf{j}\,\frac{\partial y}{\partial x}\,dx, \quad (1)$$

where dz is a very small horizontal increment in length, y is a small transverse displacement from the equilibrium position, and dy is the vertical displacement of the right end of the element relative to the left end. We should note that dz is much smaller than dy, which is turn is much smaller than dx. From Eq. (1) we have

$$|d\mathbf{r}| = \left[\left(1 + \frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial y}{\partial x}\right)^2\right]^{1/2} dx. \quad (2)$$

If the tension acting on the element of string when it is in its equilibrium position is T_0, the tension T in the displaced position must be

$$T = T_0 + \Delta T, \quad (3)$$

