Special theory of relativity and the Sagnac effect



1988 Sov. Phys. Usp. 31 861

(http://iopscience.iop.org/0038-5670/31/9/A05)

# Special theory of relativity and the Sagnac effect

A. A. Logunov and Yu. V. Chugreev

*M. V. Lomonosov Moscow State University and*

*Institute of High Energy Physics, Protvino (Moscow District)*

Usp. Fiz. Nauk **156**, 137–143 (September 1988)

The Sagnac experiment can be explained systematically without going beyond the special theory of relativity.

The Sagnac effect[1] ranks along with Michelson's experiment, measurements of the velocity of light, etc., as one of the fundamental experiments in the theory of relativity. Nevertheless, one still encounters in the literature both incorrect explanations of this effect, based on signals which are moving at velocities higher than the velocity of light[2,3] or with references to the general theory of relativity,[3] and the declaration that the Sagnac effect is puzzling and cannot be explained in a noncontradictory way.[4] We will be discussing this matter in more detail below. We therefore think it is pertinent at this point, *on the basis of methodological considerations and also to avoid any possible misconception, to emphasize once more that the Sagnac effect is of a purely special-relativistic nature.* We will of course not need any velocities higher than the velocity of light or, especially, the general theory of relativity.

We begin with a description of the Sagnac experiment. Mirrors are placed at the corners of a quadrangle on a disk. The corners are positioned with respect to each other in such a way that a light ray coming from a monochromatic source is reflected from the mirrors around a closed loop and then returns to the source. A beam splitter can be used to split the ray from the source into two rays, which would move in opposite directions around this closed loop.

Sagnac observed that if the disk was put in rotation a ray traversing the loop in the direction of the rotation would arrive at the source after a ray traversing the loop in the direction opposite the rotation, with the result that there would be a shift of an interference pattern on a photographic plate. When the rotation was reversed, the interference fringes shifted in the opposite direction.

What explanation has been given to this effect? Sagnac himself derived a theoretical value for the magnitude of the effect through a purely classical summation of the velocity of light with the linear rotation velocity for a ray moving in the direction opposite the rotation. He did the same for the ray propagating in the direction of rotation, using a corresponding subtraction. The discrepancy between this result and the experiment result is of the order of 1%.

In one form or another, frequently obscured, this explanation persisted. For example, we might cite a typical assertion in this regard by Sommerfeld in his *Optics*[3]: "Michelson's negative result of course tells us nothing about the propagation of light in *rotating* media. In this case it would be necessary to appeal, not to the special theory, but to the general theory of relativity, with its additional terms which correspond to centrifugal mechanical forces. If, however, one notes that in later experiments (of Sagnac *et al.*—note by the present authors) the only velocities which were involved satisfied $v \ll c$, and the only effects which were involved were of first order in $v/c$, then one can make do without any theory of relativity and carry out the calculations at a purely classical level."

That explanation is actually in the spirit of old ether concepts and, as Yilmaz[4] has correctly pointed out, is wrong, since it allows velocities higher than the velocity of light. Furthermore, it contradicts the relativistic law for adding velocities.

For clarity, let us follow Yilmaz and examine the circular path which would be traced out by the rays in a Sagnac experiment carried out with an infinite number of mirrors. According to the classical law for adding velocities in a rotating coordinate system, the velocities of the light would be $c \pm \omega r_0$, where $\omega$ is the rotation frequency, and $r_0$ is the radius of the path.

It is then obvious that the magnitude of the effect is given by

$$\Delta_S = \frac{2\pi r_0}{c-\omega r_0} - \frac{2\pi r_0}{c+\omega r_0} \approx \frac{4\pi r_0^2 \omega}{c^2} = \frac{4\omega S}{c^2} ,$$

which agrees well with experiment. Here $S$ is the area of the closed loop around which the rays propagate.

We see that this result has been achieved at the cost of introducing an anisotropy in the velocity of light and of actually allowing velocities higher than the velocity of light. This anisotropy contradicts the relativistic law for adding velocities (even in first order), and it contradicts the constancy of the velocity of light. In other words, although these ideas do generate a prediction which agrees (in a first approximation) with the correct prediction, it is internally untenable. With this fact in mind, Yilmaz labeled the Sagnac effect "puzzling."[4]

A team of experimentalists at the University of Maryland, headed by C. O. Alley, is planning measurements of the velocity of light with the goal of seeking a possible anisotropy.[1)]

In the present paper we will show that an explanation of the Sagnac effect is completely within the capabilities of the special theory of relativity and that none of the following need be invoked: the general theory of relativity, velocities higher than the velocity of light, or any other postulates. We will see in detail how to calculate the time between the arrival of the rays at the source, working in a fixed inertial frame of reference. We will also do this using a noninertial frame of

 0038-5670/88/090861-04$01.80  

reference which is rotating along with the apparatus. As should be expected, the results of the calculations agree.

