Chapter 1
A Potpourri of Comments about the Fiber Optic Gyro for Its Fortieth Anniversary: How Fascinating It Was and Still Is!

Hervé C. Lefèvre
iXblue

1.1 Introduction

Forty years after its first experimental demonstration by Vali and Shorthill in 1976,[1] the fiber optic gyro (FOG) remains a fascinating device. Reciprocity and serendipity have allowed it to reach a unique performance of long-term bias stability: pure, unaided strapdown inertial navigation with a drift of less than one nautical mile in a month, which corresponds to a rate bias stability of 10 μdeg/h, i.e., 0.3 nanoradian in terms of interferometer phase stability.[2]

This chapter begins with the historical context of the Sagnac effect and proposes that it be renamed, as we shall see, as the Sagnac–*Laue* effect. It then recalls the various points that make the FOG potentially perfect and finishes with a potpourri of comments that should be outlined.

1.2 Historical Context of the Sagnac–Laue Effect

If Huygens proposed in the 17th century a wave theory of light, Newton imposed his views of a corpuscular theory in the early 18th century. It is only in the early 19th century that Young's double-slit experiment reopened the wave theory, knowing that it was not easily admitted: you did not contradict Newton! It required the exceptional quality of the theoretical and experimental work of Fresnel to convince the physics community. However, for the thinking of the time, a wave needed some kind of propagation medium, as for acoustic waves. It was called "luminiferous Aether," and light was seen as propagating at a constant velocity *c* with respect to this fixed aether.

Even when Maxwell showed in 1864 the electromagnetic nature of light waves, aether was not questioned. It required the famous Michelson and Morley experiment in 1887 to have a clear demonstration that the concept of aether should be revised, which yielded the special theory of relativity in 1905: based on earlier theoretical works of Lorentz, Poincaré, Planck, and Minkowski, Einstein abandoned the concept of aether and stated that light propagates at the same velocity *c* in any inertial frame of reference in linear translation, despite its own velocity. This revolutionary conceptual leap was very difficult to accept by many, and Sagnac's experiment (Fig. 1.1), the basis of present optical gyroscopes, was actually performed to demonstrate that aether did exist, as clearly stated in the title of the publication:[3] "The luminiferous Aether demonstrated by the effect of the relative wind of Aether in an interferometer in uniform rotation."

It is now understood that Sagnac's experiment, which takes place in vacuum (actually in air, but it can be considered as in vacuum), can be either explained by relativity or aether theory, and it does not allow one to demonstrate which theory is right or wrong since the explanation is evaluated from a "rest" frame.

It has not been well known in our community, but it was explained very clearly by von Laue (Fig. 1.2) as early as 1911,[4] two years before Sagnac's publication.[3] The Sagnac effect should be renamed the Sagnac–Laue effect!

Another important point of aether theory was the hypothesis made by Fresnel in 1818 regarding the drag of aether by matter, which was demonstrated experimentally by Fizeau in 1851.[5] In the "rest" frame, the

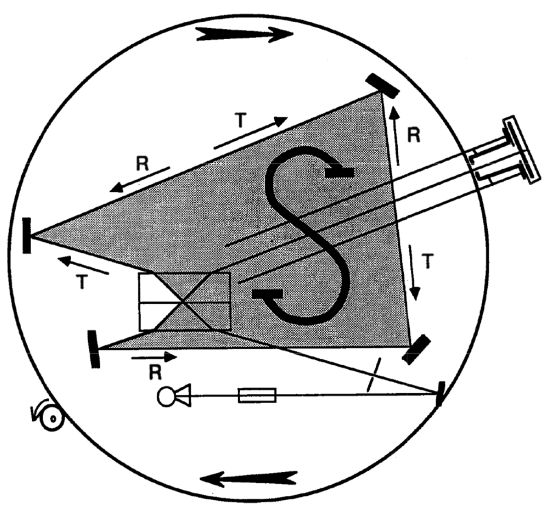

Figure 1.1 Original figure of Sagnac's experiment (S stands for *surface*, i.e., "area" in French).

