RTO-AG-339
AC/323(SCI)TP/9

RTO-AG-339

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION

BP 25, 7 RUE ANCELLE, F-92201 NEUILLY-SUR-SEINE CEDEX, FRANCE

© RTO/NATO 1999

Single copies of this publication or of a part of it may be made for individual use only. The approval of the RTA Information Policy Executive is required for more than one copy to be made or an extract included in another publication. Requests to do so should be sent to the address above.

RTO AGARDograph 339

Optical Gyros and their Application

(Gyroscopes optiques et leurs applications)

This AGARDograph has been sponsored by the Systems Concepts and Integration Panel (SCI) of RTO.

Published May 1999

Distribution and Availability on Back Cover

RTO-AG-339
AC/323(SCI)TP/9

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION

BP 25, 7 RUE ANCELLE, F-92201 NEUILLY-SUR-SEINE CEDEX, FRANCE

RTO AGARDograph 339

Optical Gyros and their Application

(Gyroscopes optiques et leurs applications)

edited by

D. Loukianov
R. Rodloff
H. Sorg
B. Stieler

This AGARDograph has been sponsored by the Systems Concepts and Integration Panel (SCI) of RTO.

The Research and Technology Organization (RTO) of NATO

RTO is the single focus in NATO for Defence Research and Technology activities. Its mission is to conduct and promote cooperative research and information exchange. The objective is to support the development and effective use of national defence research and technology and to meet the military needs of the Alliance, to maintain a technological lead, and to provide advice to NATO and national decision makers. The RTO performs its mission with the support of an extensive network of national experts. It also ensures effective coordination with other NATO bodies involved in R&T activities.

RTO reports both to the Military Committee of NATO and to the Conference of National Armament Directors. It comprises a Research and Technology Board (RTB) as the highest level of national representation and the Research and Technology Agency (RTA), a dedicated staff with its headquarters in Neuilly, near Paris, France. In order to facilitate contacts with the military users and other NATO activities, a small part of the RTA staff is located in NATO Headquarters in Brussels. The Brussels staff also coordinates RTO's cooperation with nations in Middle and Eastern Europe, to which RTO attaches particular importance especially as working together in the field of research is one of the more promising areas of initial cooperation.

The total spectrum of R&T activities is covered by 7 Panels, dealing with:

- SAS Studies, Analysis and Simulation
- SCI Systems Concepts and Integration
- SET Sensors and Electronics Technology
- IST Information Systems Technology
- AVT Applied Vehicle Technology
- HFM Human Factors and Medicine
- NSPG NATO Simulation Policy Group (Modelling and Simulation)

These Panels are made up of national representatives as well as generally recognised 'world class' scientists. The Panels also provide a communication link to military users and other NATO bodies. RTO's scientific and technological work is carried out by Technical Teams, created for specific activities and with a specific duration. Such Technical Teams can organise workshops, symposia, field trials, lecture series and training courses. An important function of these Technical Teams is to ensure the continuity of the expert networks.

RTO builds upon earlier cooperation in defence research and technology as set-up under the Advisory Group for Aerospace Research and Development (AGARD) and the Defence Research Group (DRG). AGARD and the DRG share common roots in that they were both established at the initiative of Dr Theodore von Kármán, a leading aerospace scientist, who early on recognised the importance of scientific support for the Allied Armed Forces. RTO is capitalising on these common roots in order to provide the Alliance and the NATO nations with a strong scientific and technological basis that will guarantee a solid base for the future.

The content of this publication has been reproduced
directly from material supplied by RTO or the authors.

Printed on recycled paper

Published May 1999

Copyright © RTO/NATO 1999
All Rights Reserved

ISBN 92-837-1014-2

Printed by Canada Communication Group Inc.
(A St. Joseph Corporation Company)
45 Sacré-Cœur Blvd., Hull (Québec), Canada K1A 0S7

Optical Gyros and their Application

(RTO AG-339)

Executive Summary

This AGARDograph is written by experts from France, Germany, Russia and the United States. Only the radical change of the political and military situation within the last decade made it possible to have the subject Optical Gyros covered by scientists and engineers from East and West. Until this change, Optical Gyros and their research and development were mostly handled as top secret. The more it can be seen as a giant step forward that we have in our hands an unclassified and nevertheless in depth treatment of Optical Gyros from scientists from East and West.

