Relativity and aberration

Thomas E. Phipps, Jr.
908 South Busey Avenue, Urbana, Illinois 61801

(Received 9 March 1988; accepted for publication 11 July 1988)

Certain physics textbooks convey a false impression that the stellar aberration angle depends on the relative velocity of the light source and observer. Counterfactual consequences are deduced from such an assumption and a more acceptable alternative, consistent with the relativity principle and with some other textbooks, is discussed.

I. INTRODUCTION

The student of special relativity encounters a bewildering array of physical conceptions concerning the phenomenon of stellar aberration. Some textbooks play it safe by saying almost nothing about the physics; they are content to mention the subject purely as an illustration of Lorentz transformation properties of the light propagation vector. Other textbooks bring the Sun into the discussion—although it must be obvious that what is observed on Earth cannot depend in an essential way on what might be observable on the Sun. Thus one textbook[1] refers to "an observer on the sun," another[2] states that "We have to compare the direction of the incoming light with respect to two frames of reference, that of the sun and that of the earth," while a third[3] asks us to "Consider the inertial frame in which the sun and the star are at rest" (surely a neat trick).

At worst, such expositions leave students slightly baffled; at best, students are so thoroughly confused as to be forced to think for themselves. Much more pernicious are those textbooks that leave them flatly misinformed, specifically, those that assert or imply stellar aberration to be a consequence of the *relative motion of the light source and observer*. Thus Arzeliès states,[4] "A star emits light which is received by a telescope on the Earth. We regard the star and the Earth as being at rest in two Galilean systems K_2 and K_1 respectively; K_2 has a velocity v with respect to K_1." Tonnelat states that[5] "Aberration phenomena... are explained by a modification in the direction of light rays as a result of the relative source–observer motion." Møller states that[6] "... we take S and S' to denote the systems of coordinates in which the fixed stars and the earth respectively are at rest." (This is another neat trick since "fixed stars" are a medievalism without counterpart in our restless modern universe. Apparently, Møller means to station the light source in one inertial system and the observer in another system.)

The purpose of the present article is to expose and deal with the physical misconception that relative motion between radiation source (star, emitter) and sink (observer, detector) is responsible for observed stellar aberration. The accepted first-order formula, common to both classical and relativistic treatments, is that aberration angle $= v/c + O(v^2/c^2)$ rad and involves a parameter v interpreted as the component of *some* relative velocity transverse to the observer's line of sight. No formula becomes part of science (shared knowledge) until all its parameters acquire agreed physical interpretations. We shall presently—through an operational definition of v in this formula—seek a basis for consensus concerning what relative velocity is physically involved.

Recognition of the fallacy in supposing source–sink relative motion to be responsible for stellar aberration is far from new. The earliest explicit mention (of which the present author is aware) of the conflict of such an interpretation with observation is that by Herbert Ives,[7] who noted in 1950 that "The idea sometimes met with that aberration... may be described in terms of the relative motions of the bodies concerned, is immediately refuted by the existence of spectroscopic binaries with velocities comparable with that of the earth in its orbit. These exhibit aberrations no different from other stars."

In other words, if it were true that the relative velocity of the source and detector affects aberration, the observed phenomenon would differ for differently moving sources. The intuitive idea that the vast *distance* of stellar light sources from the Earth—which causes them to be apparently "fixed" on the celestial sphere—can explain the identity of the aberration constant for all such sources is refuted by the facts that (a) distance between the source and observer is not parametrized in any theory (either classical or relativistic) and (b) if distance mattered at all, the tremendously varying actual distances of different stars from Earth would result in variable, not constant, aberration as observed from Earth.

This article will (a) reiterate Ives's[7] point for the benefit of a new generation of physicists; (b) emphasize the gravity of the problem through approximate quantification; and (c) point out an alternative interpretation of the v parameter, given in some textbooks, which agrees with the facts of observation.