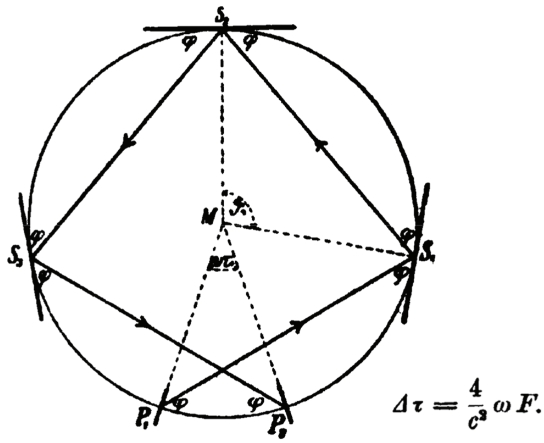

Figure 1.2 Original figure of von Laue's publication of 1911,[4] explaining the Sagnac–Laue effect and giving the formula of the effect (F stands for *Fläche*, i.e., "area" in German).

velocity v of a wave that propagates in a medium with an index of refraction n and that moves at a speed v_m is no longer c/n but

$$v = \frac{c}{n} + \left(1 - \frac{1}{n^2}\right) v_m. \tag{1.1}$$

This Fresnel–Fizeau drag effect was explained by von Laue in 1907[6] as related to the law of the addition of speeds of special relativity:

$$v = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}, \tag{1.2}$$

where v_1 is the speed of an object in a frame moving at v_2 with respect to the "rest" frame, and v is the speed of this object in this "rest" frame. One sees that Eq. (1.2) yields Eq. (1.1) to first order, considering $v_1 = c/n$, the speed of light in the moving medium, $v_2 = v_m$, the speed of the medium with respect to the rest frame, and $v_m << c$. The Fresnel–Fizeau drag effect is actually a *relativistic* effect!

The Sagnac–Laue effect does not depend on the index of refraction n of the propagation medium because of this Fresnel-Fizeau drag effect,[7] as it is clearly experienced in the fiber optic gyro. The path difference looks longer in a medium (Fig. 3), but the velocities are different because of the Fresnel–Fizeau drag of the medium moving with respect to the rest laboratory frame. The speed of the co-rotating wave is increased, and the one of the counter-rotating wave is decreased. If the Sagnac–Laue effect in vacuum can be explained by aether or relativity theory, in a medium, it is related to the Fresnel–Fizeau drag effect and is, then, also a relativistic effect.

Note that Fizeau's experiment that demonstrated the drag effect is usually described in textbooks with a Mach–Zehnder interferometer, whereas it was actually performed with a *reciprocal* ring interferometer (Fig. 1.4). The

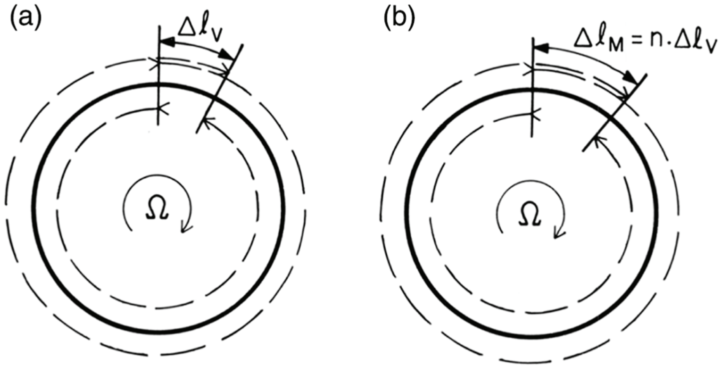

Figure 1.3 (a) Sagnac–Laue effect in vacuum, and (b) Sagnac–Laue effect in a medium: $\Delta\ell_M$ is longer than $\Delta\ell_V$, but this difference is compensated for by the difference of velocities induced by the Fresnel–Fizeau drag.

