AAPT PHYSICS EDUCATION | AMERICAN JOURNAL of PHYSICS

Introductory helium atomic spectrum analysis

Nina Abramzon and P. B. Siegel

Citation: American Journal of Physics **77**, 920 (2009); doi: 10.1119/1.3086029
View online: http://dx.doi.org/10.1119/1.3086029
View Table of Contents: http://scitation.aip.org/content/aapt/journal/ajp/77/10?ver=pdfcov
Published by the American Association of Physics Teachers

On using moments around the instantaneous center of rotation

W. F. D. Theron[a)]
Department of Applied Mathematics, University of Stellenbosch, Private Bag X1, Matieland, 7602, South Africa

(Received 3 June 2008; accepted 24 November 2008)

A previously unpublished example is presented to illustrate the advantages of using the instantaneous center as a point around which moments can be calculated. © *2009 American Association of Physics Teachers.*
[DOI: 10.1119/1.3048539]

I. INTRODUCTION

There has been some discussion as to whether the instantaneous center may be used as a safe point for the torque equation.[1,2] In particular, Tiersten[2] illustrated how "…quick and elegant solutions of many problems in plane motion of a rigid body may be obtained by taking moments around the instantaneous center of rotation" using the relation

$$M_C = I_C\alpha + \frac{1}{2}\dot{I}_C\omega. \tag{1}$$

Point C refers to the instantaneous center, M_C denotes the total moment of the forces around C, I_C is the corresponding moment of inertia, α is the angular acceleration, ω is the angular velocity, and the dot represents the time derivative.

Equation (1) is found in classic texts such as Loney[3] and Lambe,[4] and it is convenient to refer to it as *Loney's equation*. We have used Eq. (1) for many years to solve problems in rigid body dynamics, and it came as a surprise when we realized that its use is not the norm in contemporary engineering mechanics textbooks such as Ref. 5 for example.

In this note a new example is presented to illustrate the use of the instantaneous center, together with a number of remarks regarding Loney's equation.

II. EXAMPLE—A SLIDING ROD

The well known textbook problem of a rigid rod sliding down a smooth vertical wall along a smooth horizontal floor is generalized by loading the rod so that the center of mass is not at the middle of the rod, as shown in Fig. 1. Rod AB has mass m and length L, and its center of mass is at point G with $AG=\gamma L$. The moment of inertia with respect to G, I_G, is written as $I_G=\kappa_G mL^2$, with κ_G a known constant. (For example, if a uniform rod with mass m_r is loaded with a particle at B with mass $m_B=2m_r$, then $m=3m_r$, $\gamma=5/6$ and $\kappa_G=1/12$).

The rotation of the rod is measured by θ as shown in Fig. 1, and the angular velocity and acceleration are $\omega=\dot{\theta}$ and $\alpha=\ddot{\theta}$, taking anti-clockwise as positive. The forces acting on the rod are the weight mg and the normal reactions N_A and N_B. The rod has an initial angular velocity ω_1 at initial position θ_1.

The instantaneous center is located at point C, which is the intersection of lines AC and BC which are perpendicular to the velocities of points A and B, respectively. Then $(GC)^2=L^2(\gamma^2+(1-2\gamma)\cos^2\theta)$, and the moment of inertia with respect to C is

$$I_C = \kappa_C mL^2, \tag{2}$$

with $\kappa_C=\kappa_G+\gamma^2+(1-2\gamma)\cos^2\theta$. Note that for an unloaded uniform rod with $\gamma=1/2$, $GC=L/2$, and the moment of inertia around C is independent of θ. The total moment of the external forces around C is

$$M_C = mg\gamma L\sin\theta. \tag{3}$$

We substitute Eqs. (2) and (3) in Eq. (1) and cancel mL^2 and write Eq. (1) as

$$(g/L)\gamma\sin\theta = \alpha\kappa_C + \frac{1}{2}\omega^2(2\gamma-1)\sin 2\theta, \tag{4}$$

which can be written as the linear first-order differential equation

$$\kappa_C\frac{d\omega^2}{d\theta} + (2\gamma-1)\sin 2\theta\omega^2 = 2(g/L)\gamma\sin\theta. \tag{5}$$

The solution of Eq. (5) with the initial condition $\omega(\theta_1)=\omega_1$ is

$$\omega^2 = (1/\kappa_C)(\kappa_{C_1}\omega_1^2 + 2\gamma(g/L)[\cos\theta_1 - \cos\theta)], \tag{6}$$

where $\kappa_{C_1}=\kappa_G+\gamma^2+(1-2\gamma)\cos^2\theta_1$.

Alternatively, an energy principle may be used. A well known result, see Ref. 5 for example, is that the kinetic

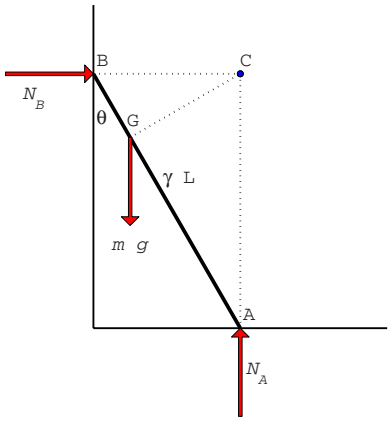

Fig. 1. An eccentrically loaded rod sliding down a smooth wall on a smooth surface.

energy of a planar system can be found with reference to the instantaneous center as

$$T = \frac{1}{2} I_C \omega^2. \tag{7}$$

If we take the floor as the reference plane, the potential energy is $V = mg\gamma L \cos\theta$, and the energy conservation results in

$$\frac{1}{2} mL^2 \kappa_C \omega^2 + mg\gamma L \cos\theta = \frac{1}{2} mL^2 \kappa_{C_1} \omega_1^2 + mg\gamma L \cos\theta_1, \tag{8}$$

which simplifies to Eq. (6). The angular acceleration can be obtained by differentiating Eq. (6) or by rewriting Eq. (1) as

$$\alpha = (1/\kappa_C)[(g/L)\gamma \sin\theta - \omega^2(\gamma - 1/2)\sin 2\theta]. \tag{9}$$

This simple example could also have been solved by other methods, but illustrates the ease with which the problem can be solved in terms of moments around the instantaneous center. Similar types of problems that are well suited to this approach are given by Tiersten[2] and Theron.[6]

III. CONCLUDING REMARKS

Equation (1) is not a relation between the vector quantities of torque and angular acceleration, and thus is not a true torque equation. Rather, it should be recognized as a derivative of an energy principle, which becomes clear by considering the manner in which Eq. (1) is derived. In the original derivation,[3] Loney started by taking moments in the standard manner and then used a scalar product to change to kinetic energy quantities. Tiersten[2] started with the energy principle that $P = dT/dt$, where P is the power (the rate at which work is being done). It is easy to show that $P = \omega M_C$ and, using Eq. (7), that Eq. (1) follows.

The torque does not equal the rate of change of the angular momentum when point C is used as the reference, except in the special case when CG is constant. This is the reason why some authors, for example Desloge,[7] strongly oppose the use of point C in the torque equation.

Also, and not unexpectedly, Eq. (1) can always be integrated to obtain the work-energy relation. Noting that $dI_C/dt = (dI_C/d\theta)\omega = I_C'\omega$ and $\alpha = (\omega^2)'/2$, Eq. (1) can be rewritten as the first-order differential equation

$$I_C \frac{\mathrm{d}}{\mathrm{d}\theta}(\omega^2) + I_C'\omega^2 = 2M_C. \tag{10}$$

As seen in the example of the sliding rod, Eq. (10) has the simple integrating factor $\exp\int (I_C'/I_C)\mathrm{d}\theta = I_C$, and the solution

$$I_C\omega^2 = 2\int M_C\,\mathrm{d}\theta + C_1, \tag{11}$$

where C_1 is a constant of integration found from the initial conditions. If we recognize that $\int M_C\,\mathrm{d}\theta$ is the work done on the system, Eq. (11) can be written as the standard relation between work and the change in kinetic energy.

There are no problems or subtle pitfalls when using moments around the instantaneous center in the various energy principles. To summarize we have

$$T = \frac{1}{2} I_C \omega^2, \tag{12a}$$

$$P = \omega M_C, \tag{12b}$$

$$\text{total work} = \int_1^2 M_C\,\mathrm{d}\theta. \tag{12c}$$

There is no reason why moments around the instantaneous center should be neglected!

ACKNOWLEDGMENTS

The author is indebted to the many colleagues and reviewers who influenced him over many years, and also to the reviewer of this note for his/her patience and constructive suggestions.