II. CONSEQUENCES OF THE SOURCE–SINK-RELATIVE-MOTION INTERPRETATION

First, to address the Ives contention, consider a highly idealized binary system consisting of a point light source in circular orbit of radius R and period T about some center of attraction at rest in an inertial system S. Suppose that (a) the observer is at rest in the same inertial system S; (b) his line of sight is normal to the orbital plane of the source; and (c) his distance from that plane is D, with D much greater than R. According to the fallacious interpretation of the relative-velocity parameter v mentioned above, the observable aberration will be of period T (since the source–observer relative velocity is of that period) and will, at first order, amount to an image angular displacement (angle of "telescope tilt") of v/c rad, where $v = 2\pi R/T$; this is the magnitude of the component of relative velocity between the source and observer transverse to the line of sight.

The angular displacement just mentioned is such that the image on the observer's celestial sphere at distance D traces out an orbit circle of enlarged radius R' centered on the true orbit of radius R. By similar triangles we have, at first order, $R'/D = v/c$, from which an orbit magnification factor $M = R'/R = 2\pi D/cT$ follows. This magnification factor can be huge if D is large and T is small. Although the observer's telescope might not resolve the true orbit of radius R (as in the "spectroscopic binary" case), the apparent orbit radius R' could be so large as to make the source's apparent motion superluminal. Since only a given number of photons is emitted by the source per period, and these must be distributed over an enlarged apparent orbit, the source light intensity is reduced by the factor M.

If the plane of the circular orbit just described is inclined to the line of sight, the orbit figure hypothetically produced by aberration of the sort just discussed becomes more complicated and may approximate a distorted figure eight. It will be apparent that such effects, due to proper motions (implying observer-relative motions) of astronomical point sources, have not been observed. Ives's[7] remark is thus confirmed. However, the confrontation with empiricism arises even more directly since incoherent light sources are thought to be *individual atoms* (or plasma constituents)—and these in stellar atmospheres are known to be in continual random high-speed motion. Hence even a star at rest in the same inertial system as the observer's telescope would exhibit visible image-spreading aberrations due to transverse motions of light-emitting source atoms if it were correct to attribute aberration to relative motion of the light source and observer.

Let us quantify this last point: Equilibrium gas theory associates the kinetic energy $kT/2$ with each particulate degree of freedom. Transverse to the observer's line of sight there are two translatory degrees of freedom of each gas atom (or particulate constituent) of the luminous stellar envelope. The transverse motional kinetic energy of an atom of mass m is therefore $kt = mv_t^2/2$. The rms transverse speed is $v_t = \sqrt{2kt/m}$. For a stellar envelope temperature of, say, 7000 K we thus find a rms average aberrational angle of $v_t/c = 3.6\times 10^{-5}/\sqrt{A}$ rad, where A is the atomic weight of whatever source–particle species is predominantly responsible for emission. Hence the image-disk angular diameter of any stellar "point source"—if the radiations of different atoms are incoherent—would be enhanced by aberration to at least $2v_t/c = 4$ arc seconds for incandescent gases no heavier than oxygen.

Here the assumption is made that any single photon arriving from an incoherent light source was emitted by a particular source atom in a definite state of motion. (In contrast, coherent radiation may involve cooperation of many source atoms.) The result further assumes an "ideal detector" able to respond to the individual photon. An actual detector might respond to groups of at least n photons. The resulting image dispersion would be divided by $\sqrt{n}$. For a photographic plate $n \approx 4$, so the dispersion disk diameter corresponding to the photographically recorded image would be of the order of 2 arc seconds rather than 4 arc seconds.

