Sagnac effect and pure geometry

Angelo Tartaglia* and Matteo Luca Ruggiero

DISAT, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy
INFN, Sezione di Torino, Via Pietro Giuria 1, Torino, Italy

(Dated: November 8, 2014)

Abstract

We show that the Sagnac effect is not necessarily due to the presence of a rotating observer, but rather to the closed path of light in space and an even inertial relative motion between the observer and the physical device forcing light to move along a closed path.

*Corresponding Author: angelo.tartaglia@polito.it
I. INTRODUCTION

The Sagnac effect takes its name from Georges Marc Marie Sagnac, a French physicist who, in 1913, partially interpreting previous experimental evidence, showed that the time of flight of light emitted by a source on a rotating platform and sent back by means of mirrors along a path closed in space was asymmetric. The asymmetry between the forward (in the sense of rotation) and backward beams, in terms of times of flight difference, is proportional to the angular velocity of the platform and appears as a phase difference measured by an interferometer. Sagnac interpreted this result as an evidence against Einstein’s relativity and in favour of a static luminiferous ether.[1][2] It was however immediately and easily shown that the Sagnac effect is indeed a fully relativistic effect.

Today the Sagnac effect is relevant for various applications, all related to the measurement of rotation rates. On the commercial side we have gyrolasers, developed since the 70’s of the last century, used on planes, ships, submarines, missile guidance systems... The name 'gyrolaser' is motivated by the fact that the devices replace old mechanical gyroscopes and use a laser to produce the counter-rotating light beams. Instead of exploiting the interference between the two opposed light rays, they measure the beat frequency of the standing wave resulting from the superposition of two waves propagating in opposite directions; that frequency is again proportional to the absolute angular velocity of the apparatus (a thorough description of gyrolaser devices can be found in the review paper [3]).

The most sensitive ring lasers (a.k.a. gyrolasers) sense the diurnal rotation of the earth, the wobbles of the terrestrial axis, the tiny rotations of the laboratory hosting the instrument, due to elastic deformations of the ground caused by liquid and solid tides and also by various surface phenomena. These rings are used for high accuracy geophysical measurements.[4]

The extremely high sensitivity obtained in most recent ring lasers is such that now the Sagnac effect is becoming important also for fundamental physics: the possibility has been suggested of using large ring lasers to detect the Lense-Thirring effect on Earth [5] and an experiment, under the acronym GINGER, is now in the phase of development and preliminary testing at the National Laboratories of the Italian INFN at the Gran Sasso site.[6][7]

The Sagnac effect plays a role also in the elaboration of the position of a receiver obtained from the GPS: a Sagnac correction is indeed required because a rotating clock (onboard a satellite or on the surface of the earth) suffers a loss of synchrony with itself at each turn.
The relevance of the relativistic effects on clocks carried around, either in the same or in the opposite sense with respect to the rotation of the earth, was experimentally verified by Hafele and Keating.[8] [9] [10]

The Sagnac effect has been deeply investigated, soon after the realization of Sagnac’s experiments, as a disproval of the theory of relativity, and also more recently, in connection with the issue of synchronization in rotating reference frames (see for instance [11] and the monograph [12]). All this is well known, but it is also accompanied by a common feeling that the Sagnac effect is essentially due to the presence of a rotating observer or rotating device: in short, to non-inertial motion. Examples of papers elaborating on various aspects of the Sagnac effect, without discussing the issue of the necessity of rotation are [13] [14] [15].

In reality, however, non-inertial motion is not necessary. Indeed, a few years ago some experimental papers have been published which claim that even if no rotation is present, a “generalized Sagnac effect” arises in a uniformly moving fiber.[16, 17] In the present paper we aim at addressing the issue, that is we want to show that the real ingredients of the Sagnac effect are two: a) a closed circuit followed by light in opposite directions; b) a relative (even inertial) motion of the emitter/receiver with respect to the physical apparatus supporting the closed path of light. We believe that this approach could be useful in teaching the foundations of relativity, since the calculations involved are simple, and they allow a deep insight into the physics of the problem, that ultimately is connected with the relativity of simultaneity.