[a]Electronic mail: wfdt@sun.ac.za

[1] F. R. Zypman, "Moments to remember—The condition for equating torque and rate of change of angular momentum," Am. J. Phys. **58**(1), 41–43 (1990).
[2] M. S. Tiersten, "Moments not to forget—The conditions for equating torque and rate of change of angular momentum around the instantaneous center," Am. J. Phys. **59**(8), 733–738 (1991).
[3] S. L. Loney, *Dynamics of a Particle and of Rigid Bodies* (Cambridge U. P., London, 1946), p. 257 (first published in 1909).
[4] C. G. Lambe, *Applied Mathematics for Engineers and Scientists* (The English Universities Press, London, 1958), p. 155.
[5] J. L. Meriam and L. G. Kraige, *Engineering Mechanics: Dynamics*, 4th ed. (J Wiley, New York, 1998).
[6] W. F. D. Theron and N. M. du Plessis, "The dynamics of a massless hoop," Am. J. Phys. **69**(3), 354–359 (2001).
[7] E. A. Desloge, *Classical Mechanics* (J. Wiley, New York, 1982), p. 230. As quoted by Zypman and by Tiersten in Refs. 1 and 2.

Introductory helium atomic spectrum analysis

Nina Abramzon and P. B. Siegel
California State Polytechnic University Pomona, Pomona, California 91768

(Received 8 December 2008; accepted 30 January 2009)

We present a spreadsheet that introduces undergraduate students to the analysis of helium atomic spectrum data.
[DOI: 10.1119/1.3086029]

I. INTRODUCTION

The measurement and analysis of the visible hydrogen spectrum is a standard experiment performed by nearly every undergraduate physics major. Students measure the four visible lines of the Balmer series, and the analysis includes verifying Balmer's formula and measuring the Rydberg constant. The hydrogen spectrum experiment is ideal because the spectral data were key ingredients in the development of quantum mechanics and can be understood using a simple relation.

The spectra of other atoms are not as simple to analyze. Adding just one more proton to the nucleus and one more orbiting electron to the atom for helium results in the complications of shielding and the effects of indistinguishable particles on the energy level structure. The energy levels of helium are discussed in upper level quantum mechanics courses as an example of identical fermions,[1] and an analysis of the helium spectrum data is usually not done by undergraduate physics students. If it is done at all, the experiment often consists of the students measuring the emitted wavelengths and using tables to identify the corresponding atomic transitions.[2] Although this experiment introduces the students to terminology in atomic physics, a quantitative analysis is absent.

In this note we present a spreadsheet calculation to be used in the laboratory to analyze the helium spectrum data. By tabulating the difference in the photon energies of helium and hydrogen, the students can measure the singlet-triplet splitting of the $2p$ level in helium and identify the d level transitions to this state.

II. HELIUM SPECTRUM EXPERIMENT

After carrying out the hydrogen spectrum experiment students collect data from helium. In both experiments data are collected using a spectrometer interfaced to a personal computer.[3] The students begin by calibrating the spectrometer using a standard source.[4] The hydrogen spectrum yields four strong lines plus one weak line of the Balmer series. The wavelengths[5] in nanometers of the Balmer lines are listed in the top row of Table I, along with the initial principal quantum number of the electron, n_i. The final quantum number for the Balmer series is $n_f=2$.

The students collect the helium spectra by first using a short collection time of 10 μs. The spectrum is taken from a helium discharge tube with a glass container. The collection time is chosen so that the most intense line, 587.6 nm, fits on the display. Although the glass absorbs in the ultra-violet, the students can identify 14 spectral lines, which are the first 14 lines listed in Table I. Figure 1 is a plot of the 10 μs helium spectrum. After measuring the wavelengths of the short collection time data, the students collect data for a longer time to observe and measure less intense emissions. The last four emission lines in Table I, listed in italics, can be observed with a 5 s collection time. In Fig. 2 we display part of the helium spectrum for a 5 s collection time, where we label the spectral lines in units of nanometers. The weakest line at 414.4 nm is small but measurable. The first column of Table I is a list in order of intensity of the 18 helium emission lines observed with our spectrometer.

III. DATA ANALYSIS

After the students have measured the helium wavelengths, we begin a discussion on how to analyze the data. We start by proposing a model for an excited helium atom. We assume that an excited state of helium consists of one electron, the screening electron, in the lowest energy level, the $1s$. The second electron, or valence electron, is in an excited state. When a transition occurs, the energy of the emitted photon is

Table I. The differences in energy of the measured hydrogen and helium visible radiation. The entries in the table are $(1240/\lambda_{He}-1240/\lambda_{H})$ eV with the wavelengths in nanometers. The uncertainty of each entry in the table is 0.006 eV. The last four rows are in italics to point out that these lines are weak and need a long collection time to be observed. The bold numbers are the transitions from the corresponding d levels to the $2p$ singlet and triplet final states.

n_i (hydrogen):	3	4	5	6	7
λ (nm) H→	656.3	486.1	434.0	410.2	*397.0*
He↓					
587.6	**0.22**	−0.44	−0.75	−0.91	−1.01
706.5	−0.13	−0.80	−1.10	−1.27	−1.37
667.8	**−0.03**	−0.69	−1.00	−1.17	−1.27
388.9	1.30	0.64	0.33	0.17	0.07
501.6	0.58	−0.08	−0.39	−0.55	−0.65
728.1	−0.19	−0.85	−1.15	−1.32	−1.42
447.1	0.88	**0.22**	−0.08	−0.25	−0.35
492.2	0.63	**-0.03**	−0.34	−0.50	−0.60
471.3	0.74	0.08	−0.23	−0.39	−0.49
402.6	1.19	0.53	**0.22**	0.06	−0.04
396.5	1.24	0.58	0.27	0.10	0.00
438.8	0.94	0.27	**−0.03**	−0.20	−0.30
382.0	1.36	0.70	0.39	**0.22**	0.12
318.8	2.00	1.34	1.03	0.87	0.77
412.1	0.12	0.46	0.15	−0.01	−0.11
370.5	1.46	0.80	0.49	0.32	**0.22**
361.8	1.54	0.88	0.57	0.40	0.30
414.4	1.10	0.44	0.14	**-0.03**	−0.13

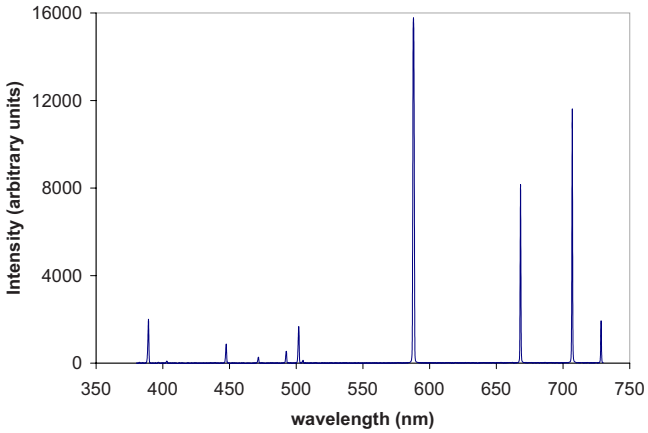

Fig. 1. A plot of the helium spectrum for a collection time of 10 μs. All the spectral lines except the last four in Table I can be observed. The data were collected using the spectrometer of Ref. 3.

approximately equal to the energy lost by the valence electron as it goes from one level to another. This model is only an approximation, because the state of the screening electron depends on the state of the valence electron. In addition, the two electrons are not independent, but satisfy the properties required of indistinguishable particles. In the extreme case that the screening electron completely screens out the nucleus, the valence electron would have the same energy levels as an electron in hydrogen.

The analysis begins by having the students perform a spreadsheet calculation. The hydrogen wavelengths are entered in the first row of the spreadsheet according to their initial principal quantum number, n_i. The helium wavelengths are entered in the first column of the spreadsheet according to their relative intensity. The students fill the spreadsheet with the energy differences ΔE_{ij} between all combinations of the emitted photons of hydrogen and helium by computing the following quantity to examine if there are any patterns in the energy differences:

$$\Delta E_{ij} = hc\left[\frac{1}{\lambda_i(\mathrm{He})} - \frac{1}{\lambda_j(\mathrm{H})}\right] = 1240\left[\frac{1}{\lambda_i(\mathrm{He})} - \frac{1}{\lambda_j(\mathrm{H})}\right] \mathrm{eV}, \quad (1)$$

where λ_i(He) is the helium wavelength (in nm) in the ith row and λ_j(H) is the hydrogen wavelength (in nm) in the jth column. Note that i and j are not indices for the ith or jth energy level. In Table I we display the results of the calculation of ΔE_{ij}.