Observed stellar image dispersions are orders of magnitude below the above value. On a clear night, any optical astronomer can refute by direct observation the source–sink-relative-velocity interpretation of the v parameter in the aberration formula. If source motion mattered at all, observational astronomy would be a vestige of its present self and there would be no such thing as a *constant of aberration*.[8]

III. DISCUSSION

Relative motion can involve different objects at the same time or the same (or an equivalent) object at different times. This is particularly apparent in the case of inertial motions, which are timeless, so that we do not have to worry about what "same time" means in the preceding sentence. Aberration exists as an observable phenomenon only in the presence of differing or changing states of motion. (For any two states of motion of the Earth that are approximately identical—1 year apart—stellar aberration is undetectable.) The foregoing considerations establish that consequences contrary to fact ensue from the assumption fostered by some textbooks that the relative velocity v appearing in the aberration formula is the relative velocity between the source and observer. Given that the formula itself is correct, what is the proper operational definition of v? Consistently with observation, there remains little alternative to the following definition.

Definition: The parameter v in the standard aberration formula is the relative velocity of the detector in different states of motion, i.e., the velocity of the detector (observer) relative to itself at different times or relative to its equivalent counterpart at the same time.

Note that the (omnidirectional) light source does not enter the discussion as to either its distance from the observer or its state of motion relative to anything whatsoever. In the present interpretation, the *relativity principle is obeyed* in that only the relative velocity of tangible objects enters the law of aberration; however, the relative velocity in question is the relative velocity of the detector in two different states of motion at different times, or the relative velocity of two physically different detectors at the same time or different times. If such detector relative velocities vanish, stellar aberration vanishes as an observable phenomenon.

A luminiferous ether (an intangible unsuited to operational definition) remains entirely outside the discussion. Further, we leave unexplored possible implications for such esoterica of physics as the aberrational "invisibility" of the Lorentz contraction—such a contraction itself a phenomenon never successfully looked for, or always successfully unseen.[9] Such invisibility purports to result from the three-dimensionality of a fast-moving extended source (whereby there arise differing propagation delays from different parts of the source). But if analysts should seek to connect such theorizings with observation, it is apparent that they must proceed with caution in interpreting the v parameter in their equations.

In summary, the stellar aberration measured by astronomers is the direct consequence of changes in the Earth's state of motion *and of nothing else*. This has been recognized by some textbook writers, perhaps most explicitly by Synge, who states[10] that "For observational purposes, the frames of reference S and S' consist of the earth itself at two different positions in its orbit around the sun."

[1]E. F. Taylor and J. A. Wheeler, *Spacetime Physics* (Freeman, San Francisco, 1966), p. 90.

[2]P. G. Bergmann, *Introduction to the Theory of Relativity* (Prentice-Hall, New York, 1946), p. 36.

[3]W. G. V. Rosser, *An Introduction to the Theory of Relativity* (Butter-

worths, London, 1964), p. 151.
[4]H. Arzeliès, *Relativistic Kinematics* (Pergamon, New York, 1966), p. 162.
[5]M.-A. Tonnelat, *The Principles of Electromagnetic Theory and of Relativity* (Reidel, Dordrecht, The Netherlands, 1966; Gordon and Breach, New York, 1966), p. 114.
[6]C. Møller, *The Theory of Relativity* (Oxford, London, 1952), 2nd ed., p. 60.
[7]H. E. Ives, J. Opt. Soc. Am. **40**, 185 (1950), footnote, p. 190.
[8]E. Whittaker, *A History of the Theories of Aether and Electricity* (Harper, New York, 1951), Vol. 1, p. 95. The "constant of aberration" is an oversimplification in that what is actually observed for any star over the course of a year is an image that traces out a projection of the Earth's orbit on the celestial sphere centered on the observer's mean line of sight to the star. The "constant" (20.5 s of arc) is the angular measure of the semimajor axis of such a projected ellipse. Since the star's line of sight changes steadily throughout the night, its figure of aberration is continually changing.
[9]T. E. Phipps, Jr., *Heretical Verities: Mathematical Themes in Physical Description* (Classic Non-fiction Library, Urbana, IL, 1987), p. 23.
[10]J. L. Synge, *Relativity: The Special Theory* (North-Holland, Amsterdam, 1964), 2nd ed., p. 147.