We begin with the case of an inertial frame of reference. We write the interval in cylindrical coordinates:

$$ds^2 = c^2\,dt^2 - dr^2 - r^2\,d\varphi^2 - dz^2. \qquad (1)$$

Let us assume, as we mentioned earlier, that the light rays are moving in the $z = 0$ plane along a circle of radius $r = r_0 = \text{const}$. For the light, the interval is zero, so we find

$$\frac{d\varphi_\pm(t)}{dt} = \pm\frac{c}{r_0}. \qquad (2)$$

The plus sign indicates the ray which is moving in the direction of the rotation, and the minus sign indicates that which is moving in the opposite direction.

Using the initial conditions $\varphi^\pm(0) = 0$, $\varphi_-(0) = 2\pi$, we find the behavior of the angles $\varphi_\pm$ of the two rays as functions of the time $t$:

$$\varphi_+(t) = \frac{c}{r_0}t,$$

$$\varphi_-(t) = 2\pi - \frac{c}{r_0}t. \qquad (3)$$

The rays meet at the time $t_1$, at which we have $\varphi_+(t_1) = \varphi_-(t_1)$. Substituting (3) into this relation, we find

$$\varphi_+(t_1) = \varphi_-(t_1) = \pi.$$

Now choosing $t_1$ as the initial time and repeating the arguments, we find that the next meeting of the rays occurs at specifically that point (in a three-dimensional space) from which the rays were emitted, i.e., at the point with the coordinates $\varphi = 0$, $r = r_0$, $z = 0$.

We wish to emphasize that this result obviously does not depend on the angular rotation velocity of the frame of reference of the source and the mirror.

By definition, the changes in the angular coordinate of the source are described by [the initial condition is $\varphi_s(0) = 0$]

$$\varphi_s(t) = \omega t. \qquad (4)$$

Consequently, the source meets the + ray at the coordinate time $t_+$, determined by the condition $\varphi_s(t_+) = \varphi_+(t_+) - 2\pi l$; i.e.,

$$t_+ = \frac{2\pi}{(c/r_0) - \omega}. \qquad (5)$$

It meets the − ray at the coordinate time $t_-$, which is determined by the condition $\varphi_s(t_-) = \varphi_-(t_-)$:

$$t_- = \frac{2\pi}{c/r_0 + \omega}. \qquad (6)$$

The form of expressions (5) and (6) might give the impression that the velocity of light is anisotropic and different from $c$ in this case. However, this is not true. The velocity of light is the same for the two rays and equal to $c$; the explanation for the difference in the times at which the rays return to the source is that over the ray propagation time the source has moved a certain distance, so the rays traverse different distances before they meet the source (the + ray travels the greater distance).

Let us find the proper-time difference between the arrival of the two rays for an observer at the source. By definition, this difference is

$$\Delta = \frac{1}{c}\int_{s(t_-)}^{s(t_+)} ds = \frac{1}{c}\int_{t_-}^{t_+}\frac{ds}{dt}\,dt,$$

where $s$ is the interval. Substituting the values of the interval at the point of the source into (7) and using (4), we find

$$ds^2 = c^2\,dt^2 - r_0^2\,d\varphi^2 = c^2\,dt^2\left(1 - \frac{r_0^2\omega^2}{c^2}\right),$$

where $\omega^2 r_0^2/c^2 < 1$.

We find the exact value of the Sagnac effect[2]:

$$\Delta = \left(1 - \frac{r_0^2\omega^2}{c^2}\right)^{1/2}(t_+ - t_-) = \frac{4\pi\omega r_0^2}{c^2\,[1 - (r_0^2\omega^2/c^2)]^{1/2}}. \qquad (8)$$

In deriving (8) we used only the absolute concepts of events in which rays meet each other or meet the source—we did not use the concept of the velocity of light with respect to the rotating frame of reference.