Figure 1.4 Reciprocal ring interferometer used by Fizeau to demonstrate the matter drag effect in 1851: a source S is placed at the focal point of a lens, and the collimated beam goes through two holes (O_1 and O_2) to yield two parallel beams that propagate in the moving water and that are swapped on a cat's eye (on the left) to return "reciprocally."

publication of Fizeau[5] clearly stated (page 352), more than 160 years ago, the importance of what we call today reciprocity, which is a key point for optical gyroscopes: "It is easy to see that, with this set-up, all the points that are on the path of one beam, are also on the one of the other beam, so that a change of density at any point of the path has the same effect on both beams, and then has no influence on the fringe shift." Beautiful!

1.3 Fascinating Serendipity of the Fiber Optic Gyro

1.3.1 The proper frequency

The first experimental demonstration of a fiber gyro[1] showed that a simple fiber ring interferometer is not perfectly reciprocal. Complementary interference fringes were observed at both ports of the interferometer, depending on the alignments of the fiber ends. All of these problems could have been a severe limitation to high performance, but they can be solved very simply with a so-called reciprocal configuration (Fig. 1.5), proposed independently by Ulrich[8] and by Arditty et al.[9] It is sufficient to feed light into the interferometer through a truly single-mode waveguide (single spatial mode and single polarization) and look at the returning interference wave, which is filtered through this same waveguide in the opposite direction. In this case, the alignments are needed solely to optimize the throughput power (and its related

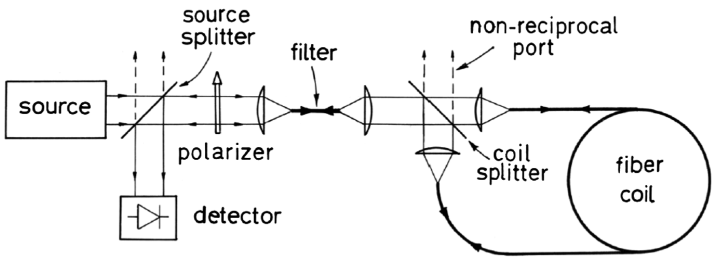

Figure 1.5 Reciprocal configuration with a single-mode filter on the common input–output port.[8,9]

signal-to-noise ratio), which requires difficult but nevertheless reasonable mechanical tolerances. It is now ensured that both returning waves have propagated along exactly the same path in the opposite direction and that they interfere perfectly in phase when the system is at rest. This simple modification has made both opposite paths perfectly identical, zero rotation giving zero phase difference.

The reciprocal configuration provides an interference signal of the Sagnac–Laue effect with perfect contrast since the phases as well as the amplitudes of both counter-propagating waves are perfectly equal at rest. The optical power response is then a raised cosine function of the rotation-induced phase difference $\Delta\phi_R$, and it is maximum at zero. To get high sensitivity, this signal must be biased about an operating point with a nonzero response slope. It is obtained with the use of a *reciprocal* phase modulator placed at one end of the coil. This coil then acts as a delay line (Fig. 1.6).[10] Because of reciprocity, both interfering waves carry exactly the same phase modulation ϕ_m but shifted in time. The delay equals the difference $\Delta\tau_g$ of the group transit time between the long and short paths that connect the modulator and the splitter. This yields a biasing modulation $\Delta\phi_m(t)$ of the phase difference, despite the reciprocity of the phase modulator:

$$\Delta\phi_m(t) = \phi_m(t) - \phi_m(t - \Delta\tau_g). \quad (1.3)$$

The ring interferometer behaves like a perfect delay line filter with a sinusoidal transfer function $2 \sin(\pi \cdot f_m \cdot \Delta\tau_g)$, that is maximum (in absolute

