Time measurement using a realizable atomic clock in a moving frame of reference

1970 J. Phys. A: Gen. Phys. 3 368

(http://iopscience.iop.org/0022-3689/3/4/007)

J. Phys. A: Gen. Phys., 1970, Vol. 3. Printed in Great Britain

Time measurement using a realizable atomic clock in a moving frame of reference

L. M. STEPHENSON

Department of Electronic and Electrical Engineering, University College, London WC1, England

MS. received 29*th January* 1970, *in revised form* 17*th March* 1970

Abstract. All realizable clocks are made up from two distinct parts. The first part consists of either an oscillatory mechanism or a device that exhibits a decay process. The second part consists of an integrating unit. The observations that may be made of a realizable atomic clock, situated in an inertial frame of reference, are examined. It is shown that special relativity theory yields two different predicted results depending on the particular experimental arrangement of the atomic clock. It is proved that the usually quoted result does not relate to a true elapse of time in a moving frame of reference; it is obtained by placing the first part of an atomic clock in the moving frame of reference and the second part of the atomic clock in the observer's frame of reference. By analysing in detail the experiment appropriate to having a complete atomic clock in the moving frame, and an identical complete atomic clock in the observer's frame, it is shown that no discrepancy is observable in the readings of the two clocks. Thus special relativity theory requires that the measured time elapse between two given events should be the same for all observers in all inertial frames of reference.

It is shown that two different time scales may be defined for a given moving frame of reference; these two time scales have been indiscriminately used in previous analyses because they have been incorrectly assumed to be the same. The invariant time scale given by the readings of identical, complete clocks assumes only a constant velocity of light; it is, therefore, the primary time scale of special relativity. The more commonly used, velocity dependent, time-scale additionally assumes a time-assigning function; this additional assumption restricts the use of what is now a secondary time scale to the class of problems involving only the steady-state radiation of electromagnetic energy.

1. Introduction

Since Einstein first formulated special relativity theory in 1905 there have been innumerable references to the 'apparent' length of a moving, rigid body. Until 1959 nearly all physicists would have firmly upheld the view that the 'observed' length of a moving, rigid body would be foreshortened in the direction of motion. However, in upholding such a statement, very little attention was paid to defining, with precision, firstly the exact length that was being observed and, secondly, the precise experimental technique which was to be used to record the apparent length.

In 1959 Terrell (Terrell 1959) wrote an excellent paper which analysed in detail the instantaneous photographic image that would be recorded by a camera when a moving, self-illuminated, rigid body was photographed. In particular, Terrell showed that such a body, if spherical in shape in its own frame of reference, would produce a circular photographic image on the camera plate. Terrell's analysis was limited to the case in which the angle subtended by the body at the camera was infinitesimal, but Penrose (1959) has produced an analysis which shows that the same conclusion is reached for a finite angle subtended at the camera.

It should be emphasized that in both of these analyses special relativity theory was used throughout and the results are generally accepted as being the correct deductions from special relativity theory for this particular, well-defined experiment. In simple terms one may state that the *combined* effect of the length foreshortening of the body in the direction of motion, and the finite velocity of light, results in the predicted circular photographic image. The photons which arrive simultaneously at the photographic plate at two diametrically opposite points on the image have not, in general, come from diametrically opposite points on the moving spherical body at the same instant in time.

Hence, looking at special relativity from a slightly different viewpoint, one may say that the real transformation of length that occurs in a moving frame of reference is necessary in order that the *apparent* shape of a moving, self-illuminated, spherical body should appear circular to a stationary observer.

Since 1905 there have also been innumerable references to the 'apparent' time that would be recorded by a moving clock. Many statements continue to be made that a 'clock', situated in a moving, unaccelerated frame of reference, will 'record' less time than an identical stationary clock. Right up to the present day very little attention has been paid to defining, with precision, the exact time unit which is being observed in the moving frame of reference and the precise experimental method which is to be used to record the passing of these time units. In 1960 Terrell (Terrell 1960) produced another paper which dealt with time transformations and time measurements; this paper also contains an extensive bibliography. However, in this later paper, Terrell makes no attempt to define the precise source of 'time-units' in the moving frame, or the precise observational technique which is to be used to count these units. Terrell claims to have shown (as have many other authors) that there is an asymmetry of 'clocks' moving in unaccelerated frames of reference and he finishes the section of his paper by stating: " ... As long as the observers continue their uniform, unaccelerated, velocities there is no basis for saying that anyone's clock is *really* indicating the passage of less time than another's clock; to do so would be to give preference to one of the coordinate frames ...".