II. THE SAGNAC EFFECT

The physical principles of the Sagnac effect are explained in full details in the well known paper by Post,[18] and they can be summarized as follows. The Sagnac effect may easily be described in classical terms if one assumes that the speed of light is \( c \) with respect to a static ether. Considering the rotating platform mentioned in the Introduction you see that it will take longer for light to reach again the emission point on the rim of the platform just because, meanwhile, the receiver will have moved forward by a distance \( \Delta l_+ = vt_+ \) where \( t_+ \) is the total time of flight and \( v \) is the velocity of the emitter with respect to the ether; the geometric length of the path is \( l \). Hence the time of flight is read from the following relation
\[
\frac{l + \Delta l_+}{c} = \frac{\Delta l_+}{v}. \quad \text{In the opposite sense the receiver will move towards the returning light so that the path will be shorter by } \Delta l_\llap{=}_- = vt_-, \text{ whence } \frac{l - \Delta l_-}{c} = \frac{\Delta l_-}{v}. \quad \text{Solving for } t_+ \text{ and } t_- \text{ gives the time of flight difference:}
\]

\[
\Delta t = t_+ - t_- = \frac{2lv}{c^2 - v^2} \quad (1)
\]

An equivalent deduction can be made adopting the viewpoint of the rotating observer. In his/her reference frame light will be expected to have speed \(c - v\) in one sense (forward) and \(c + v\) in the other; now the path is the same \(l\) for both. The times of flight are again different and a trivial calculation brings back again to formula \((1)\). Add that in the example we have \(v = \omega R\) with \(\omega\) representing the angular velocity of the platform and \(R\) the distance of the observer from the rotation axis; simplify the geometry assuming that the path of light is a circle\(^1\) at radius \(R\) so that \(l = 2\pi R\); consider that in all practical cases we have \(v \ll c\), and Eq. \((1)\) will be transformed into:

\[
\Delta t \simeq 4 \frac{A \cdot \omega}{c^2} \quad (2)
\]

Here \(A\) is the area enclosed in the path of the light beams, represented as a vector, whose magnitude is \(\pi R^2\), perpendicular to the plane containing the trajectory \(^2\); \(\omega\) is also a vector and the dot product projects the area in a plane orthogonal to the rotation axis. Equation \((2)\) is the famous Sagnac formula, where the \(\simeq\) symbol is currently replaced by a plain \(=\).

As said, however, the Sagnac effect does not need any ether at all. Special (and general) relativity is essentially geometry so we may use a geometric approach in order to work out what happens. In the simplified configuration of the classical example the situation may be represented in a spacetime graph like the one in Fig. \((1)\).

Everything is drawn on the surface of a cylinder and the viewpoint is that of an inertial observer at rest with respect to the axis of the rotating platform.\(^{[19]}\) The vertical axis of the figure is \textit{not} the rotation axis of the turntable; it is the coordinate time axis. The worldline of the rotating observer is a helix and the two light beams are represented by two opposite helices as well, which, after one turn, intersect the worldline of the observer at two different

\(^{[1]}\) This would have been practically impossible in Sagnac’s time, but is feasible today with modern optical fibers, as it is done in commercial gyrolasers.

\(^{[2]}\) Actually the planarity is not needed.
FIG. 1: Spacetime diagram of the Sagnac effect. $OAB$ is the world line of the rotating observer. The line is a helix drawn on a three-dimensional cylinder. Along the vertical axis inertial time is measured. The other two lines (helices) represent two light rays. $A$ and $B$ are the events where the observer is reached by the counter-, respectively, co-rotating beams.

events ($A$ and $B$, in the figure). The interval between $A$ and $B$ is the proper time difference measured by the observer between the arrival times of the two light rays.

A cylinder is indeed a flat surface, so we may have a simpler view cutting the cylinder along one generatrix, then opening it in a flat spacetime strip like in Fig. 2. Now helices become straight lines and the geometry is evident and simple; the only trick to remember is that the right border of the strip is identified with the left border, so that lines reaching one border reappear at the corresponding point on the other.

The properties of straight lines and triangles are the usual ones, recalling that this is not Euclidean space, but Minkowski spacetime so that the squared length of the hypotenuse is now a squared proper time multiplied by $c^2$ and is given by the difference between the squared coordinate time span times $c^2$ and the square of the travelled distance. Pure geometry [19] leads again, at the lowest approximation order, to Eq. (2).