The students are then asked to identify any patterns in the ΔE_{ij} array. As seen in Table I, 0.22 eV occurs five times and −0.03 eV occurs four times, once in each column indicated in bold. The interpretation of these results leads to the following discussion. Consider first the 0.22 eV difference, which is the same value for n_i=3, 4, 5, 6, and 7 in hydrogen with n_f=2. One way to have the same energy difference for these five lines is to have the binding energy of the valence electron in helium for n_i=3, 4, 5, 6, and 7 be the same as in hydrogen and the n_f=2 level for the valence electron in helium be shifted 0.22 eV compared to hydrogen. The orbital quantum number ℓ plays an important role. For larger values of ℓ the electron's location is farther from the nucleus. Because $\Delta\ell$=1 for dipole radiation, the largest values of ℓ for the transitions are $7d\rightarrow 2p$, $6d\rightarrow 2p$, $5d\rightarrow 2p$, $4d\rightarrow 2p$, and $3d\rightarrow 2p$. The five values of 0.22 eV in Table I can be understood if the ℓ=2 (d) states for the valence electron in helium have the same energy to the accuracy of our measurements as the corresponding states for the electron in hydrogen. That is, for the ℓ=2 states of the helium valence electron, the screening electron successfully screens the nucleus so that the valence electron experiences the potential from just one proton. The 0.22 eV shift is therefore due to the shift of the $2p$ level of the valence electron from that of the $2p$ level of hydrogen.

The −0.03 eV differences can be similarly understood if the $2p$ level of helium's valence electron is shifted up 0.03 eV from the energy of the $2p$ level of hydrogen. Thus,

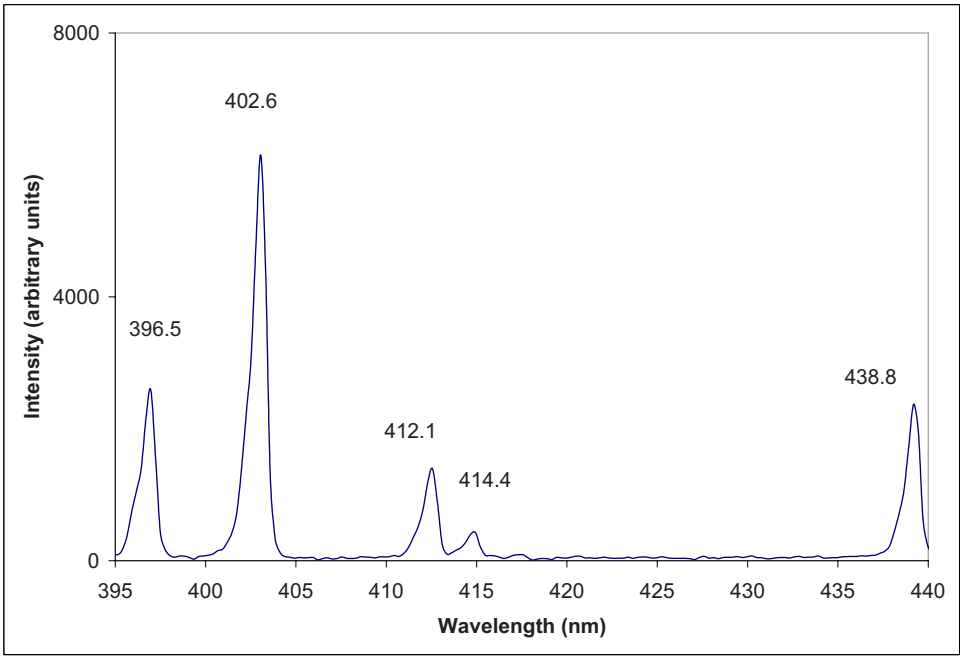

Fig. 2. A plot of part of the helium spectrum for a collection time of 5 s.

the multiple occurrences of 0.22 eV and −0.03 eV in Table I can be explained if the $\ell=2$ energies of helium's valence electron are not shifted significantly from the corresponding levels in hydrogen, and the $2p$ level for helium's valence electron is split: one state being 0.22 eV lower than the corresponding hydrogen state and the other 0.03 eV higher. We note that the expected −0.03 eV singlet entry in the $n_i=7$ column corresponds to a wavelength of 400.93 nm. This wavelength in the helium spectrum is not observable with our spectrometer because the $7d\rightarrow 2p$ singlet transition is very weak and is obscured by the much stronger $5d\rightarrow 2p$ triplet line at 402.6 nm.

The problem concludes with a discussion of the splitting of the $2p$ valence electron level in helium. The largest contribution to the splitting is due to the exchange interaction, which is a consequence of the collective properties of the identical electrons.[6] It is also referred to as singlet-triplet splitting. The triplet level, or orthohelium, has a lower energy than the singlet level, or parahelium. The total splitting measured in the experiment is 0.25 eV and agrees with the published value for the $1s2p$ level splitting in helium.[5]

It is important to have adequate resolution of the wavelengths for the experiment to be successful. An accuracy of ± 0.5 nm for the wavelength is desirable. An uncertainty of ± 0.5 nm is $\approx 0.1\%$ error in λ and consequently 0.1% error in the energy of the emitted photon. The largest photon energies are 3 eV, so the uncertainty in the photon energies is around 0.003 eV. Because the entries in Table I are the difference of two photon energies, their uncertainties are around ± 0.006 eV. This resolution is enough to recognize the pattern of 0.22 eV, but insufficient to see any further splitting of the levels. The spectrometers of Ref. 3 are rated to have an uncertainty of 0.7 nm in wavelength and are ideal for the experiment.

In conclusion, the experiment allows students to identify nine of the helium spectral lines: $1s\ n_id\rightarrow 1s2p$ triplet for $n_i=3$, 4, 5, 6, and 7, and $1s\ n_id\rightarrow 1s2p$ singlet for $n_i=3$, 4, 5, and 6. This problem is ideal for an upper division physics laboratory and can also be used in a sophomore level atom physics laboratory.

[1] R. L. Liboff, *Introductory Quantum Mechanics*, 3rd ed. (Addison-Wesley, Reading, MA, 1998).
[2] M. C. Lo Presto, "A closer look at the spectrum of helium," Phys. Teach. **36**, 172–173 (1998).
[3] We use the Ocean Optics HR4000 spectrometer.
[4] We use a standard source provided by the HR4000 spectrometer. Alternatively, we could use the accepted values of the Balmer lines of hydrogen to calibrate the spectrometer.
[5] The NIST Atomic Spectra Data website is ⟨physics.nist.gov/PhysRefData/ASD/lines-form.html⟩.
[6] A. Goswami, *Quantum Mechanics*, 2nd ed. (Brown, Dubuque, IA, 2001).

Evolution and the second law of thermodynamics

Emory F. Bunn[a)]
Department of Physics, University of Richmond, Richmond, Virginia 23173

(Received 11 December 2008; accepted 26 March 2009)

Skeptics of biological evolution often claim that evolution requires a decrease in entropy, giving rise to a conflict with the second law of thermodynamics. This argument is fallacious because it neglects the large increase in entropy provided by sunlight striking the Earth. A recent article provided a quantitative assessment of the entropies involved and showed explicitly that there is no conflict. That article rests on an unjustified assumption about the amount of entropy reduction involved in evolution. I present a refinement of the argument that does not rely on this assumption. © *2009 American Association of Physics Teachers.*

[DOI: 10.1119/1.3119513]

I. INTRODUCTION

Styer recently addressed the claim that evolution requires a decrease in entropy and therefore is in conflict with the second law of thermodynamics.[1] He correctly explained that this claim rests on misunderstandings about the nature of entropy and the second law. The second law states that the total entropy of a closed system must never decrease. However, the Earth is not a closed system and is constantly absorbing sunlight, resulting in an enormous increase in entropy, which can counteract the decrease presumed to be required for evolution. This argument is known to those who defend evolution in evolution-creationism debates,[2] but it is usually described in a general, qualitative way. Reference 1 filled this gap with a quantitative argument.

In the following I present a more robust quantitative argument. We begin by identifying the appropriate closed system to which to apply the second law. We find that the second law requires that the rate of entropy increase due to the Earth's absorption of sunlight, $(dS/dt)_{\text{sun}}$, must be sufficient to account for the rate of entropy decrease required for the evolution of life, $(dS/dt)_{\text{life}}$ (a negative quantity). As long as

$$\left(\frac{dS}{dt}\right)_{\text{sun}}+\left(\frac{dS}{dt}\right)_{\text{life}}\geqslant 0, \tag{1}$$

there is no conflict between evolution and the second law.

Styer estimated both $(dS/dt)_{\text{sun}}$ and $(dS/dt)_{\text{life}}$ to show that the inequality (1) is satisfied, but his argument rests on an unjustified and probably incorrect assumption about $(dS/dt)_{\text{life}}$.[1] I will present a modified version of the argument, which does not depend on this assumption and which shows that the entropy decrease required for evolution is orders of magnitude too small to conflict with the second law of thermodynamics.