Optical Gyros replacing the fast rotating rotor gyros as sensors for angular displacement or angular rate opened new possibilities in applications as well as in production. Especially in the field of inertial navigation, Optical Gyros replaced the mechanical gyro. Using them in Strapdown Inertial Navigation Systems in combination with Satellite Navigation Systems, offers new systems for dual usage in military and civil applications.

This volume can be divided into four main parts. The first one gives a detailed description of the laser and fiber optical gyro theory, the main sources of errors and the methods used to reduce their influence. The analysis of many types of modern laser gyros allowing various requirements to be met is presented in the second part. The reader can also find here some areas of practical applications of laser gyros. The third part is devoted to fiber optical gyros that have been developed recently. They are so successful both technologically and commercially that they are in great demand. The last part of the book considers special applications of optical gyros.

Gyroscopes optiques et leurs applications

(RTO AG-339)

Synthèse

Cette AGARDographie a été rédigée par des spécialistes français, allemands, russes et américains. Il est à noter que seul le changement radical de la situation politique et militaire au cours de la dernière décennie a permis l'étude conjointe du sujet des gyroscopes optiques par des scientifiques et des ingénieurs des pays de l'est et des pays de l'Alliance. Avant ce changement, les gyroscopes optiques, ainsi que les activités de recherche et développement y associées, étaient classifiés très secret. D'ailleurs on peut mesurer le progrès accompli avec un tel document non-classifié, rédigé par des scientifiques des pays membres de l'OTAN et des pays de l'ancien pacte de Varsovie, traitant de façon approfondie les gyroscopes optiques.

La mise en oeuvre de gyroscopes optiques en remplacement des gyroscopes à rotation rapide en tant que capteurs de déplacement et de vitesse angulaires a ouvert de nouvelles possibilités d'application et de production. Dans le domaine de la navigation inertielle surtout, les gyroscopes optiques ont remplacé les gyroscopes mécaniques. Placés dans des centrales inertielles liées associées à des systèmes de navigation par satellite, ils sont devenus de nouveaux systèmes pour des applications militaires et civiles.

Ce volume peut être sous-divisé en quatre grandes parties. La première donne la description détaillée de la théorie des gyroscopes laser et des gyroscopes à fibres optiques, ainsi qu'une description des principales sources d'erreur et des méthodes utilisées pour les réduire. Dans la deuxième partie, les différents types de gyroscopes modernes sont analysés selon leurs différents critères d'utilisation. Le lecteur y trouvera également la description de certains domaines d'application des gyroscopes laser. La troisième partie est consacrée aux gyroscopes à fibres optiques récents. Ceux-ci sont très demandés en raison de leur grand succès technologique et commercial. La dernière partie de ce volume examine des applications spéciales relatives aux gyroscopes optiques.

Click inside the blue boxes to view the corresponding section

Contents

	Page
Executive Summary	**iii**
Synthèse	**iv**
List of Authors/Editors	**vi**
	Reference
1. Introduction by D. Loukianov, R. Rodloff, H. Sorg and B. Stieler	**1**
2. Physical Background and Technical Realization by R. Rodloff	**2**
3. Fundamentals of the Ring Laser Gyro by F. Aronowitz	**3**
4. Multioscillator Ring Laser Gyroscopes and their Applications by C.H. Volk, S.C. Gillespie, J.G. Mark and D.A. Tazartes	**4**
5. Zeeman Laser Gyroscopes by V.V. Azarova, Yu.D. Golyaev, V.G. Dmitriev, M.S. Drozdov, A.A. Kazakov, A.V. Melnikov, M.M. Nazarenko, V.N. Svirin, T.I. Soloviova and N.V. Tikhmenev	**5**
6. Laser Gyros With Total Reflection Prisms by Y.V. Bakin, G.N. Ziouzev and M.B. Lioudomirski	**6**
7. Application of the Sagnac Effect in the Interferometric Fiber-Optic Gyroscope by H.C. Lefèvre	**7**
8. Passive All-Fiber Open Loop Gyroscope by G. Trommer	**8**
9. Miniature Fiber Optic Gyro. Fizoptika Implementation by V. Listvin, V. Logozinski and V. Solomatin	**9**
10. Progress in Fiber-Optic Gyro Development and Applications by W. Auch	**10**
11. Progress in Fiber-Optic Gyroscope Applications II with Emphasis on the Theory of Depolarized Gyros by G.A. Sanders and B. Szafraniec	**11**
12. Dynamic Ring Laser Goniometer by Yu.V. Filatov, D.P. Loukianov, P.A. Pavlov, M.N. Burnashev and R. Probst	**12**
13. Gyroscopic Goniometry During Quasistatic and Kinematic Conditions Using a New and Conventional Methods for Aiding by B. Stieler	**13**
14. Application of a Laser Gyro in Track Measuring Systems by D.P. Loukianov, A.V. Mochalov and M. Rechel	**14**
15. A System for Measuring Deformations of Large-Sized Objects by A.V. Mochalov	**15**