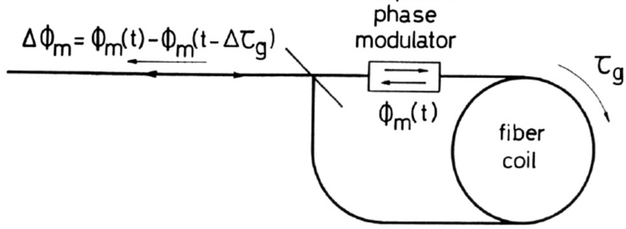

Figure 1.6 Generation of the biasing phase modulation $\Delta\Phi_m$ with a reciprocal modulator, using the delay difference $\Delta\tau_g$ through the fiber coil.

value) at the so-called proper (or *eigen*) frequency f_p, defined as $f_p = 1/(2\Delta\tau_g)$ and its odd harmonics, and that is null at DC and all of the even harmonics. However, a phase modulator is not perfect. It has some spurious harmonic content and parasitic intensity modulation, and the first example of serendipity in the fiber optic gyro is the fact that at the proper frequency f_p, all of these phase modulator defects are completely washed out.[11] The product proper frequency times the coil fiber length is about 100 kHz·km.

1.3.2 Perfection of the digital phase ramp

This biasing modulation–demodulation detection scheme provides very good bias stability since it preserves the reciprocity of the ring interferometer. However, if a high-performance gyroscope must have a stable and low-noise bias, it also requires good accuracy over the whole dynamic range, not only about zero. The measurement of interest is the integrated angle of rotation and not simply the rate. Any past error will affect the future information. This constraint implies the need for an accurate measurement at any rate (i.e., an accurate scale factor). Furthermore, the intrinsic response of an interferometer is sinusoidal, whereas the desired rate signal of a gyroscope should be linear.

This problem is solved with a closed-loop (or phase-nulling) signal-processing approach, as proposed independently by Davis and Ezekiel[12] and Cahill and Udd.[13] The demodulated biased signal (or open-loop signal) is used as an error signal that is fed back into the system to generate an additional feedback phase difference $\Delta\phi_{FB}$ that is opposite to the rotation-induced phase difference $\Delta\phi_R$. The total phase difference is servo-controlled on zero, which provides good sensitivity, since the system is always operated about a high-slope point. With such a closed-loop scheme, the new measurement signal is the feedback value $\Delta\phi_{FB}$. This yields a linear response with good stability since this feedback value $\Delta\phi_{FB}$ is independent of the returning optical power and of the gain of the detection chain. The current consensus holds that the best closed-loop method is the digital phase ramp proposed by Arditty et al. (Fig. 1.7),[14,15] on which serendipity also applies!

In addition, there is the possibility of a second servo loop to control the modulation efficiency (also called V_π, the voltage that induced a π radian phase shift) by controlling the resets on the 2π radian.[14,15] A better way is the so-called *four-state* modulation that allows one to get a permanent V_π control without waiting for a reset (Fig. 1.8).[16]

It is possible to directly view the reason why the digital ramp technique tolerates many defects (such as an imperfect 2π reset, quantization, nonlinearity of the electronic drive, nonlinearity of the phase modulator response, low frequency modulator drift, etc.) without degrading the scale factor performance. The actual phase ramp shift $\phi_{PR\ actual}(t)$ created by the modulator is the sum of an ideal phase ramp shift $\phi_{PR\ ideal}(t)$ and a defect $\phi_{PR\ defect}(t)$. The induced phase difference in the interferometer is

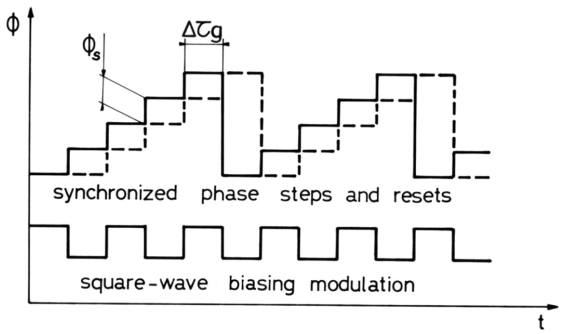

Figure 1.7 Digital phase ramp and resets synchronized with the square-wave biasing phase modulation.