A point which needs special emphasis is that the essence of the special theory of relativity is the postulate of a pseudo-Euclidean geometry of space-time; the principle of relativity and the "postulate of the constancy of the velocity of light" are secondary, particular consequences of this fundamental position. It is the postulate of a pseudo-Euclidean geometry which allows us to examine the effects in noninertial frames of reference, while remaining exactly in the special theory of relativity. We will demonstrate this point below, and we will show that calculations of the Sagnac effect in a rotating (noninertial) frame of reference differ in no fundamental way from corresponding calculations in an inertial system. Consequently, the general theory of relativity simply is not involved here. The reader interested in more details about this point might look in Ref. 5.

We will now show that if an experimentalist wished to measure the velocity of light in this experiment with respect to a fixed frame of reference or with respect to a rotating frame he would always find that the result was exactly $c$. We first recall that what can be checked by direct experiment is a so-called physical velocity, which should be distinguished from a coordinate velocity, which we might say has a mathematical rather than physical meaning.

We thus consider the interval of the pseudo-Euclidean Minkowski space:

$$ds^2 = \gamma_{ik}\,dx^i dx^k = \gamma_{00}c^2\,dt^2 + 2\gamma_{0\alpha}c\,dt\,dx^\alpha + \gamma_{\alpha\beta}\,dx^\alpha dx^\beta, \qquad (9)$$

where $\gamma_{ik}$ is the metric tensor, for which the Riemann curvature tensor is zero. We can identically transform interval (9) to the form

$$ds^2 = \left[(\gamma_{00})^{1/2}c\,dt + \frac{\gamma_{0\alpha}}{(\gamma_{00})^{1/2}}\,dx^\alpha\right]^2 - \left(\frac{\gamma_{0\alpha}\gamma_{0\beta}}{\gamma_{00}} - \gamma_{\alpha\beta}\right)dx^\alpha\,dx^\beta,$$

which is of the same form as the interval in an ordinary inertial frame of reference,

$$ds^2 = c^2\,d\tau^2 - dl^2. \qquad (10)$$

The role of physical time is thus being played by the quantity $d\tau = (\gamma_{00})^{1/2}dt + \frac{\gamma_{0\alpha}}{c}\,dx^\alpha(\gamma_{00})^{1/2}$, which is equal to $ds/c$ in the case $dl = 0$; the role of the square of the physical distance is played by the quantity

$$dl^2 = \left(\frac{\gamma_{0\alpha}\gamma_{0\beta}}{\gamma_{00}} - \gamma_{\alpha\beta}\right)dx^\alpha\,dx^\beta,$$

which is equal to $-ds^2$ in the case $d\tau = 0$. It is clear from these definitions that both $d\tau$ and dl can be measured, since they can be expressed in terms of an absolute value: an interval. It also follows that the velocity which is measurable experimentally is the quantity $dl/d\tau$. Since an invariant definition of light signals in the special theory of relativity is of the form $ds^2 = 0$, we find on the basis of (10)

$$\left|\frac{dl}{d\tau}\right| = c. \tag{11}$$

This result means that in whatever type of frame of reference—inertial or nonertial—that the experimentalist measures the velocity of light, the local value of this velocity will be constant, equal to $c$, in absolute value everywhere. In the case of an inertial frame, the quantity $d\tau$ is a total differential, and we can say that the physical velocity of light is globally constant. The coordinate velocity of light, $dx^\alpha/dt$, on the other hand, can have any value except zero and infinity.

To calculate experimentally measurable times, distances, etc., we need to know the components of the metric tensor $\gamma_{ik}$, along with the coordinates. We thus conclude from this discussion that the physical velocity of light with respect to a fixed frame of reference is

$$\mathbf{v} = \frac{d}{d\tau} = \pm c\mathbf{n}_\varphi,$$

where $\mathbf{n}_\varphi$ is a unit vector in the azimuthal direction. In this frame of reference, the physical and coordinate velocities of light are the same.