List of Authors/Editors

Editors

Prof.Dr. D. Loukianov
St.Petersburg Electrotechnical University
Dept. of Navigation, Control & Mechanics
Prof. Popov St. 5
St.Petersburg, 197376
Russia

Prof.Dr. H. Sorg
Universität Stuttgart
Institut A für Mechanik
Pfaffenwaldring, 9
70550 Stuttgart
Germany

Dr. R. Rodloff
DLR
Institut für Flugführung
Lilienthalplatz, 7
38108 Braunschweig
Germany

Prof.Dr. B. Stieler
DLR
Institut für Flugführung
Lilienthalplatz, 7
38108 Braunschweig
Germany

Authors

F. Aronowitz (Chapter 3)
11430 Manzanitz Trail
Dewey, AZ 86327
USA

W. Auch (Chapter 10)
SFIM Industries Deutschland
Gottlieb-Daimler Str. 60
71711 Murr
Germany

V.V. Azarova, Yu.D. Golyaev, V.G. Dmitriev,
M.S. Drozdov, A.A. Kazakov, A.V. Melnikov,
M.M. Nazarenko, V.N. Svirin, T.I. Soloviova,
N.V. Tikhmenev (Chapter 5)
Research & Development Institute "Polyus"
3 Vvendensky str.
117342 Moscow
Russia

Y.V. Bakin, G.N. Ziouzev,
M.B. Lioudormirski (Chapter 6)
Bauman Moscow State Technical University
No. 5 2nd Baumanskaja St.
Moscow, 107005
Russia

Yu. V. Filatov, P.A. Pavlov,
M.N. Burnashev (Chapter 12)
Autonomous Navigation Department
St. Petersburg State Electrotechnical University
Professor Popov Str. 5
St. Petersburg, 197376
Russia

H.C. Lefèvre (Chapter 7)
EuroFOG-Photonetics
52, avenue de l'Europe, BP 39
78160 Marly-le-Roi
France

V. Listvin, V. Logozinski, V. Solomatin (Chapter 9)
Fizoptika Co.
Post Box 16
Moscow 109387
Russia

A.V. Mochalov (Chapters 14 & 15)
St. Petersburg Electrotechnical University
Professor Popov Str. 5
St. Petersburg 197376
Russia

R. Probst (Chapter 14)
Physikalisch Technische Bundesanstalt
Bundesallee 100
38116 Braunschweig

A. Sanders, B. Szafraniec (Chapter 11)
Honeywell Inc. HTC
21111 N. 19th Avenue
Phoenix, Arizona 85027
USA

B. Stieler (Chapter 13)
DLR
Institut für Flugführung
Lilienthalplatz, 7
38108 Braunschweig
Germany

G. Trommer (Chapter 8)
LFK (Lenkflugkörpersysteme-GmbH)
FTE 5
Postfach 80 11 49
81633 München
Germany

C.H. Volk, S.C. Gillespie, J.G. Mark,
D.A. Tazartes (Chapter 4)
Litton Guidance and Control Systems
5500 Canoga Ave
Woodland Hills, CA 91367-6698
USA

1. Introduction

by

D. Loukianov, R. Rodloff, H. Sorg, B. Stieler

The second half of the 20th century is marked not only by social changes; but also intensive and thorough investigations in fundamental and applied science. Two aspects are to be of the paramount importance in the field of navigation: 1) creation of global satellite navigation systems, such as Transit, GPS and GLONASS; 2) the advent of a new generation of inertial navigation sensors – laser and fiber optical gyros (LG and FOG). These optical devices, making use of fundamental properties of electromagnetic waves, provide a basis for a new line of investigations in inertial navigation and make it possible to give up the fast rotating rotor. This situation opens a possibility to use new progressive technology in mass production.