If one prefers to proceed analytically the starting point can be the line element of Minkowski spacetime written using cylindrical coordinates in space: $ds^2 = c^2 dt^2 - dr^2 - r^2 d\phi^2 - dz^2$. Introducing the constraints $r = R = constant$ and $z = constant$, corresponding to all physics taking place on the rim of a disk perpendicular to the $z$ axis, the problem is reduced to two effective dimensions and the line element becomes:
FIG. 2: Same as Fig. 1. The cylinder has been cut along a generatrix passing through the rotating observer at time 0, and opened. For convenience, in order to make the picture more compact, two replicas of the opened cylinder are shown side by side: one enrolled to the right, the other to the left. Points $O'$ and $O''$ coincide with $O$. Four windings are shown. The vertical straight line is the world line of an inertial observer at rest with the axis of the disk. $AB$ is the Sagnac effect expressed in terms of proper time of the rotating observer.

$$ds^2 = c^2dt_0^2 - R^2d\phi_0^2$$

(3)

The label $0$ means that the coordinates are those of the inertial observer.

To account for the rotation of the platform it is convenient to introduce axes rotating together with the platform so that the angular coordinate becomes $\phi = \phi_0 - \omega t_0$. Then we need a Lorentz transformation between the inertial observer at rest with the axis of the platform and the inertial frame instantaneously comoving with the observer on the rim.
of the turntable. In Special Relativity the Lorentz transformations leave the line element unchanged, so that in the rotating reference frame it becomes:

\[ ds^2 = (c^2 - \omega^2 R^2)dt^2 - R^2d\phi^2 - 2R^2\omega d\phi dt \]  

(4)

On writing the previous expression in the form

\[ ds^2 = g_{00}dt^2 + g_{\phi\phi}d\phi^2 + 2g_{0\phi}dt d\phi \]  

(5)

and considering that for light we have \( ds = 0 \), it is possible to work out \( dt \):

\[ dt = \frac{-g_{0\phi}d\phi \pm \sqrt{g_{\phi\phi}^2d\phi^2 - g_{\phi\phi}g_{00}d\phi^2}}{g_{00}} \]  

(6)

We are interested in solutions located in the future, so we choose \( dt > 0 \), i.e. the + sign in the formula. Under the simplified assumption that light moves along a circumference, if we integrate in the two opposite directions (counterclockwise, \( d\phi > 0 \); clockwise, \( d\phi < 0 \)) from the emission to the absorption events, we get the expression for the co-rotating \( (t_+) \) and counter-rotating \( (t_-) \) times of flight, and the difference between them turns out to be

\[ \Delta t = t_+ - t_- = -2 \int \frac{g_{0\phi}}{g_{00}} R d\phi \]  

(7)

where \( \ell \) is the circumference of radius \( R \).\(^3\) Moreover, the observer measures the proper-time difference:

\[ \Delta \tau = -2\sqrt{g_{00}} \int \frac{g_{0\phi}}{g_{00}} R d\phi \]  

(8)

On substituting the expressions (4) of the metric coefficients, we get the following expressions for the coordinate and proper times of flight difference

\[ \Delta t = 4\frac{\pi R^2 \omega}{c^2(1 - \omega^2 R^2/c^2)} \]

\[ \Delta \tau = 4\frac{\pi R^2 \omega}{c^2\sqrt{1 - \omega^2 R^2/c^2}} \]

At the lowest approximation order in \( \omega R/c \) the two coincide, and they are in agreement with Eq. (2), with \( A = \pi R^2 \hat{u}_z \). The unit vector \( \hat{u}_z \) points along the \( z \) direction.

\(^3\)In the most general case \( \ell \) is the integration path and, using arbitrary coordinates, the formula would be \( \Delta t = -2 \int_\ell \frac{g_{0i}}{g_{00}} dx^i \) (see e.g. [20]).
III. SAGNAC WITHOUT ROTATION

Despite the widespread tendency to ascribe the Sagnac effect to rotating systems it is easy to show that rotation is not essential. Consider a source/receiver of light in inertial motion at the speed $v$ with respect to a set of mirrors rigidly fastened to one another or to an optical fiber so that they guide the light beams emitted by the moving source along a closed path in space. The situation is schematically shown in Fig. 3 where an optical fiber is envisaged.