We now consider the same physical process—the propagation of rays in opposite directions around a circle—in a noninertial frame of reference which is rotating at an angular velocity $\omega$. To find the form of the interval in this system, we transform coordinates:

$$\begin{aligned}
\varphi_n &= \varphi_0 - \omega t_0\\
t_n &= t_0\\
r_n &= r_0\\
z_n &= z_0
\end{aligned} \tag{12}$$

In terms of the new coordinates $t_n$, $r_n$, $\varphi_n$, $z_n$ we find the interval to be (for simplicity, we are omitting the subscript n)

$$ds^2 = \left(1 - \frac{\omega^2 r^2}{c^2}\right) c^2\, dt^2 - \frac{2\omega r^2}{c}\, d\varphi c\, dt - dr^2 - r^2 d\varphi^2 - dz^2. \tag{13}$$

It is not possible to realize a rotating frame of reference with $r \geqslant c/\omega$ physically since in the limit $r \to c/\omega$ the inertial mass and moment of inertia of such a frame would become infinite. Although the four-dimensional geometry remains pseudo-Euclidean, the geometry of three-dimensional space, which is determined (as we have already noted) by the metric tensor $\varkappa_{\alpha\beta} = \gamma_{0\alpha}\gamma_{0\beta}/\gamma_{00} - \gamma_{\alpha\beta}$, is not Euclidean in this case. Let us calculate the curvature tensor of such a space:

$$R^{(3)}_{\lambda\mu\nu\sigma} = \frac{1}{2}\left(\partial^2_{\sigma\mu}\varkappa_{\lambda\nu} - \partial^2_{\sigma\lambda}\varkappa_{\mu\nu} - \partial^2_{\mu\nu}\varkappa_{\lambda\sigma} + \partial^2_{\nu\lambda}\varkappa_{\mu\sigma}\right) + \varkappa_{\eta\rho}\left(\Gamma^\eta_{\nu\lambda}\Gamma^\rho_{\mu\sigma} - \Gamma^\eta_{\sigma\lambda}\Gamma^\rho_{\mu\nu}\right);$$

here $\Gamma^\eta_{\nu\lambda}$ denotes the connectedness of the three-dimensional space and is given by

$$\Gamma^\eta_{\nu\lambda} = \frac{1}{2}\varkappa^{\eta\sigma}\left(\partial_\nu\varkappa_{\sigma\lambda} + \partial_\lambda\varkappa_{\sigma\nu} - \partial_\sigma\varkappa_{\lambda\nu}\right),$$

where the matrix $\varkappa^{\eta\sigma}$ is the inverse of $\varkappa_{\alpha\beta}$. Substituting the effective metric $\varkappa_{\alpha\beta} = \mathrm{diag}\,\{1, r^2/[1 - (\omega^2 r^2/c^2)], 1\}$ into the equation for the curvature tensor, and calculating the connectedness $\Gamma^\eta_{\nu\lambda}$, we find that it has only one independent nonvanishing component,

$$R^{(3)}_{r\varphi r\varphi} = \frac{3\omega^2 r^2}{c^2\,[1 - (\omega^2 r^2/c^2)]}.$$

We see that the curvature of the three-dimensional space is nonzero and that the coordinate $r$ cannot take on values greater than $c/\omega$, since the coefficient $\gamma_{00}$ would change sign, and that change would be physically inadmissible.

A characteristic manifestation of this deviation from a Euclidean nature also in the three-dimensional geometry of the space is the known fact that the ratio of the circumference of a circle to its radius is not equal to $2\pi$:

$$\frac{1}{r}\int_{l(0)}^{l(2\pi)} dl = \frac{2\pi}{[1 - (\omega^2 r^2/c^2)]^{1/2}} > 2\pi.$$

Let us examine the Sagnac effect in this noninertial frame of reference. We will carry out the calculations by the old method. As before, the paths traced out by the light rays are circles of radius $r_0 = \text{const}$ which lie in the $z = 0$ plane. From condition (13) we find, using the relation $ds_2 = 0$, the behavior of the angle $\varphi$ as a function of the coordinate time $t$:

$$\frac{d\varphi\pm}{dt} = -\omega \pm \frac{c}{r_0}. \tag{14}$$

Using the initial conditions $\varphi_+(0) = 0$, $\varphi_-(0) = 2\pi$, we find

$$\begin{aligned}
\varphi_+(t) &= \frac{ct}{r_0}\left(1 - \frac{\omega r_0}{c}\right),\\
\varphi_-(t) &= 2\pi - \frac{ct}{r_0}\left(1 + \frac{\omega r_0}{c}\right).
\end{aligned} \tag{15}$$

The first meeting of the rays occurs at the time $t_1$, which corresponds to the value $\varphi_+(t_1) = \varphi_-(t_1)$, i.e., to the value $\varphi_1 = \pi[1 - (\omega r_0/c)]$ of the angular variable. Similar arguments lead us to the conclusion that the second meeting of the rays occurs "at the angle"

$$\varphi_2 = 2\pi\left(1 - \frac{\omega r_0}{c}\right), \tag{16}$$

i.e., at an angular distance of $-2\pi r_0\omega/c$ from the source.