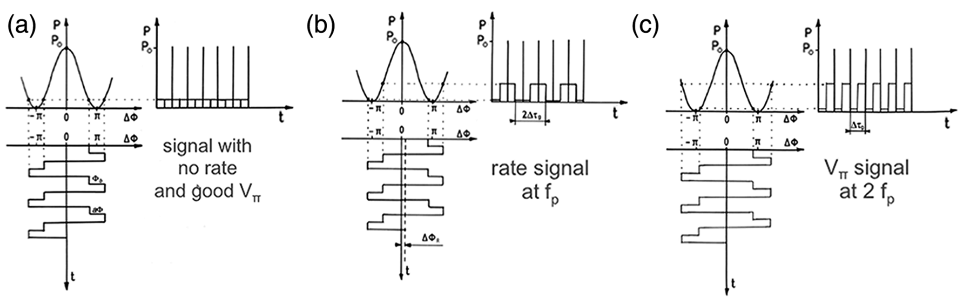

Figure 1.8 Principle of four-state modulation:[16] (a) Signal with no rate and good modulation efficiency (good V_π), (b) rate signal at the proper frequency (f_p), and (c) V_π error signal at twice the proper frequency ($2f_p$).

$$\Delta\phi_{PR\ actual}(t) = \Delta\phi_{PR\ ideal}(t) + \Delta\phi_{PR\ defect}(t). \tag{1.4}$$

Following the general equation (1.3) of any phase modulation, the defect $\Delta\phi_{PR\ defect}(t)$ of the phase difference is given by the delay difference $\Delta\tau_g$ through the coil with

$$\Delta\phi_{PR\ defect}(t) = \phi_{PR\ defect}(t) - \phi_{PR\ defect}(t - \Delta\tau_g). \tag{1.5}$$

Since the mean value of a difference is the difference of the mean values, the mean defect $<\Delta\phi_{PR\ defect}(t)>$ is

$$\begin{aligned} < \Delta\phi_{PR\ defect}(t) > &= < \phi_{PR\ defect}(t) - \phi_{PR\ defect}(t - \Delta\tau_g) > \\ &= < \phi_{PR\ defect}(t) > - < \phi_{PR\ defect}(t - \Delta\tau_g) > , \end{aligned} \tag{1.6}$$

and both mean values $< \phi_{PR\ defect}(t)>$ and $<\phi_{PR\ defect}(t - \Delta\tau_g)>$ are perfectly equal because of reciprocity, over a phase ramp period. Therefore, on average,

$$< \Delta\phi_{PR\ defect}(t) > = 0. \tag{1.7}$$

1.3.3 The optical Kerr effect

As analyzed by Ezekiel et al.,[17] an important case of truly nonreciprocal effect may arise due to the optical Kerr effect, which is nonlinear. Reciprocity is indeed based on the linearity of the propagation equation, but an imbalance in the power levels of the counter-propagating waves can produce a small nonreciprocal phase difference because of propagation nonlinearity induced by the high optical power density in the very small silica fiber core. Slow variations in the splitting ratio of the power divider feeding the sensing coil may therefore translate directly into bias drift. Experimentally, a power difference of 1 μW (e.g., arising from a 10^{-3} splitting imbalance of a 1-mW source) gives a nonreciprocal index difference as small as 10^{-15}, but when integrated along a kilometer of fiber, this produces a phase difference of a few 10^{-5} rad, at least three orders of magnitude above the theoretical sensitivity limit.