The purpose of the present paper is threefold. Firstly, it will be shown that the usual analysis which is made to deduce the time elapse in a moving frame of reference is a legitimate analysis that may be related to a meaningful experiment; but it will be proved that the result of this analysis does not relate to the reading that would be obtained on a physically realizable atomic clock situated in a moving frame. Secondly, it will be shown that a complete, physically realizable, atomic clock *does not* show any discrepancy when subjected to unaccelerated motion, as compared with an identical stationary clock. Finally it will be shown that the μ-meson decay process is such that longer 'apparent lifetimes' will be observed for μ-mesons travelling at high velocities. The author believes that uncritical thinking has occurred in mentally equating the performance of moving clocks, on the one hand, and the observations of μ-mesons on the other hand. It is essential to be precise about the definition of a realizable clock, and the nature of the μ-meson decay process, so that a correct relativistic analysis of an experiment involving either of these phenomena may be made. Whereas it is perfectly true to state that an atomic source of electromagnetic radiation, and the decay process of atomic particles, may equally well be used *to form the basis* of a realizable clock, these phenomena are not clocks as such. It is essential to add an integrating device if one wishes to form a complete, physically realizable, clock.

In the analysis which follows special relativity theory will be used throughout.

2. The usual analysis which is made to deduce the readings of a clock that is situated in a moving frame of reference

The very simplest relativistic problem will be considered, consisting of a stationary frame of reference which contains an observer O. This observer O sees a frame of reference A travel past him at a velocity $+v$ along the x axis. When A has reached a point at a distance equal to L_0 from O the velocity of A is reversed to $-v$, and at some later time the observer O sees A re-pass him. (This problem is an elementary example of the so-called 'clock-paradox' but it will be shown later that the usual analysis does not relate to the reading of a *complete* clock.)

The first point to be cleared up is the acceleration period that occurs when the velocity of A is reversed. Terrell (1960), and many other authors, have shown that the acceleration period is unimportant in any consideration of what occurs *during* the constant-velocity period, although there may well be an *additional* effect. The acceleration period may, in any case, be entirely eliminated by using three inertial frames of reference and arranging a suitable, physically realizable, signalling system (Terrell 1960).

To return to the simple problem outlined above, it is then usually stated that the observer O 'sees' that a 'clock' situated in A is running slow, such that

$$\Delta\tau_a = \Delta\tau_0(1-v^2/c^2)^{-1/2} \tag{1}$$

and that O then 'observes' that the total 'time' elapsed on A's 'clock', between the time when A first passes O and the time when A re-passes O, will be

$$\frac{2L_0}{v}\left\{\Delta\tau_0\left(1-\frac{v^2}{c^2}\right)^{-1/2}\right\}^{-1} \tag{2}$$

and for $\Delta\tau_0 = 1$, as defined in O's frame of reference:

$$T_1 = \frac{2L_0}{v}\left(1-\frac{v^2}{c^2}\right)^{1/2} \quad \text{seconds.} \tag{3}$$

Now if O also has an identical clock situated in his own frame of reference, he will observe that the time elapsed, between the same events, is

$$\frac{2L_0}{v}(\Delta\tau_0)^{-1} \tag{4}$$

and again for $\Delta\tau_0 = 1$:

$$T_2 = \frac{2L_0}{v} \quad \text{seconds.} \tag{5}$$

3. The construction of realizable clocks

There is no single, unambiguous, way in which a 'time-unit', generated in frame A, may be observed by the stationary observer O. Any realizable clock consists of two distinct parts. The first part is usually characterized by some device or mechanism that exhibits oscillatory motion, although this first part of a clock may equally well be formed by some device that exhibits the decay of some quantity. Two very basic examples of the first part of a clock are the balance wheel and hairspring of a watch, or the sand and orifice of an egg timer.