II. ENTROPY PROVIDED BY SUNLIGHT

Let us begin by justifying the inequality (1). The Earth maintains an approximately constant temperature by absorbing energy from the Sun and radiating energy at an almost equal rate. To consider the application of the second law of thermodynamics to these processes we first identify a closed system that is large enough that these energy flows may be considered to be internal to the system. Let us take our system to be the Earth, the Sun, and the outgoing thermal radiation emitted by these bodies. We will ignore interactions of this radiation with bodies other than Earth and Sun and consider the outgoing radiation from each to form an ever-expanding spherical halo. In this system no entropy is produced by emission of radiation from the Sun, because this process is a flow of energy from the Sun to its radiation field at the same temperature. The same applies to radiation emitted by the Earth. Entropy production occurs only when radiation from the Sun is absorbed on the Earth, because this absorption represents energy flow between parts of the system at different temperatures.

Let $T_\odot$ and $T_\oplus$ be the temperatures of Sun and Earth, respectively, and let P be the solar power absorbed by Earth. (To be precise, P is the net flow from Sun to Earth, including the backward flow of energy from Earthshine being absorbed on the Sun, but the latter contribution is negligible.) The Earth gains entropy at the rate $P/T_\oplus$, and the Sun's radiation field loses entropy at the rate $-P/T_\odot$. The rate of entropy production is

$$\left(\frac{dS}{dt}\right)_{\text{sun}} = \frac{P}{T_\oplus} - \frac{P}{T_\odot} \approx \frac{P}{T_\oplus}, \tag{2}$$

where the last approximate equality reflects the fact that the Sun's temperature is much larger than the Earth's. If we assume that the evolution of life requires the entropy to decrease at the rate $(dS/dt)_{\text{life}}$, the second law of thermodynamics applied to this system gives Eq. (1).

By using values for the solar constant[3] and Earth's albedo,[4] Styer[1] found that Earth absorbs solar radiation at a rate of $P=1.2\times10^{17}$ W. If we use $T_\oplus=300$ K as a rough estimate of Earth's temperature, we find that

$$\left(\frac{dS}{dt}\right)_{\text{sun}} = \frac{P}{T_\oplus} = 4\times10^{14}\ \text{(J/K)/s} = (3\times10^{37}k)\ \text{s}^{-1}, \tag{3}$$

where k is the Boltzmann constant.

In this estimate we did not include any entropy increase due to thermalization of the radiant energy emitted by the Earth. If we assume that this radiation eventually thermalizes with the cosmic background (CMB) radiation in deep space, then an additional, much larger entropy increase results: $(dS/dt)_{\text{CMB}}=P/T_{\text{CMB}}=4\times10^{16}$ (J/K)/s. We may not include this entropy production in accounting for evolution. One reason is that this thermalization probably never occurs: the mean free path of a photon in intergalactic space is larger than the observable Universe and is probably infinite.[5] In any case, even if thermalization does occur, it happens far in the future and at great distances from Earth and so is not available to drive evolution on Earth. For this reason we may ignore the existence of distant thermalizing matter in defining the system to which we apply the second law. The argument in Sec. IV, which concludes that inequality (1) is satisfied, would be strengthened if this extra entropy were included.[6]

III. EVOLUTIONARY DECREASE IN ENTROPY

We now consider $(dS/dt)_{\text{life}}$. Styer's argument relies on the assumption: "Suppose that, due to evolution, each individual organism is 1000 times 'more improbable' than the corresponding individual was 100 years ago. In other words, if Ω_i is the number of microstates consistent with the specification of an organism 100 years ago, and Ω_f is the number of microstates consistent with the specification of today's 'improved and less probable' organism, then $\Omega_f=10^{-3}\,\Omega_i$. I regard this as a very generous rate of evolution, but you may make your own assumption."[1]

The fact that no justification is provided for this assumption undermines the persuasive power of the argument. Moreover, far from being generous, a probability ratio of $\Omega_i/\Omega_f=10^3$ is probably much too low. One of the central ideas of statistical mechanics is that even tiny changes in a macroscopic object (say, one as large as a cell) result in exponentially large changes in the multiplicity (that is, the number of accessible microstates).

I will illustrate this idea by some order of magnitude estimates. First, let us address the precise meaning of the phrase "due to evolution." If a child grows up to be slightly larger than her mother due to improved nutrition, we do not describe this change as due to evolution, and thus we might not count the associated multiplicity reduction in the factor Ω_i/Ω_f. Instead we might count only changes such as the turning on of a new gene as being due to evolution. However, this narrow view would be incorrect. For this argument we should do our accounting in such a way that all biological changes are included. Even if a change such as the increased size of an organism is not the direct result of evolution for this organism in this particular generation, it is still ultimately due to evolution in the broad sense that all life is due to evolution. All of the extra proteins, DNA molecules, and other complex structures that are present in the child are there because of evolution at some point in the past if not in the present, and they should be accounted for in our calculation.

To see that this broad sense of evolution is the correct one to apply, consider a thought experiment. Suppose that the entropy reduction due to life in this broad sense were calculated and found to be greater than the entropy provided by sunlight. Creationists would justifiably cite this finding as proof that evolution is in conflict with the second law.

We now make some estimates of the required multiplicity reduction. We consider the case of an *E. coli* bacterium. We will first consider the reduction in multiplicity associated with building this organism from scratch. Following Styer,[1] we will then imagine a series of ever-simpler ancestors of this organism at 100 year intervals, stretching back over the 4 billion year history of evolution. Each organism is somewhat more improbable than its ancestor from the previous century, and the product of all of these multiplicity reductions must be sufficient to account for the total required multiplicity reduction.

The entropy reduction associated with the evolution of life comes in many forms, consisting of the construction of complex structures from simpler building blocks. For simplicity, we will consider just one portion of this process, namely, the construction of proteins from their constituent amino acids. Because we will neglect other processes (such as the synthe-

sis of the amino acids in the first place and the formation of other macromolecules), we will greatly underestimate the required multiplicity reduction.

An *E. coli* bacterium has about 4×10^6 protein molecules.[7] (This number refers not to the number of distinct types of protein, but to the total number of protein molecules in the cell.) We will find the multiplicity cost of assembling all of these molecules by first considering the multiplicity cost of assembling a single protein molecule. Imagining assembling the protein one amino acid at a time. At each step we must take an amino acid that was freely moving through the cell and place it in a specific position relative to the others that have already been assembled. If the amino acids were previously in a dilute solution in the cell, then the multiplicity loss due to each such step is approximately n_Q/n, where n is the number density of amino acids and n_Q is the density at which the amino acids would reach quantum degeneracy.[8] This ratio is large: amino acids in a cell are far from degenerate. To assemble a protein with N_a amino acids, we would repeat this process N_a-1 times, resulting in the large number $\Omega_i/\Omega_f\sim(n_Q/n)^{N_a-1}$. For instance, if $n_Q/n=10$ (much too low) and $N_a=300$ (about the average size of a protein[7]), the multiplicity ratio is $\sim10^{299}$ for the production of a single protein molecule.

If we use this conservative estimate for the multiplicity change associated with the formation of a single protein molecule, we estimate the multiplicity reduction required to assemble all of the proteins in the bacterium to be $\sim(10^{299})^{4\times10^6}\sim10^{10^9}$. If the entire 4 billion years (4×10^7 centuries) of biological evolution were required to achieve this number, we would require a multiplicity reduction of $(10^{10^9})^{1/(4\times10^7)}=10^{25}$ in each century, not 10^3.

These estimates are extremely rough. For example, they neglect the internal degrees of freedom of the protein (which are far fewer than those of the free amino acids), and entropy changes due to the energy absorbed or emitted during the formation of chemical bonds. To include the latter we note that the multiplicity change associated with a chemical reaction is $e^{\mu/kT}$. The chemical potential μ in a chemical reaction is of order 1 eV ($\sim10^{-19}$ J) or more, implying multiplicity changes of order $e^{40}\approx10^{17}$ for each chemical bond formed or broken at biological temperatures. Because hundreds of chemical bonds must be formed in assembling each protein molecule, the resulting factor will again be exponentially large. This sort of number is the ante to enter this particular game.

Because bacteria appeared very early in evolution, we should assume a time period much shorter than 4 billion years, and hence a still larger multiplicity reduction per century would be required. Similarly, if we considered more than just the formation of proteins, or if we considered a large multicellular organism, the required factor would be much greater.