Exceptional properties of optical gyros (high accuracy, wide dynamic range, non-sensitivity to linear acceleration, et al.) stimulated the evolution of various highly-accurate strapdown inertial navigation systems (SINS) which are in great demand on the market today. Integration of SINS based on optical gyros with Satellite Navigation Systems, such as GPS and GLONASS, allows new positive properties necessary for dual usage in military and commercial technology. It is necessary to note considerable and extremely important progress in FOG in the last few years due to deeper understanding of physical processes in electromagnetic wave propagation in optical fiber and a wide use of integral optics. All this permits realization of an optimal architecture of FOG and using of micromechanical sensors to create a new modification of low life-cycle cost, small-sized and low production cost SINS.

The most important scientific and practical results in the field of laser and fiber gyros have been obtained through military and industrial complexes of leading states. This fact remained a serious obstacle to wide international cooperation for a long time, particularly in the sphere of modern high technologies. A new era was opened in 1990 after the historical Declaration at the London Summit Meeting in the light of the changes that had taken place in Europe. In 1991 the Alliance furthered this process with the Declaration on Peace and Cooperation and publication of a new Strategic Concept to bring NATO's overall strategy in line with future needs. As a result of these steps AGARD's Technical Cooperation Program was worked out and the initiative of publishing the monography "Optical Gyros and their Application" by an international body of authors was supported.

This volume consists of 15 chapters that are divided into four main parts. The first one gives a detailed description of the laser and fiber optical gyro theory, the main sources of errors and the methods used to reduce their influence. The analysis of many types of modern laser gyros (LG) allowing various requirements to accuracy, dimension, weight and cost to be satisfied is presented in the second part. The reader can also find here some areas of practical application of LG. The third part is devoted to fiber optical gyros (FOG) that have been developed recently. They are so successful technologically and commercially that are in great demand both in Russia and abroad. The last part of the book considers special applications of optical gyros for laser dynamic goniometry and metrology, fundamental and applied researches (such as nuclear physics, aerodynamics, etc.), railway track surveying and estimation of deformation of some objects.

The authors wish to thank those leading experts (France, Germany, Russia and United States) who have made significant contributions to every aspect of this AGARDograph.

2. Physical Background and Technical Realization.

R. Rodloff
German Aerospace Research Establishment
Institute of Flight Guidance
Braunschweig

2.1 Sagnac Effect

The situation we want to consider might be a little bit strange: An "observer" inside a rotating system has the task to get an information about the absolute rotation of his system without getting in touch with the outer world - and without using a mechanical gyro. In a first attempt he might think to measure centrifugal forces which act on a test specimen with the mass m. If he has additional knowlegde about the distance r between his position and the centre of rotation it should be easy to calculate the rotation rate ω by using the expression:

$$F_{centrifugal} = m \cdot r \cdot \Omega^2 .$$

But if the value of r is not known, - and generally that's the case - our "observer" has a problem: he can try to find this point inside the rotating system where the rotation dependent force will vanish [1],- that's the centre of rotation, - but if this point is not inside of his system he will fail.

A solution of this problem could be a "travelling time experiment", which enables the "observer" inside the rotating box to measure the travelling time of a signal which propagates around a closed loop. The result of this experiment depends (of course !) on the geometrical properties of the loop (circumference L or loop area A), on the signal velocity u and - at a glance hard to believe - on the rotation rate Ω of the entire system !

This is the core of the Sagnac effect.

In a second step this experiment can be improved by measuring the travelling time difference of two counterpropagating signals which use the same loop. The advantage is tremendous:

- the sensitivity against the rotation rate is doubled;
- the output is now independent from the signal velocity.

This is the way how the Sagnac effect is actually used !