The second part of a realizable clock is completely separate from the first part in principle, although this may not be obvious on a superficial examination. This second part consists of an integrating device that records the total number of periods of oscillation of the first part of the clock, or the fractional decay that has occurred in the first part of the clock.

The clock that is easiest to analyse in detail, and which also forms the basis of all present-day time standards, is the atomic clock. A typical atomic clock consists of an atomic source of electromagnetic radiation that exhibits a characteristic frequency. The total number of periods of this frequency, which have elapsed from some particular starting time, are then usually recorded on a digital read-out indicator.

If any particular clock is situated in a moving frame of reference, then an observer is free to make two possible observations. Firstly, he may examine the oscillatory motion, or the decay process, associated with the first part of the given clock. Secondly, he may examine the actual reading on the face of the indicator that records the output of the second part of the given clock. It is wrong to assume, prior to carrying out a detailed analysis, that the observations will necessarily yield identical results. It is even wrong to assume that the separate parts of different types of clock will all yield the same result; it is clear, for example, that pendulum clocks must provisionally be placed in a separate category.

We will proceed to carry through the analysis for the particular case of an atomic clock.

4. Measurement

4.1. *The first measurement that may be made using an atomic frequency source and a digital read-out indicator*

In the first realizable experiment that is to be analysed the first part of an atomic clock, namely the source of electromagnetic energy, will be placed in frame of reference A and the electromagnetic energy radiated by this source will be examined by O. The relative motion will be exactly as was described in the first paragraph of § 2. The observer O will use a digital read-out indicator to determine the number of periods he observes of the electromagnetic energy coming from A. The observer O is thus using the second part of an atomic clock as his detection apparatus.

Observer O knows that the time unit will be transformed *in A's* frame, and as the source of electromagnetic energy is in A's frame, he *considers* the frequency to be

$$f_{\mathrm{a}} = f_0(1 - v^2/c^2)^{1/2}. \tag{6}$$

Observer O now wishes to *record* the number of periods of the electromagnetic energy coming from A. Using the relativistic Doppler effect formula, the number of periods which O calculates he should receive, between the two coincidences of O's frame and frame A, is

$$N_3 = L_0 f_0 \left\{ \left(\frac{1}{v} + \frac{1}{c}\right) \left(\frac{1 - v/c}{1 + v/c}\right)^{1/2} + \left(\frac{1}{v} - \frac{1}{c}\right) \left(\frac{1 + v/c}{1 - v/c}\right)^{1/2} \right\}.$$

Note here particularly that the observer records red-shifted energy for a time $L_0(1/v + 1/c)$ and he records blue-shifted energy for a time $L_0(1/v - 1/c)$. Hence

$$N_3 = \frac{2L_0 f_0}{v}\left(1 - \frac{v^2}{c^2}\right)^{1/2}. \tag{7}$$

But O's definition of time is given by $T = N/f_0$, and hence

$$T_3 = \frac{2L_0}{v}\left(1-\frac{v^2}{c^2}\right)^{1/2} \quad \text{seconds.} \tag{8}$$

4.2. *The second measurement that may be made using an atomic frequency source and a digital read-out counter*

We will follow the initial set-up of the experiment outlined in § 4.1 and place the atomic frequency source in frame A. Likewise the observer O will use the second part of the atomic clock to record the number of periods he observes of the radiation coming from frame A. However, the position of the observer O in the stationary frame of reference will now be changed so that the motion of A appears to be transverse. This can be achieved if, prior to the experiment, O moves away from the origin of the stationary frame, along a line perpendicular to the x axis. The length L_0 that frame A is seen by O to travel along the x axis is made infinitesimal. Equation (6) is still valid and the number of periods that O *records* of the electromagnetic energy coming from A is

$$N_4 = \frac{2L_0 f_0}{v}\left(1-\frac{v^2}{c^2}\right)^{1/2} \tag{9}$$

and hence

$$T_4 = \frac{2L_0}{v}\left(1-\frac{v^2}{c^2}\right)^{1/2} \quad \text{seconds.} \tag{10}$$