The Kerr-induced rotation-rate error results in fact from a complex four-wave mixing process[17] and not simply from an intensity self-dependence of the propagation constant of each counter-propagating wave. It has been shown[18,19] that if the mean value $<I>$ of the modulated intensity I equals its standard deviation, the Kerr-induced phase shift is balanced for both counter-propagating waves, which nulls out the effect on the phase difference—and this is actually the statistic of intensity fluctuation of a broadband spontaneous-emission source, clearly the ultimate serendipity!

1.3.4 Technological serendipity: erbium ASE fiber source and proton-exchanged $LiNbO_3$ integrated-optic circuit

As we just saw, serendipity is found in the mathematics of the signal-processing method and in the physics of the optical Kerr effect, but also in the technology:

- An amplified spontaneous emission (ASE) source,[20,21] derived from telecom erbium-doped fiber amplifier (EDFA) (Fig. 1.9), is a

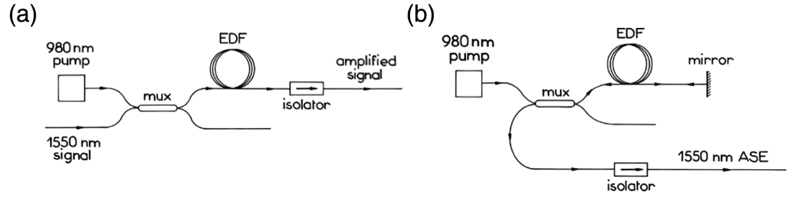

Figure 1.9 (a) Schematics of an erbium-doped fiber amplifier (EDFA) with a pump laser diode, wavelength multiplexer (mux), erbium-doped fiber (EDF), and isolator; and (b) an ASE source with the same components.

quasi-ideal source for the fiber gyro because it features a broadband and stable spectrum with unpolarized emission, which is very beneficial to reduce birefringence-induced non-reciprocity, as analyzed by Pavlath et al.[22]

- A proton-exchanged lithium niobate ($LiNbO_3$) integrated-optic circuit[23] provides excellent polarization rejection (Fig. 1.10), which is crucial since birefringence-induced non-reciprocity is related to the amplitude rejection ratio of the polarizer and not to the intensity, as analyzed by Kintner.[24]

1.4 Potpourri of Comments

1.4.1 OCDP using an OSA

The technique of optical coherence domain polarimetry (OCDP), also called distributed polarization extinction ratio (PER), has been a powerful tool to measure accurately the polarization and birefringence properties of the components of a fiber gyro since it measures the amplitude ratio.[25,26] It was first based on path-matched interferometry (Fig. 1.11) and has yielded several commercially available, dedicated instruments.[27]

However, a path-matched readout interferometer is actually a scanning interferometer, which is the base of Fourier-transform spectroscopy. One measures the full coherence function of the tested light, and through the Fourier transform the spectrum is retrieved. Conversely, a "classical" optical spectrum analyzer (OSA), grating based in particular, measures the spectrum, and through the (inverse) Fourier transform the coherence function can be retrieved. Since an OSA is now a very common multi-purpose test instrument following the development of wavelength domain multiplexing (WDM) telecoms, it is tempting to use an OSA for OCDP instead of a dedicated "single-purpose" test instrument.

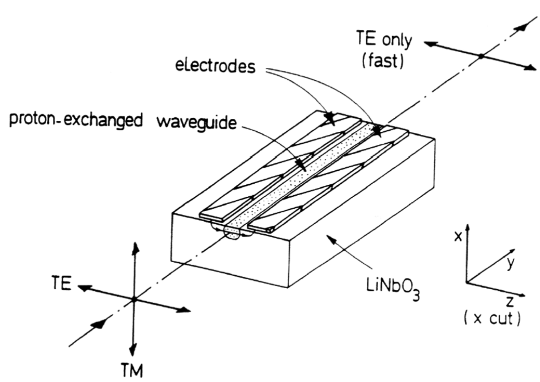

Figure 1.10 Proton-exchanged $LiNbO_3$ waveguide that guides a single-TE polarization.