The behavior of the angular coordinate of the source in this case is described by the trivial equation $\varphi_s = \text{const} = 0$.

The coordinate time $t_+$, at which the + ray meets the source, is found, as before, from the condition $\varphi_s(t_+) = 0 = \varphi_+(t_+) - 2\pi$:

$$t_+ = \frac{2\pi r_0}{c - \omega r_0}, \tag{17}$$

Correspondingly, the time $t_-$ is found to be

$$t_- = \frac{2\pi r_0}{c + \omega r_0}.$$

The interval of proper time between the two events in which the rays arrive at the point of the source is found with the help of definition (7):

$$\Delta = \frac{1}{c}\int_{t_-}^{t_+}\frac{ds}{dt}\,dt = \left(1 - \frac{\omega^2 r_0^2}{c^2}\right)^{1/2}(t_+ - t_-) = \frac{4\pi\omega r_0^2}{c^2\,[1 - (\omega^2 r_0^2/c^2)]^{1/2}}.$$

This is the same as the result calculated in the fixed frame of reference. Interestingly, in the nonrelativistic approximation the expression for the shift of the interference fringes which follows from the expression for $\Delta$ is also valid in a medium; i.e., it does not depend on the refractive index of the optical fiber, the group-velocity dispersion, and so forth.[6]

Let us calculate the physical velocity of light for this process. From (10) and (14) we find

$$\frac{dl}{d\tau} = \pm c n_{\varphi}.$$

We also note that in a rotating frame of reference the coordinate velocity of light is anisotropic: $d\varphi/dt = -\omega \pm c/r_0$.

We wish to stress that the detector does not detect changes in the frequency of the rays, since the metric coefficient $\gamma_{00}$, which is responsible for the redshift, is constant over the entire propagation path of the rays.

Interesting in this connection is the shift of the frequency of light in this noninertial frame of reference. If a light ray of frequency $\nu_0$ is directed away from the rotation axis along the radius, its frequency for an observer on the disk will increase in accordance with

$$\frac{\nu(r)-\nu_0}{\nu_0} = \frac{1}{[1-(\omega^2 r^2/c^2)]^{1/2}} - 1 \underset{r \ll c/\omega}{\approx} \frac{1}{2}\frac{\omega^2 r^2}{c^2} \qquad (18)$$

in the blue direction. This point has been verified experimentally within[7] $10^{-2}\%$.

We have thus shown that the Sagnac effect can be explained without the need to modify the special theory of relativity, without the need to use velocities greater than the velocity of light, and without the need to resort to the general theory of relativity. All that is needed is strict adherence to the special theory of relativity.

We wish to thank H. Yilmaz and C. O. Alley for useful discussions.

[1)] Private communication from C. O. Alley.

[2)] In a calculation of the realistic Sagnac effect, in which case the path traced out by the light ray is a broken line, one would have to take into account the deformation of the centrifuge caused by centrifugal forces.


[1] M. G. Sagnac, J. Phys. (Paris) **4**, 177 (1914).

[2] S. I. Vavilov, *Collected Works* (In Russian), Vol. 4, Izd. Akad. Nauk SSSR, M., 1956.

[3] A. Sommerfeld, *Optics*, Academic Press, N.Y., 1954 [Russ. transl., IL, M., 1953].

[4] H. Yilmaz, in: *Proceedings of the Fourth Marcel Grossman Meeting on General Relativity* (ed. R. Ruffini), Elsevier Science Publ., Rome, 1986, p. 1753.

[5] A. A. Logunov, *Lectures on the Theory of Relativity and Gravitation* (in Russian), Nauka, M., 1987.

[6] S. Ezekiel and H. J. Arditty Eds., *Fiber-Optic Sensors and Related Technologies*, Springer-Verlag, Berlin, Heidelberg, New York (Springer Series in Optical Science, Vol. 32).

[7] H. J. Hay, J. P. Schiffer, T. E. Cranshaw, and P. A. Egelstaff, Phys. Rev. Lett. **4**, 165 (1960).


Translated by Dave Parsons