If the observer O is also provided with a *complete* atomic clock, consisting of an identical frequency source coupled directly to a digital read-out indicator, in his own, stationary, frame of reference, then for either experiment 4.1 or experiment 4.2 this indicator will read:

$$N_5 = \frac{2L_0 f_0}{v} \tag{11}$$

and

$$T_5 = \frac{2L_0}{v} \quad \text{seconds.} \tag{12}$$

4.3. *Summary of the results that have been obtained in* §§ 4.1 *and* 4.2

An atomic frequency standard was placed in the moving frame A and the observer made measurements on this using the second half of an atomic clock. It has been shown that the reading obtained is independent of the position of the second half of the atomic clock in the observer's frame of reference, as would be expected (i.e. $T_3 = T_4$). This result is the usual transverse Doppler-shift result. It correctly applies to any measurement of the frequency shift of electromagnetic waves in the absence of longitudinal motion (e.g. the Ives–Stillwell experiment). Note particularly, however, that the usual result quoted for a 'time elapse' in a moving frame of reference is given by T_1 in equation (3). As $T_1 = T_3 = T_4$ the *realizable experiment* appropriate to the usually quoted result corresponds to having the first half of an atomic clock in the moving frame of reference and the second half of an atomic clock in the observer's frame of reference. When the frequency standard is in one frame and the integration is performed in another frame it is clearly incorrect to associate the result T_1 with a 'time elapse' in either of the frames of reference.

It should be noted that Born (1963) has clearly stated, without giving a reason, that the type of analysis given in § 2 is not appropriate to the consideration of time intervals in a moving frame. Unfortunately Born did not suggest an experiment or an analysis which *would* be relevant to a realizable clock travelling in a moving frame.

5. The third measurement that may be made using an atomic frequency source and a digital read-out indicator

In this third experiment to be analysed we shall be guided by the inevitable and logical conclusion that the detailed analysis in §§ 3 and 4 produces. To measure the true time-elapse in the moving frame of reference we must place in that frame of reference an atomic frequency source that is directly coupled to a digital read-out indicator. We now have a complete atomic clock in the moving frame of reference A. To determine what this clock will read we must first make a *complete* transformation into frame A, to understand the working of the clock in that frame; having done this we must consider what the effect will be of observing the reading of an atomic clock in frame A from the stationary frame containing the observer O.

In A's frame of reference:

$$f_{\mathrm{a}} = f_0(1-v^2/c^2)^{1/2} \tag{6}$$

$$\Delta x_{\mathrm{a}} = \Delta x_0(1-v^2/c^2)^{1/2}. \tag{13}$$

Equation (13) requires careful consideration. It indicates that the length unit, for example the Bohr radius, in A's frame of reference is smaller than in the stationary frame, when measured along the x axis. If the length *unit* used by A is smaller than that used by O, then the length *in space* which O saw to be L_0 will appear to an observer in A to be larger than L_0. Thus

$$L_{\mathrm{a}} = L_0(1-v^2/c^2)^{-1/2}. \tag{14}$$

It is not at all unusual for equations (13) and (14) to be confused.

The atomic clock in A's frame of reference, for either the experiment described in § 2 or the experiment described in § 4.2, will now give a reading on its digital read-out indicator that is equal to

$$N_6 = \frac{2L_{\mathrm{a}}f_{\mathrm{a}}}{v} = \frac{2L_0 f_0}{v} = N_5. \tag{15}$$

The digital read-out indicator of an atomic clock in frame A is now reading the same as the digital read-out indicator of an *identical* clock situated in the stationary frame. Although this result is most encouraging, the analysis is still not complete. In the first place, what will be the effect of the relative motion if the observer O now *examines* the digital read-out indicator in A's frame? The digits appearing in the window of A's indicator represent a time elapse in A's frame of reference, but the same numbers appearing in the window of an identical indicator situated in A's frame could be a statement of, say, the fine-structure constant as measured in A's frame. If a number is inscribed on the surface of a material substance, then that number, whatever it may represent physically, cannot be affected by relative motion. All observers, as they pass near to the digital indicator, will record the same number, provided they are each equipped with suitable photographic recording apparatus.

We may thus conclude that the reading of A's clock read-out indicator, as observed by O, will always be the same as the reading of the indicator of an *identical* clock situated in O's frame of reference.