2.1.1 Historical Review.

The first ringinterferometer and the first experiment concerning light propagation in rotating media were not carried out by G. Sagnac, but by F. Harress. /1/. Figure 2.1 shows his experiment as a redrawing from the original picture /2/. He used a rotating ring of prisms as a guiding system for the two counter propagating light wavetrains. Lightsource and readout optics in the centre of the ring are not rotating with the whole system , which was the reason for many problems[2]. The main goal of his experiments was the investigation of the dragging coefficient to : "... get a

[1] for the moment we will forget about the problem how to distinguish between the gravitational - , translational - and the rotational forces.
[2] readout difficulties for instance.

decision whether Fresnel's $(n^2-1)/\ n^2$ or Lorentz's $(n^2-1)/n^2-(\lambda/v)\cdot(dv/d\lambda)$ expression for the dragging coefficient was the right one." (ref. to /2/)

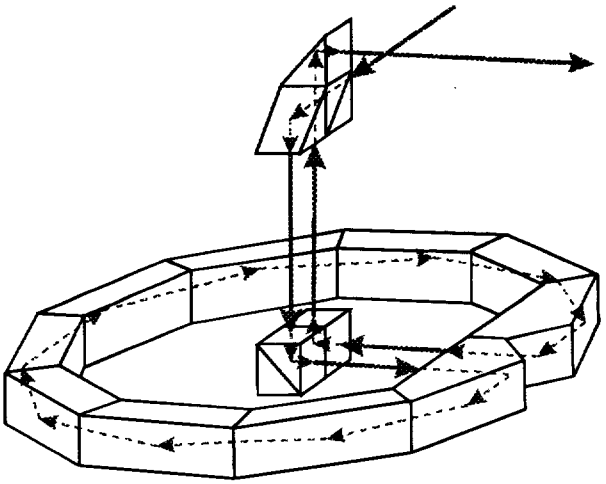


Fig. 2.1: Ringinterferometer used by Harress. (Redrawing from the original picture in /2/).

From today's point of view even this kind of question is something doubtful, but the most critical mistake of Harress' attempt was the assumption, that the pathlengths of the two counterpropagating wavetrains within the ringinterferometers are equal. Using this assumption it happened to him to neglect the main effect and as a consequence he tried to interpret his experimental results only on the basis of dragging effects. /5/ Additional mistakes in his theoretical discussion were later corrected by Paul Harzer /5/ and Albert Einstein /6,7/.

In 1920 Max von Laue wrote a paper which threw light on the fact that the time difference between the counterpropagating wavetrains within the rotating interferometer is completely independent from the refraction index of the glass prisms /8/. M. von Laue pointed out that "... the main reason that the propagation time difference is independent from the refraction index ... .. is the Lorentz-transformation". With this investigation the theoretical fundament of the observed propagation time difference was completed.

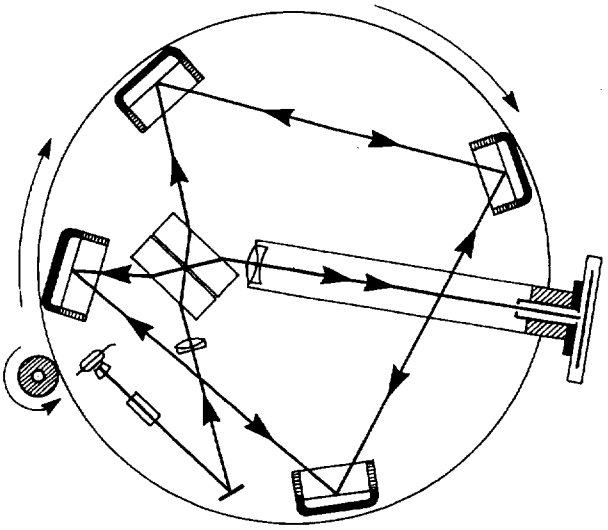


Fig. 2.2: Sagnac- Ringinterferometer (Redrawing from the original picture in /3/).

The mistakes Harress made when he tried to interpret his experimental results are possibly the reason that the basic effect for modern optical gyros are not named after him, but after G. Sagnac /3,4/, although the latter published his results two years later. Fig 2.2 shows the arrangement of Sagnac's interferometer.

In contrast with the ringinterferometer used by Harress (Fig.: 2.1) the medium for the counterpropagating light beams was not glass but air. Moreover the light source and the readout optics were corotating with the interferometer. To calculate the interference fringe shift he already used the correct formula, but he interpreted the result of his experiments as a proof of the existence of the so called 'ether' which is resting relative to the earth. He did that in spite the fact that Einstein showed already in 1905 that the concept of an 'ether' which serves as a carrier of electromagnetic wave was not necessary /13/. In the year 1913 the discussions concerning the 'ether' theory were not finished, - actually for some people it is still going on today ! -, and with his interpretation Sagnac was in good company; nevertheless it must be stated that he also failed in the attempt to find a correct physical description, which was first delivered by Max von Laue in 1920.