Rough as these arguments are, they establish that there is reason to doubt the factor 10^3 that plays an important role in Styer's argument, rendering his argument unpersuasive.[9] To strengthen the argument we should set a robust upper limit on $|(dS/dt)_{\text{life}}|$, or equivalently on the total entropy reduction $|\Delta S_{\text{life}}|$, in a way that does not depend on such an assumption.

IV. A ROBUST ARGUMENT

Let us establish an upper limit on $|\Delta S_{\text{life}}|$ by estimating the relevant quantities in a way that is certain to overestimate the final result. Consider the entropy difference between two systems: Earth as it is at present, and a hypothetical dead-Earth on which life never evolved. We will assume that dead-Earth and Earth are identical, except that every atom in Earth's biomass is located in dead-Earth's atmosphere in its simplest molecular form. When considering the entropy of Earth, we will assign zero entropy to the biomass. That is, we will imagine that to turn dead-Earth into Earth, it is necessary to pluck every atom required for the biomass from the atmosphere and place it into its exact present-day quantum state. These assumptions maximize the entropy of dead-Earth and minimize that of Earth, so the difference between the two entropies grossly overestimates the required entropy reduction for the production of life in its present form.

With these assumptions, we can use the standard thermodynamic result $\mu/T=-\partial S/\partial N$, which implies $\Delta S=-N\mu/T$, to estimate the entropy difference as

$$\Delta S_{\text{life}}=S_{\text{earth}}-S_{\text{dead-earth}}\approx\frac{N_b\mu}{T}, \tag{4}$$

where N_b is the number of molecules required to make up the biomass and μ is a typical chemical potential for a molecule in the atmosphere of dead-Earth. If we use standard relations for an ideal gas,[10] we find $\mu/kT\sim-10$, so that $\Delta S_{\text{life}}<0$ as expected. We can obtain a value for N_b from the estimate[11] that the total carbon biomass of Earth is $\sim10^{15}$ kg. Even if we increase this value by a generous factor of 100 to account for other elements, we still have fewer than 10^{43} molecules. We conclude that the entropy reduction required for life on Earth is far less than

$$|\Delta S_{\text{life}}|\sim10^{44}k. \tag{5}$$

If we compare this value with the rate of entropy production due to sunlight in Eq. (3), we find that the second law, in the form of Eq. (1), is satisfied as long as the time required for life to evolve on Earth is at least

$$\Delta t=\frac{|\Delta S_{\text{life}}|}{(dS/dt)_{\text{sun}}}\sim10^7\ \text{s}, \tag{6}$$

or less than a year. Life on Earth took 4 billion years to evolve, so the second law of thermodynamics is safe.[12]

ACKNOWLEDGMENTS

The author thanks Andrew Bunn, H. Franklin Bunn, and Ovidiu Lipan for helpful discussions. Two anonymous referees provided very useful comments that significantly sharpened the arguments in this article.

[a]Electronic address: ebunn@richmond.edu

[1]D. F. Styer, "Entropy and evolution," Am. J. Phys. **76**, 1031–1033 (2008).

[2]The TalkOrigins archive has a list of resources on the subject at ⟨www.talkorigins.org/faqs/thermo.html⟩.

[3]D. Labs and H. Neckel, "The solar constant (a compilation of recent measurements)," Sol. Phys. **19**, 3–15 (1971).

[4]P. R. Goode, J. Qiu, V. Yurchyshyn, J. Hickey, M.-C. Chu, E. Kolbe, C. T. Brown, and S. E. Koonin, "Earthshine observations of the Earth's reflectance," Geophys. Res. Lett. **28**, 1671–1674 (2001).

[5]E. W. Kolb and M. S. Turner, *The Early Universe* (Addison-Wesley, Reading, MA, 1990), p. 354.

[6]The reader who wishes to compare this article with that of Ref. 1 may find it helpful to note that Ref. 1 calculates both $(dS/dt)_{\text{sun}}$ and

$(dS/dt)_{\mathrm{CMB}}$. In the quantitative conclusion [Eqs. (5) and (6)], Ref. 1 correctly uses $(dS/dt)_{\mathrm{sun}}$, referring to this quantity as the "entropy throughput."

[7] According to B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson, *Molecular Biology of the Cell* (Garland, New York, 1994), 3rd ed., p. 90, an *E. coli* bacterium has a volume of 2×10^{-12} cm^3 and is about 15% protein by mass. If we assume that the density of the bacterium is that of water, the total mass of protein is 3×10^{-13} g. The molecular weights of amino acids are about 150 g/mol, and thus the total number of amino acids in these proteins is $\approx1.2\times10^{9}$. An average protein contains 300 amino acids per protein (p. 118), yielding 4×10^{6} proteins in the bacterium.

[8] Imagine that there are n_a amino acids in solution, with N available quantum states. Nondegeneracy means that $N\gg n_a$. The multiplicity is $\Omega(n_a)=\binom{N}{n_a}$. Taking one molecule out of solution causes the multiplicity to decrease by a factor $\Omega(n_a)/\Omega(n_a-1)=(N-n_a+1)/n_a\approx N/n_a=n_Q/n$.

[9] If we adopt a much narrower view of which changes are due to evolution, these estimates would not apply. Even in this case, the factor of 10^3 would still lack justification.

[10] See, for example, Eq. (3.63) in D. V. Schroeder, *An Introduction to Thermal Physics* (Addison-Wesley Longman, Reading, MA, 2000).

[11] W. B. Whitman, D. C. Coleman, and W. J. Wiebe, "Prokaryotes: The unseen majority," Proc. Natl. Acad. Sci. U.S.A. **95**, 6578–6583 (1998).

[12] Creating all of life in 6 days might be thermodynamically problematic, however.

The Proca equations derived from first principles

Michel Gondran[a)]
University Paris Dauphine, Lamsade, 75 016 Paris, France

(Received 25 January 2009; accepted 27 April 2009)

We show how the Proca equations can be deduced from first principles in a way that is similar to the derivations of the Dirac relativistic electron equation and Maxwell equations. © *2009 American Association of Physics Teachers.*

[DOI: 10.1119/1.3137042]

I. INTRODUCTION

The Schrödinger and Klein–Gordon equations can be directly deduced from first principles using the relation between the energy E, mass m, and momentum $\vec{\mathbf{p}}$, and the correspondence principle where E and $\vec{\mathbf{p}}$ are replaced by the operators $i\hbar\partial/\partial t$ and $-i\hbar\vec{\nabla}$. The nonrelativistic result $(E-(\vec{\mathbf{p}}^2/2m)-V)\Psi=0$, where ψ is a wave function, gives the Schrödinger equation, and the relativistic result $(E^2-c^2\vec{\mathbf{p}}^2-m^2c^4)\Psi=0$ gives the Klein–Gordon equation,

$$\left[\left(\frac{i\hbar\partial}{\partial t}\right)^2-c^2(-i\hbar\nabla)^2-m^2c^4\right]\Psi=0. \tag{1}$$

The Dirac relativistic equation[1] is also derived from first principles, but indirectly by factorization of Eq. (1), with a four component wave function Ψ, which leads to the introduction of spin $\frac{1}{2}$. Gersten showed[2] how the Maxwell equations can be obtained by the factorization of Eq. (1), with $m=0$ and a three component wave function Ψ, which is interpreted as corresponding to a spin 1 particle.

The aim of this note is to show how the Proca equations,[3,4] which describe a massive spin 1 particle, can also be derived from Eq. (1) using a factorization similar to that for Maxwell's equations. The Proca equations can be used for massive photons[4,5] and for the London penetration depth in a superconductor.[4,6] Contrary to Maxwell's equations, it is necessary to introduce a potential to transform a second-order differential equation, such as the Klein–Gordon equation, into a first-order differential equation such as the Proca equations.

II. DIRAC AND MAXWELL EQUATIONS

In his seminal 1928 paper,[1] Dirac decomposed the relativistic energy-momentum relation corresponding to Eq. (1) with a four component column (bispinor) wave function Ψ into

$$\left[EI^{(4)}+\begin{pmatrix}mc^2I^{(2)} & c\vec{\mathbf{p}}\cdot\vec{\sigma}\\ c\vec{\mathbf{p}}\cdot\vec{\sigma} & -mc^2I^{(2)}\end{pmatrix}\right]\times\left[EI^{(4)}-\begin{pmatrix}mc^2I^{(2)} & c\vec{\mathbf{p}}\cdot\vec{\sigma}\\ c\vec{\mathbf{p}}\cdot\vec{\sigma} & -mc^2I^{(2)}\end{pmatrix}\right]\Psi=0, \tag{2}$$

where $I^{(4)}$ is the 4×4 unit matrix, $I^{(2)}$ is the 2×2 unit matrix, and $\vec{\sigma}=(\sigma_x,\sigma_y,\sigma_z)$ is the Pauli spin one-half vector matrix.