![Figure 3](image_url)

FIG. 3: $O$ represents a source/receiver of light. Light is sent both forward and backward (dashed lines). The irregular line joining $F$ with $R$ represents an optical fiber of length $l_1$. $O$ is moving with the velocity $v$ with respect to the fiber. The total length of the path is of course $L = l_1 + l_2$.

Repeating the simple reasoning made at the beginning of sec. II we easily arrive to the same result written in formula (1). The only additional clarification is that Eq. (1) is expressed from the view point of an observer at rest with the fiber (or array of mirrors); in the proper time of the moving emitter/receiver it would rather be (we set $\beta = \frac{v}{c}$):

$$\Delta \tau = 2 \frac{Lv}{c^2 \sqrt{1 - \beta^2}} \sim 2 \frac{Lv}{c^2}$$  \hspace{1cm} (9)

No rotation, no acceleration, no enclosed area appear. What matters is just the existence of relative (even inertial) motion and of a closed path in space. Of course, with reference to Fig. 3, $v$ is parallel to the light rays in $RF$.

The effect we have just described and the related formula (9) have been verified by experiment by Wang, Zheng, Yao, and Langley (WZYL).[16, 17]
Using the geometrical approach the result is rather obvious. In spacetime the world lines are drawn again on a cylinder. The only difference is that on Fig. 1 the cylinder had a circular cross section, now the cross section is rather irregular. In any case the cylinder is a flat bidimensional Minkowski surface and, when opening it in a plane, the graph is exactly like in Fig. 2.

Formula (1) appears also in Ref. [21] where it is related to the relativity of simultaneity typical of special relativity.

Another aspect to clarify is related to the form (2) of the effect. The accidental presence of the area $A$ suggested to some people [22] [23] an interpretation of the Sagnac effect (at least for matter waves) as a sort of Aharonov-Bohm effect: however, this is just an analogy, which holds at lowest relativistic order only (see e.g. [20]).

IV. A SIMPLE EXAMPLE

In order to get a deeper insight, let us consider a simple example where the calculation can be carried out in full details. Our interferometer is at rest in the laboratory and is made of four mirrors $M_1$, $M_2$, $M_3$, $M_4$, at the corners of a rectangle: if we use Cartesian coordinates, their positions are given in Fig. 4, and the interferometer path length is $L = 4a + 4b$. An observer, moving at speed $v$ to the right in the laboratory frame, sends two counter propagating light signals when he is at the origin. The paths of the counterclockwise (CCW) and clockwise rays (CW) are depicted in Fig.s 5 and 6, respectively, as seen in the laboratory frame. Note that in the laboratory frame, the CW path is not closed, and the CCW path overlaps with itself. What we want to show, by the following calculation, is that both paths are closed in the frame moving with the observer.

In terms of Cartesian spacetime coordinates $(ct, x, y)$ in the laboratory frame, the events sequence along the CCW path is

- $e_0$: $(0, 0, 0)$ (emission)
- $e_1$: $(a, a, 0)$ (reflection by $M_1$)
- $e_2$: $(a + 2b, a, 2b)$ (reflection by $M_2$)
- $e_3$: $(3a + 2b, -a, 2b)$ (reflection by $M_3$)
FIG. 4: The four mirrors $M_1, M_2, M_3, M_4$ that form the interferometer are at rest in the laboratory frame.

FIG. 5: CCW rays path, as seen in the laboratory; the symbol ♢ denotes the position of the observer when he sends the signal and when he is reached by the signal. To draw the figure, we chose $\beta = 20/101$, so that $\gamma = 101/99$, and $b = a/4$.

- $e_4$: $(3a + 4b, -a, 0)$ (reflection by $M_4$)
- $e_5$: $(ct_1, vt_1, 0) = \left(\frac{4(a+b)}{1-\beta}, \frac{4(a+b)}{1-\beta} \beta, 0\right)$ (reception)

In particular, the total propagation time is obtained from $ct_1 = vt_1 + 4a + 4b$.

As for the CW path, we have the following events:

- $e_6$: $(0, 0, 0)$ (emission)
- $E_1$: $(a, -a, 0)$ (reflection by $M_4$)
FIG. 6: CW rays path, as seen in the laboratory; the symbol $\blacklozenge$ denotes the position of the observer when he sends the signal and when he is reached by the signal. To draw the figure, we chose $\beta = 20/101$, so that $\gamma = 101/99$, and $b = a/4$.