There is one loose end that now needs to be tidied up—it is this one point that has caused all the confusion in the past and, up to this stage, the wording of the present paper has been carefully chosen so that no ambiguities or double-definitions entered the analysis. Following on from equation (15), examine now what happens if an observer on A decides, very reasonably, to define his time scale in the following way:

$$T_a = \frac{N}{f_a} \tag{16}$$

which follows on from O's definition of time that was made following equation (7). Under these circumstances

$$T_6 = T_5(1-v^2/c^2)^{-1/2} \tag{17}$$

and we are apparently back to having time scales transformed in the moving frame of reference. The answer to the problem is simple. Equation (17) relates the readings of two *different* atomic clocks; and equation (16) tells us that the 'gear-train' ratio of the mechanism which connects the oscillatory source and the read-out indicator differs in the two clocks by a factor equal to f_0/f_a. It is for this reason that the analysis for the comparison of *identical* clocks should be made in terms of the number of oscillations of a given type of source which occur between any two given events. As this point is so important let us re-state it in a slightly different way. The reading on the output indicator of an oscillatory clock is basically a count of the number of periods of oscillation. If we wish the clock to read in some arbitrary unit, say seconds, then we introduce an extra gear-train just prior to the indicator which ensures that $T = N/f_0$. We will call this clock Y. If we place clock Y in a moving frame of reference it still contains the same gear-train that we chose to put in it and the transformation between oscillatory periods and the time reading in seconds *is still* $T = N/f_0$. If, in the moving frame of reference, we construct a clock and decide that we wish the clock to record time as given by $T_a = N/f_a$, then a *different* extra gear-train will have to be incorporated. We will call this clock Z. Time scale T_a now goes slower than time scale T, but this is not very surprising because we have deliberately made the clocks run at different rates by using different gear-trains. The frequencies f_0 and f_a are *really* different in the two frames of reference, but the time scales T and T_a have *artificially* been made different. It is essential to appreciate that in a moving frame of reference the time unit and the secondary time scale T_a are affected only by the time transformation of special relativity. The primary time scale, *which directly gives the time elapse between events*, is affected by both the time *and length* transformations of special relativity. The primary time scale is T and is the same in all inertial frames of reference. Hence, any given complete clock will measure the same time elapse between two events in all inertial frames of reference. This statement will apply to pendulum clocks and biological clocks provided the term complete is extended to cover all the local environmental conditions which have a direct bearing on the action of the clock.

6. The apparent 'lifetime' of μ-mesons

The average lifetime of μ-mesons may be stated, in terms of an atomic source, as being equal to n periods of that source, thus:

$$t = nf^{-1} \tag{18}$$

the average lifetime of μ-mesons situated in frame A will be

$$t_a = nf_a^{-1} = n\{f_0(1-v^2/c^2)^{1/2}\}^{-1}. \tag{19}$$

Note particularly here the fundamental difference between equations (16) and (19). Equation (16) is a statement referring to the arbitrary time scale to be associated with a complete clock. Equation (19) is a statement of the frequency transformation appropriate to the first part of a clock which makes use of a decay process—it is equivalent to the statement made in equation (6) for the frequency transformation appropriate to the first part of an oscillatory clock.

The average distance travelled by a μ-meson, situated in frame A, as observed by a stationary observer O will be

$$l_0 = vnf_0^{-1}(1-v^2/c^2)^{-1/2}. \tag{20}$$

The apparent path length travelled by μ-mesons is thus increased by the factor of $(1-v^2/c^2)^{-1/2}$ in accordance with established observations.

The observation of the increased path length travelled by μ-mesons is another example of the first part of the clock (the decay process) occurring in the moving frame of reference, and the second part of the clock (the integration associated with the length measurement) being in the stationary frame of reference. Under these circumstances relative motion has an effect, but one should not talk about the time elapse as occurring in either the moving frame or the stationary frame.

7. Conclusions

It has been shown that complete, identical, clocks will not indicate any discrepancy in their readings when subjected to unaccelerated motion. It has also been shown that this statement is consistent with the established observations of the increased path lengths traversed by μ-mesons which enter the Earth's atmosphere at high velocities.

Attention has also been directed to two sources of error that have arisen in past analyses relating to the reading of a clock situated in a moving frame of reference. The first error has been simply one of not defining precisely the nature of the clock that is to be examined and the nature of the experimental technique which is to be used to carry out the examination. This first error is closely related to the error that was made up until 1959 when analysing the effect which relativistic length contraction would have on actual observations.