Equation (2) is satisfied if the equation,

$$\left[EI^{(4)}-\begin{pmatrix}mc^2I^{(2)} & c\vec{\mathbf{p}}\cdot\vec{\sigma}\\ c\vec{\mathbf{p}}\cdot\vec{\sigma} & -mc^2I^{(2)}\end{pmatrix}\right]\Psi=0 \tag{3}$$

is satisfied. The Dirac equation is obtained by replacing E and $\vec{\mathbf{p}}$ in Eq. (3) with the operators $i\hbar\partial/\partial t$ and $-i\hbar\vec{\nabla}$.

In 1999 Gersten[2] decomposed the free photon relativistic equation into

$$\left(\frac{E^2}{c^2}-\vec{\mathbf{p}}^2\right)I^{(3)}\vec{\Psi}=\left[\frac{E}{c}I^{(3)}-\vec{\mathbf{p}}\cdot\vec{\mathbf{S}}\right]\left[\frac{E}{c}I^{(3)}+\vec{\mathbf{p}}\cdot\vec{\mathbf{S}}\right]\vec{\Psi}-\begin{pmatrix}p_x\\ p_y\\ p_z\end{pmatrix}(\vec{\mathbf{p}}\cdot\vec{\Psi})=0, \tag{4}$$

where $I^{(3)}$ is the 3×3 unit matrix, $\vec{\Psi}$ is a three component

column wave function, and $\vec{\mathbf{S}}$ is a spin one vector matrix with components

$$S_x=\begin{pmatrix}0&0&0\\0&0&-i\\0&i&0\end{pmatrix},\quad S_y=\begin{pmatrix}0&0&i\\0&0&0\\-i&0&0\end{pmatrix},$$
$$S_z=\begin{pmatrix}0&-i&0\\i&0&0\\0&0&0\end{pmatrix}, \tag{5}$$

and the properties

$$[S_x,S_y]=iS_z,\quad [S_z,S_x]=iS_y,\quad [S_y,S_z]=iS_x,\quad \vec{\mathbf{S}}^2=2I^{(3)}. \tag{6}$$

Equation (4) will be satisfied if the two equations,

$$\left[\frac{E}{c}I^{(3)}+\vec{\mathbf{p}}\cdot\vec{\mathbf{S}}\right]\vec{\Psi}=0, \tag{7}$$

$$\vec{\mathbf{p}}\cdot\vec{\Psi}=0, \tag{8}$$

are simultaneously satisfied. Maxwell's equations are obtained by replacing E and $\vec{\mathbf{p}}$ in Eqs. (7) and (8) with the operators $i\hbar\partial/\partial t$ and $-i\hbar\vec{\nabla}$, and the wave function with $\vec{\Psi}=\vec{\mathbf{E}}-i\vec{\mathbf{B}}$, where $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$ are the electric and magnetic fields, respectively. With the identity $(-i\hbar\vec{\nabla}\cdot\vec{\mathbf{S}})\vec{\Psi}=\hbar\vec{\nabla}\times\vec{\Psi}$, Eqs. (7) and (8) become

$$i\frac{\hbar}{c}\frac{\partial(\vec{\mathbf{E}}-i\vec{\mathbf{B}})}{\partial t}=-\hbar\vec{\nabla}\times(\vec{\mathbf{E}}-i\vec{\mathbf{B}}), \tag{9}$$

$$\vec{\nabla}\cdot(\vec{\mathbf{E}}-i\vec{\mathbf{B}})=0, \tag{10}$$

which are the Maxwell equations if the electric and magnetic fields are real.

III. PROCA EQUATIONS

In his 1936 paper,[3] Proca found relativistic wave equations for a massive spin 1 particle. We can derive these equations from the relativistic condition on the energy E, mass m, and momentum $\vec{\mathbf{p}}$,

$$\left(\frac{E^2}{c^2}-\vec{\mathbf{p}}^2-m^2c^2\right)I^{(3)}\vec{\Psi}=0, \tag{11}$$

where $\vec{\Psi}$ is a three component column wave function. Equation (11) can be decomposed into

$$\left[\frac{E}{c}I^{(3)}-\vec{\mathbf{p}}\cdot\vec{\mathbf{S}}\right]\left[\frac{E}{c}I^{(3)}+\vec{\mathbf{p}}\cdot\vec{\mathbf{S}}\right]\vec{\Psi}-\begin{pmatrix}p_x\\p_y\\p_z\end{pmatrix}(\vec{\mathbf{p}}\cdot\vec{\Psi})-m^2c^2\vec{\Psi}=0, \tag{12}$$

where $\vec{\mathbf{S}}$ is a spin one vector matrix defined by Eqs. (5) and (6).

If $m\neq 0$, there exists a 4-potential $(\varphi,\vec{\mathbf{A}})$ defined by the equations,

$$\vec{\mathbf{p}}\cdot\vec{\Psi}=i\frac{m^2c^2}{\hbar}\varphi, \tag{13}$$

$$\left[\frac{E}{c}I^{(3)}+\vec{\mathbf{p}}\cdot\vec{\mathbf{S}}\right]\vec{\Psi}=i\frac{m^2c^2}{\hbar}\vec{\mathbf{A}}. \tag{14}$$

If we use these two equations in Eq. (12) and substitute E and $\vec{\mathbf{p}}$ by the operators $i\hbar\partial/\partial t$ and $-i\hbar\vec{\nabla}$, we find the equation that describes the wave function in terms of the potential

$$\vec{\Psi}=-\frac{1}{c}\frac{\partial\vec{\mathbf{A}}}{\partial t}-\vec{\nabla}\cdot\varphi-i\vec{\nabla}\times\vec{\mathbf{A}}\equiv\vec{\mathbf{E}}-i\vec{\mathbf{B}}. \tag{15}$$

Other equations corresponding to Eqs. (13) and (14) are obtained by replacing E and $\vec{\mathbf{p}}$ with the operators $i\hbar\partial/\partial t$ and $-i\hbar\vec{\nabla}$

$$\vec{\nabla}\cdot\vec{\Psi}=-\frac{m^2c^2}{\hbar^2}\varphi, \tag{16}$$

$$\frac{1}{c}\frac{\partial\vec{\Psi}}{\partial t}-i\vec{\nabla}\times\vec{\Psi}=\frac{m^2c^2}{\hbar^2}\vec{\mathbf{A}}. \tag{17}$$

We deduce the Proca equations from Eqs. (15)–(17) if $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$ are real fields

$$\vec{\nabla}\cdot\vec{\mathbf{E}}=-\frac{m^2c^2}{\hbar^2}\varphi,\quad \vec{\nabla}\cdot\vec{\mathbf{B}}=0, \tag{18}$$

$$\frac{1}{c}\frac{\partial\vec{\mathbf{E}}}{\partial t}-\vec{\nabla}\times\vec{\mathbf{B}}=\frac{m^2c^2}{\hbar^2}\vec{\mathbf{A}},\quad \frac{1}{c}\frac{\partial\vec{\mathbf{B}}}{\partial t}+\vec{\nabla}\times\vec{\mathbf{E}}=0. \tag{19}$$

IV. CONCLUSION

In this note it has been shown that the set of Proca equations can be derived from first principles in a way similar to that used to find the Schrödinger, Klein–Gordon, Dirac, and Maxwell equations.

As demonstrated by Lévy-Leblond,[7] the origin of the electron spin in Dirac's equation is not relativistic but is due to the use of multicomponent wave functions. It is the same for the Maxwell and Proca equations.

Because $m\neq 0$, a 4-potential $(\varphi,\vec{\mathbf{A}})$ exists in Proca's equation, which is not allowed for the photon case. This potential is the source term that defines the wave function.

[a]Electronic mail: michel.gondran@polytechnique.org

[1] P. Dirac, "The quantum theory of the electron," Proc. R. Soc. London **A117**, 610–624 (1928).
[2] A. Gersten, "Maxwell equations as the one-photon quantum equation," Found. Phys. Lett. **12**, 291–298 (1999).
[3] A. Proca, "Sur la théorie ondulatoire des électrons positifs et négatifs," J. Phys. Radium **7**, 347–353 (1936).
[4] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).
[5] T. Prokopec and R. Woodard, "Vacuum polarization and photon mass in inflation," Am. J. Phys. **72**, 60–72 (2004).
[6] L. H. Ryder, *Quantum Field Theory* (Cambridge U. P., Cambridge, 2003).
[7] J. M. Lévy-Leblond, "Nonrelativistic particles and wave equations," Commun. Math. Phys. **6**, 286–311 (1967).