- $E_2$: $(a + 2b, -a, 2b)$ (reflection by $M_3$)
- $E_3$: $(3a + 2b, a, 2b)$ (reflection by $M_2$)
- $E_4$: $(3a + 4b, a, 0)$ (reflection by $M_1$)
- $E_5$: $(ct_2, vt_2, 0) = \left(\frac{4(a+b)}{1+\beta}, \frac{4(a+b)}{1+\beta} \beta, 0\right)$ (reception)

Again, the total propagation time is obtained from $ct_2 = 4a + 4b - vt_2$. As expected, the CCW is longer than the CW one, since $t_1 > t_2$, so the two signals reach the observer at different times.

To obtain the coordinates of these events in the frame co-moving with the observer, we use the Lorentz transformations:

$$ct' = \gamma (ct - \beta x), \quad x' = \gamma (x - \beta ct), \quad y' = y$$

(10)

where $\gamma = (1 - \beta^2)^{-1/2}$. As a consequence, the Cartesian coordinates $(ct', x', y')$ of the events for the CCW path are

- $e_0$: $(0, 0, 0)$ (emission)
- $e_1$: $\left(\gamma (1 - \beta) a, \gamma (1 - \beta) a, 0\right)$ (reflection by $M_1$)
- $e_2$: $\left(\gamma (1 - \beta) a + 2\gamma b, \gamma (1 - \beta) a - 2\gamma \beta b, 2b\right)$ (reflection by $M_2$)
- $e_3$: $\left(\gamma (3 + \beta) a + 2\gamma b, -\gamma (1 + 3\beta) a - 2\gamma \beta b, 2b\right)$ (reflection by $M_3$)
- $e_4$: $\left(\gamma (3 + \beta) a + 4\gamma b, -\gamma (1 + 3\beta) a - 4\gamma \beta b, 0\right)$ (reflection by $M_4$)

11
For the CW path we have the following coordinates:

- \( e_0: \ (0,0,0) \) (emission)
- \( E_1: \ (\gamma (1 + \beta) a, -\gamma (1 + \beta) a, 0) \) (reflection by \( M_4 \))
- \( E_2: \ (\gamma (1 + \beta) a + 2\gamma b, -\gamma (1 + \beta) a - 2\gamma \beta b, 2b) \) (reflection by \( M_3 \))
- \( E_3: \ (\gamma (3 - \beta) a + 2\gamma b, \gamma (1 - 3\beta) a - 2\gamma \beta b, 2b) \) (reflection by \( M_2 \))
- \( E_4: \ (\gamma (3 - \beta) a + 4\gamma b, \gamma (1 - 3\beta) a - 4\gamma \beta b, 0) \) (reflection by \( M_1 \))
- \( E_5: \ (4\gamma (1 - \beta) (a + b), 0, 0) \) (reception)

These collections of events show that in the moving frame the two paths are closed.

In particular, CCW path is indeed a polygon whose vertices have the following \((x', y')\) coordinates: \((0,0), (\gamma (1 - \beta) a, 0), (\gamma (1 - \beta) a - 2\gamma \beta b, 2b), (-\gamma (1 + 3\beta) a - 2\gamma \beta b, 2b), (-\gamma (1 + 3\beta) a - 4\gamma \beta b, 0)\). The shape of this polygon is depicted in Fig. 7 and its length is \( \ell_1 = 4\gamma (1 + \beta) (a + b) \); since the traversal time is read from event \( e_5, t'_1 = 4\gamma (1 + \beta) \frac{(a+b)}{c} \), we see that the average traversal speed is \( c \).

On the other hand, the CW path is a polygon whose vertices have the following coordinates: \((0,0), (-\gamma (1 + \beta) a, 0), (-\gamma (1 + \beta) a - 2\gamma \beta b, 2b), (\gamma (1 - 3\beta) a - 2\gamma \beta b, 2b), (\gamma (1 - \beta) a - 4\gamma \beta b, 0)\). The shape of this polygon is depicted in Fig. 8, and its length is \( \ell_2 = 4\gamma (1 - \beta) (a + b) \); again, by reading the traversal time \( t'_2 = 4\gamma (1 - \beta) \frac{(a+b)}{c} \) from \( E_5 \), we see that the average traversal speed is \( c \).