The second error that attention has been directed to is that of making double, mutually contradictory, definitions in the same analysis. One is perfectly entitled to relate the number of periods of oscillation of a given clock, the frequency of oscillation, and a time interval, by the formula $T = N/f$. However, one must not then make a second definition of this type, to be applied in a moving frame of reference, and then apply it directly to the reading of an *identical*, complete clock. The act of making this second definition is to mathematically create two different time scales in the moving frame of reference. The secondary time scale compensates for the *real* change in the *time unit* in the moving frame of reference and may be applied directly only to a radiated, single frequency, electromagnetic wave (or to a clock whose gear-train has been modified in the ratio f_0/f_a). If one wishes to discuss the readings of identical clocks it is preferable to analyse the number of periods of oscillation, of a given type of clock, which occur between two stipulated events.

All that has been said in this paper is directly related to an assumption that has always been made in the past concerning the time-assigning function in special relativity theory. It has mistakenly been presumed that a time-assigning function must be assumed *in addition* to the assumption of a constant velocity of light. The assumed time-assigning function is generally expressed in the form $t_2 = \frac{1}{2}(t_1 + t_3)$ for an out-and-return signal. This assumption is not essential; its inclusion simply limits the analysis to a consideration of *direct* measurements on radiated, unmodulated, electromagnetic waves in the steady state, and it is therefore a chosen restriction on special relativity theory.† The secondary time scale T_a will correctly apply to this restricted class of problem.

The primary time scale T does *not* depend on any assumed time-assigning function; it is the basic time scale of special relativity theory and assumes only a constant velocity of light. The primary time scale T measures the true time elapse as determined by any complete clock no matter whether its action is based on atomic, mechanical, chemical or biological processes.

One may confirm the essence of all that has been stated in this paper by a very simple alternative approach. If one defines the time elapse between two given events in terms of the number of periods of oscillation of a given type of atomic source, or in terms of the fractional decay of a given type of decay process, then the time elapse is being stated as a ratio or, in other words, as a *dimensionless number*. For example:

$$\text{Number of ticks of a clock} = \frac{\text{time units between the two events in frame A}}{\text{time units between consecutive ticks in frame A}}.$$

Although the transformations of special relativity theory will affect the time unit, and hence the oscillation frequency of the atomic source in a moving frame of reference, the transformations cannot affect the time elapse between two events when this time elapse is expressible as a dimensionless ratio. This simple alternative approach cannot bring out the details of the analysis covered in §§ 3–6 but, like many simple approaches, it is extremely powerful in its application. One may immediately extend the result to cover non-inertial frames of reference. General relativity theory also requires that all dimensionless ratios, derived from two or more quantities measured in a given frame of reference, shall be unaffected by transformation. One may therefore conclude that the primary time scale, which directly gives the time elapse between events, must remain the same in all frames of reference. A predicted universal primary time scale of this nature in no way affects the observed gravitational frequency shift of electromagnetic energy which is radiated from a region remote from the observer.

The existence of a universal primary time scale, together with the existence of a relativistic change in the time unit, is important in view of the doubts that have been cast on whether the special and general theories of relativity are truly relativistic theories in the Machian sense. It is still possible that an evaluation of the, at present, indeterminate boundary conditions of general relativity theory may lead eventually to a unified field theory. The initial approach to a novel type of boundary condition evaluation has recently been proposed by the author (Stephenson 1969).

† This assumption is equivalent to including equation (16) in the analysis. Equation (16) gives the correct result for measurements on radiated electromagnetic waves but it must not be applied directly when considering the reading of a complete clock.

Acknowledgments

I am very grateful to Professor C. W. Kilmister for the helpful comments he made after reading the first draft of this paper.

References

BORN, M., 1963, *Nature, Lond.*, **197,** 1287.
PENROSE, R., 1959, *Proc. Camb. Phil. Soc.*, **55,** 137–9.
STEPHENSON, L. M., 1969, *J. Phys.* A: *Gen. Phys.*, **2,** 475–6.
TERRELL, J., 1959, *Phys. Rev.*, **116,** 1041–5.
—— 1960, *Nuovo. Cim.*, **16,** 457–68.