A simple design for an experiment to determine the Lorentz force on a current-carrying conductor in a magnetic field

C. M. Romo-Kröger[a)]
Basic Sciences Institute, Catholic University of Maule, Casilla 617, Talca, Chile

(Received 25 January 2009; accepted 28 April 2009)
[DOI: 10.1119/1.3138699]

A popular experiment to determine the Lorentz force on a current-carrying conductor in a magnetic field requires the use of a balance.[1–3] An alternative apparatus is presented, which gives a direct determination of the Lorentz force on a segment of wire in the interior of a solenoid. Although the apparatus uses the gravitational force, it does not require the use of a balance. The experiment allows students to gain more experience with the magnetic field and magnetic force.[4]

Figure 1 is a schematic of the experimental setup. A trapeze was constructed of 1/16 in. diameter varnished copper wire, which is commonly used to make electrical coils. The trapeze can be previously prepared or constructed by students in the laboratory. Figure 1 shows the trapeze hanging from two horizontal pins with its horizontal lower segment inside a solenoid. A panoramic view of the experimental setup is shown in Fig. 2.

The upper ends of the rectangular trapeze should be stripped for good electrical contact. The two pins from which the trapeze hangs and where two nippers connect with the rest of the circuit are nailed to two vertical wood columns. The pins are horizontal, one in front of the other. A press tightens the system, with the two columns and the solenoid in between. The whole system is fixed to the worktable by another press. The internal cavity of the solenoid contains four plastic walls that form an interior space of square cross section. Within this cavity is the lower segment of the trapeze. It can move with a pivot around the pins. When the lower segment touches one wall, the distance x that it travels before touching the other wall is equal to the wall separation.

To avoid exceeding the maximum allowable current, two separate power supplies were used for the hanging trapeze and the solenoid. Ordinary ammeters were used to measure the current in the trapeze and in the solenoid.

The following measurements were obtained in a preliminary experiment. The currents through the trapeze and solenoid (number of turns $N=800$ and length $L=4$ cm) were $I_0=0.8$ and $I=4A$, respectively. I_0 is the current through the trapeze when the lower segment barely touches the opposite wall due to the magnetic force. The distance between the internal walls in the solenoid is $x=2$ cm. The length of the rectangular trapeze is $l=21.5$ cm and its lower segment is $l_0=1.4$ cm. The total mass of the trapeze is 4.24 g and consists of a longitudinal part with $m=3.8$ g, a transverse segment with $m_0=0.24$ g, and hangers with mass of 0.20 g. The following calculations were performed with these data.

The magnetic field inside the solenoid is parallel to the vertical axis and its magnitude is

$$B=\frac{\mu_0 NI}{L}, \tag{1}$$

where μ_0 is the magnetic permeability in air. The magnitude of the Lorentz force for the current-carrying segment l_0 is

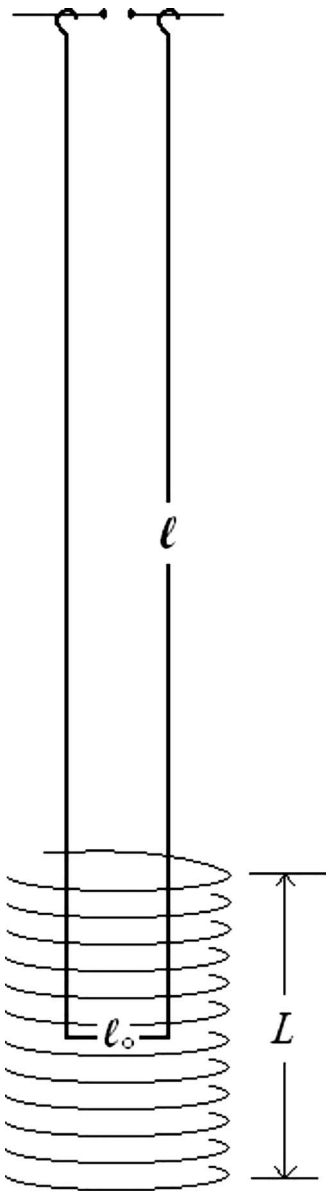

Fig. 1. The hanging trapeze of length l with its horizontal segment of length l_0 at the interior of the solenoid (length L).

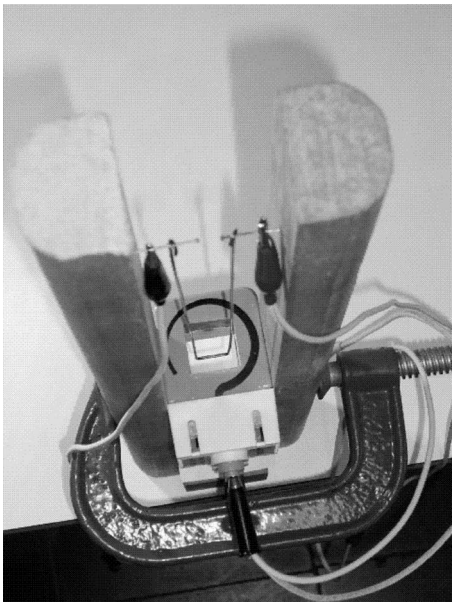

Fig. 2. Photograph of the experimental apparatus showing the pins where the trapeze hangs and where the electrical connections are made. The lower horizontal segment of the wire is visible inside the solenoid and is free to move between the interior walls of the solenoid. The presses that hold the entire system are also seen.

$$F_M = I_0 l_0 B, \tag{2}$$

and its direction is horizontal and perpendicular to l_0. The magnitude of F_M is found to be $F_M = 112$ dyn. This force generates a torque τ_M on the trapeze.

The vertical gravitational forces lead to a torque τ_G that opposes the torque τ_M. In Fig. 3(a) the trapeze is viewed on edge with the magnetic and gravitational forces acting on it. The torques on the trapeze are due to F_M and the force $m_0 g$ at a distance l from the pivot, and to mg at a distance $l/2$. In equilibrium, the sum of the torques is zero so that

$$mg(l/2)\sin\theta + m_0 g l \sin\theta - F_M l \sin(90° - \theta) = 0. \tag{3}$$

A corresponding calculation leads to the magnetic force of $F_M = 97.6$ dyn. Thus, a value of F_M is obtained from the Lorentz force in Eq. (2) and another value is deduced from the balance of torques in Eq. (3). A comparison of the two values of F_M gives an error of less than 13%, which demonstrates that the method works and could be refined by more accurate measurements.

Many questions can be asked at the completion of this experiment. For example, on the basis of the Lorentz force, are the lateral parts of the trapeze subject to forces? Does the angle of inclination cause a force on the lateral parts? How much do the horizontal components of the B field above the solenoid affect the measurements? What percentage of the mass attributable to the hangers of the trapeze is needed for the correct determination of the torques? What are the sources of uncertainty?

Errors can be discussed as follows. The current I_0 at which the trapeze reaches the opposite wall can be determined with a little practice. By increasing the current slowly, a subtle bounce on the wall is produced, which lets the magnitude of I_0 be found accurately. When the current I_0 flows, the trapeze is inclined at an angle θ. Then the vertical field B produces forces in the lateral branches, but these forces cancel each other. In contrast, the horizontal component of the B field above the solenoid (fringe field) produces a torque that opposes the torque at the lower segment. This contribution is missing in Eq. (3). Thus the value we calculated, 97.6 dyn, is due to the torque of the field B acting on the lower segment minus this new torque. This effect can be seen in Fig. 3(b) bytaking the cross product $\vec{I}_0 \times \vec{B}$ in each section of the trapeze. Although the field strength is very low outside the solenoid, the fringe field introduces a systematic error, which can be reduced by the use of a solenoid with a length similar to the length of the trapeze, wrapping it almost entirely. A sheet of iron could also be used as a magnetic shield. The mass of the hangers affects the determination of the gravitational torque. In addition, an uncertainty in the position of the pivot occurs when the trapeze rotates around the pins. These sources of error should be analyzed by the students.

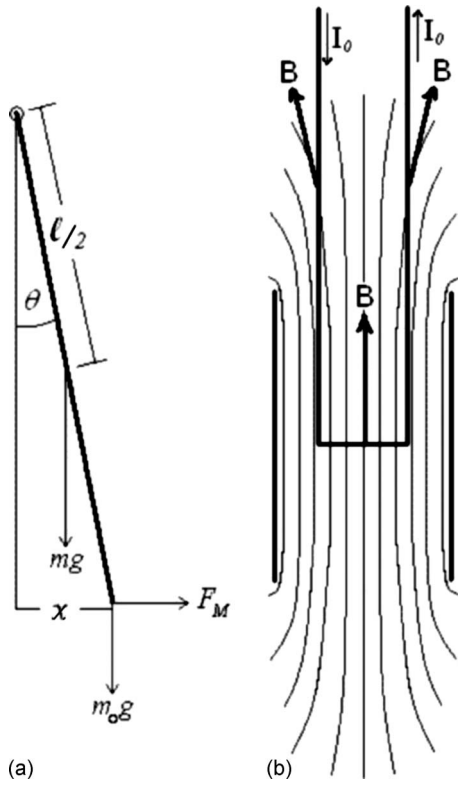

Fig. 3. (a) Force diagram for the trapeze suspended from its upper part. At the angle θ the total torque on the trapeze is zero due to the actions of the magnetic force F_M and the gravitational forces $m_0 g$ at the distance l and mg at the distance $l/2$. (b) Magnetic field lines inside and near the solenoid showing that the horizontal component of the B field produces a torque that opposes the torque at the interior segment.