We remark that we can easily explain the tilt of the lateral edges in Fig.s 7 and 8 in terms of the motion of the observer: indeed, when the light ray moves upward at right angles to the x axis in the laboratory frame, it has to move upward and left in the observer’s frame; similarly, when it moves downward, it has to move downward and left in the observer’s frame.

The moving observer measures the proper time difference between the propagation of the CCW and CW rays

\[
\Delta \tau = t'_1 - t'_2 = 8\gamma \frac{(a + b)}{c} \beta \simeq 2\frac{Lv}{c^2} \quad (11)
\]
This result is in agreement with the general formula (9). Moreover, this calculation allows to understand the origin of the time difference arising in the moving frame: the two light paths, which are congruent except for their end pieces in the laboratory frame, are no longer congruent in the observer’s moving frame. Consequently, since they have different lengths, light must take different time to propagate through them.

FIG. 7: CCW rays path, as seen in the observer’s frame: both emission and reception take place at the origin (0, 0). To draw the figure, we chose $\beta = 20/101$, so that $\gamma = 101/99$, and $b = a/4$.

FIG. 8: CW rays path, as seen in the observer’s frame: both emission and reception take place at the origin (0, 0). To draw the figure, we chose $\beta = 20/101$, so that $\gamma = 101/99$, and $b = a/4$.

This example shows how to explain the Sagnac-effect in the case of an inertial observer: indeed, a similar approach can be used to do the same thing, for an accelerated one. However, while in our case the observer’s space is Euclidean and everything is unambiguous, care is needed in defining the spatial geometry for an arbitrary accelerated observer: spatial geometry can be defined in terms of congruence of world lines (see e.g. [24]), so it is not simply related to one observer, rather it depends on a collection of observers.

Things are simple in the case of the uniformly rotating disk, where the ordinary Sagnac effect usually is explained: a natural choice of observers that enables to define the spatial geometry is given by the observers sitting at fixed points on the disk (see again [24] and
In this case, both the co-rotating and the counter-rotating rays propagate along the same curve. Then, the Sagnac effect can be explained in terms of different speeds of light in the co-rotating and counter-rotating directions, if clocks along the rim of the disk are Einstein-synchronized, or in terms of a time gap, if a different synchronization procedure is used (see e.g. [26] and [15]).

V. POSSIBLE APPLICATIONS

Not considering the conceptual aspects, the phenomenon we have interpreted and WZYL have measured prompts possible interesting applications, as suggested in [17].

As we see from Eq. (9) the time of flight difference for light is proportional to the relative velocity of the source with respect to the fiber or the mirrors. Now consider a ring laser where the active segment is not rigidly connected to the rest of the annular cavity; imagine, for instance, that the active cavity is attached to the rest via an elastic support allowing for relative vibrations. What is usually measured in a ring laser is a beat frequency of two counter-propagating beams. The beat frequency is easily obtained from Eq. (9) and is:

$$\nu_b = \frac{v}{\lambda}$$

(12)

where \( \lambda \) is the wavelength of the laser. We have a simple speedometer for instantaneous velocities where 1 m/s corresponds roughly to beat frequencies of the order of MHz; but in the obvious case that the active element is subject to an acceleration, a simple differentiation of the output converts the device into an accelerometer.

Of course practical analyses on the stability of the laser modes should be made, but the idea is appealing.

VI. CONCLUSION

We have shown that the Sagnac effect is due to the closure of the path followed by light and to the relative motion of the observer with respect to the physical system obliging the beam to bend and come back to the observer. After having introduced the foundations of the effect, we have studied in full details a simple example to focus on the origin of the time delay in the inertial frame of the moving observer, and we have made a comparison with
the case of the rotating observer. We have proved that the Sagnac effect is not peculiar to rotations and accelerated motion, rather it originates from the closure of the two space paths as seen in the frame co-moving with the emitter/receiver, and from relative motion between the emitter/receiver and the mirrors (or physical device); indeed, its foundations are related to the relativity of simultaneity. On the practical side one may think of interesting applications based on the use of ring lasers as linear accelerometers.

Acknowledgments

The authors are indebted to one of the referees for suggesting the example in section IV.


Phys. 80, 772-781 (2012)