Several improvements in the apparatus can be made. Students can experiment with changing the currents in the trapeze and in the solenoid. They can use solenoids with different characteristics (number of turns, length, and diameter). It is good to have made wire trapezes of different geometries. The rectangular trapeze could be constructed as a multiple loop to increase the force with the same current, which may lead to more accurate results. The lower segment of the trapeze could be made of different lengths and could also form an angle with respect to the horizontal to study the dependencies on the length that appears in Eq. (2) and the angle implicit in the Lorentz force.

ACKNOWLEDGMENTS

The contributions of students Ferid Hasbun and Roger Mendez in the design of the experiment are recognized. The contributions of anonymous referees are also recognized.

[a]Electronic mail: cromo@ucm.cl

[1]PSSC, Physical Science Study Committee, *Laboratory Guide for Physics* (D. C. Heath and Company, Boston, MA, 1960).

[2]PASCO Physics 2008–2009, *Catalog and Experiment Guide* (PASCO Scientific, Roseville, CA, 2008).

[3]PHYWE, Laboratory Experiments, Physics Experiment No. 4.1.06, www.nikhef.nl/~h73/kn1c/praktikum/phywe/LEP/Experim/4_1_06.pdf.

[4]R. Chabay and B. Sherwood, "Restructuring the introductory electricity and magnetism course," Am. J. Phys. **74**, 329–336 (2006).

Comment on "A 'local observables' method for wave mechanics applied to atomic hydrogen," by Peter J. Bowman [Am. J. Phys. 76 (12), 1120–1129 (2008)]

Hans C. Ohanian[a)]

Department of Physics, University of Vermont, Burlington, Vermont 05405

(Received 9 March 2009; accepted 26 March 2009)

[DOI: 10.1119/1.3119514]

In a baffling paper[1] Bowman advocates that we abandon operators in quantum mechanics and calculate the angular momentum and magnetic moment of electron states directly from the electron currents in the wave fields, treating these as though they were classical quantities. He performs such a calculation for the states of the hydrogen atom and makes the bizarre claim that the total angular momentum of the ground state is $\hbar$, rather than $\hbar/2$.

It is obvious that this result must be wrong. The ground state wave function is spherically symmetric in its dependence on the spatial coordinates, so the angular momentum has no orbital contribution and must be entirely attributed to the electron's spin. If the angular momentum of the ground state were $\hbar$, then the spin of the electron would have to be $\hbar$;[2] that is, the electron would have to be what we conventionally call a particle of spin 1. This spin value for the electron would lead to disastrous mathematical and physical consequences. It would mean that Bowman's calculation is self-contradictory, because he uses the Dirac equation appropriate to spin 1/2 and not the Proca equation[3] required for a massive particle of spin 1. The spectral lines of hydrogen and alkali atoms would display fine-structure triplets, instead of the observed doublets. Additionally, the gyromagnetic ratio of the electron would be 1, instead of the observed value of 2.00…, which is found from measurements of the electron magnetic resonance frequency.

If Bowman's result is wrong, where is the mistake in his calculation? It turns out that the mistake is elementary. He attributes the angular momentum to a circulatory flow of mass, obtained by multiplying the Pauli electron current density $\mathbf{k}_{\text{pauli}}$ by the mass,

$$m_e \mathbf{k}_{\text{pauli}} = \frac{\hbar}{2}[2\,\text{Im}(\psi^\dagger \nabla \psi) + \nabla \times (\psi^\dagger \sigma \psi)] \tag{1a}$$

$$= \frac{\hbar}{2i}[\psi^\dagger \nabla \psi - (\nabla \psi^\dagger)\psi] + \frac{\hbar}{2} \nabla \times (\psi^\dagger \sigma \psi). \tag{1b}$$

The Pauli electron current density is the nonrelativistic limit of the Dirac electron current density, and the use of a nonrelativistic approximation in hydrogenlike atoms (in which the electron speeds are as large as $c/137$, and even larger for atoms of high Z) is questionable. However, Bowman's fatal mistake is not this nonrelativistic approximation, but a confusion over the source of angular momentum. The electron current density (or some current density proportional to it) is not the source of angular momentum. Instead, the source of angular momentum is the momentum density, which is given by the T^{0k} components of the energy-momentum tensor (what in electromagnetic theory we would call the Poynting vector). *This momentum density is not equal to the electron current density multiplied by the mass.*

In a paper published in 1986 (which Bowman references, but evidently chose to ignore),[4] I performed a calculation similar to Bowman's, with the goal of providing an intuitive interpretation of the spin. In this calculation I used the correct (and relativistically exact) momentum density $\mathbf{G}$ for the Dirac field, extracted from the symmetrized energy-momentum tensor:[5]

$$\mathbf{G} = \frac{\hbar}{2i}[\psi^\dagger \nabla \psi - (\nabla \psi^\dagger)\psi] + \frac{\hbar}{4} \nabla \times (\psi^\dagger \sigma \psi). \tag{2}$$

The magnetic moment is the volume integral of $-\frac{1}{2}e\mathbf{x} \times \mathbf{k}_{\text{pauli}}$, whereas the angular momentum is the volume inte-

gral of $\mathbf{x}\times\mathbf{G}$. Comparison of Eqs. (1b) and (2) shows that the factor multiplying the spin term in Eq. (1b) is twice as large as in Eq. (2). Because of this extra factor of 2, the electron spin contributes twice as much to the magnetic moment as it contributes to the angular momentum. This difference accounts for the nonclassical gyromagnetic ratio associated with the electron spin. Hence Bowman's calculation of the magnetic moment of the ground state of the hydrogen atom is correct, but his calculation of the angular momentum is off by a factor of 2. There are similar mistakes in his calculations for all the other states.

Bowman proposes that a direct measurement of the angular momentum of the ground state of hydrogen could serve as an *experimentum crucis* that discriminates between his "local observables" scheme and standard quantum mechanics. My correction of his calculation shows that such a measurement would do no such thing—the angular momentum of the ground state is $\hbar/2$, no matter which way we calculate it. Although it is instructive to calculate physical quantities by integration over the wave fields, these calculations (when done correctly) will always agree with the results obtained by the operator methods of standard quantum mechanics. Thus, there is no need for a reformulation of quantum mechanics.

ACKNOWLEDGMENT

The author would like to acknowledge cordial and helpful correspondence from Peter Bowman about this comment.

[a]Electronic mail: hohanian@uvm.edu

[1]P. Bowman, "A 'local observables' method for wave mechanics applied to atomic hydrogen," Am. J. Phys. **76**, 1120–1129 (2008).

[2]This result can be easily confirmed by applying Bowman's calculation to a free-electron wave function.

[3]G. Wentzel, *Quantum Theory of Fields* (Interscience, New York, 1949), p. 75.

[4]H. C. Ohanian, "What is spin?," Am. J. Phys. **54**, 500–505 (1986).

[5]The symmetrized energy-momentum tensor for the Dirac field is given by G. Wentzel, *op. cit.*, p. 170. Wentzel uses a slightly different notation, with an extra factor of $i\beta$ included in $\psi^\dagger$. The tensor given by Wentzel can be rewritten in the form of Eq. (2) (the manipulations in Ref. 4 used the Dirac equation for a free electron, but it is easy to check that Eq. (2) remains valid when the Dirac equation includes electromagnetic fields). Symmetrization of the energy-momentum tensor was first proposed by F. J. Belinfante, "On the spin angular momentum of mesons," Physica (Utrecht) **6**, 887–898 (1939). If instead of Belinfante's symmetric energy-momentum tensor, we adopt an asymmetric tensor such that the spin cannot be attributed to the moment of the momentum density, then an extra spin density must be added "by hand" to account for all or part of the electron spin. In essence, this spin density means that the electron is regarded as a spinning pointlike entity with an "internal" angular momentum (somewhat a là Goudsmit and Uhlenbeck). Such a description of the spin is strongly favored by some physicists who wish to use the spin density as the source for a torsion of space time.