Discussions concerning the Sagnac-Effect at that time are superposed by a great number of experimental problems which prevent the view for the basic mechanisms. Experimental physicists raised many questions, such as : "how will the beam diameter influence the result ?" - "in which way will the readout optics distort the effect ?" - "is the law of reflection still valid in a rotating system ?" - "are the centrifugal forces of importance ?" - or: "are the small air gaps between the prisms of the Harress - interferometer a problem and will they distort the measured effect ?" All these problems were seriously discussed /5,8,9,10/, but beyond those more or less theoretical difficulties, the experimenters had a lot of practical problems with their devices. Especially the rather high rotation rates of the interferometers drove them to despair. B. Pogany for instance reports on an experiment with a 8o kg device which rotates with 1600 revolutions (!) per minute. At a critical rate of 600 rpm ".. the whole device was exposed to heavy vibrations and all optical parts were damaged." /9/ (In a final version Pogany rotated his ring interferometer with 2000 rpm !)

2.1.2 The Sagnac-Effect: a Simplified Model.

Figure 2.3a shows the basic elements of a ring interferometer. Two wavetrains, created by a beamsplitter, are travelling around the ring interferometer in opposite directions. If the beams are combined (superposed) after one circle they form an interference fringe pattern, which is made visible on a screen or registrated by means of a photodetector.

The light source, the beam guiding system (mirrors, prisms or glass fibre) , combining optics, screen and/or photodetector, - all these elements are mounted on a platform. If the whole system rotates around an axis perpendicular to the plane of the counter propagating wavetrains, the fringe pattern will be shifted proportional to the rotation rate.

The actual effect is based on a travelling time - , or phase-difference between the two wavetrains. This leads to a shift of the interference fringe pattern, and this again can easily be detected.

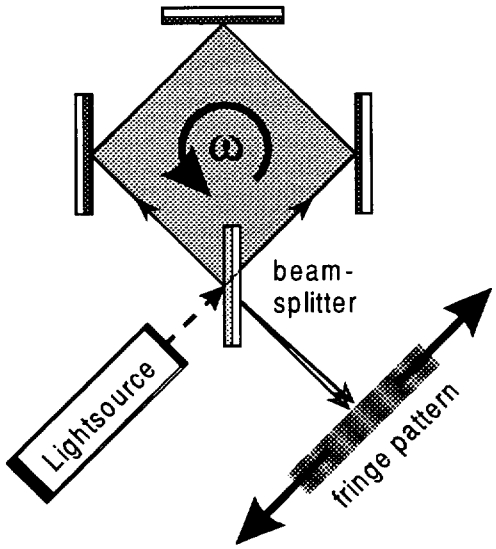


Fig.: 2.3a: Sketch of a Sagnac-ringinterferometer.

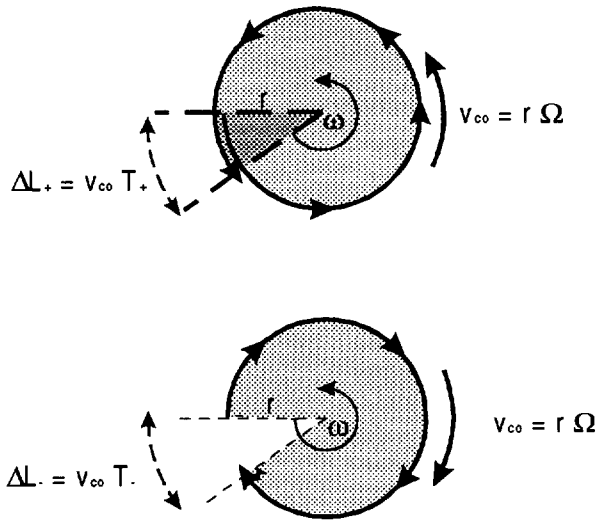


Fig.: 2.3b: Sketch for a simplified calculation of the Sagnac-effect.

A very simple calculation of the magnitude of this effect is based on circular beam guiding configuration (Fig 2.3 b). The following parameters are used:

Radius of the beam guiding system	r
Circumference of the beam guiding system:	$L = 2\pi \cdot r$
Area of the beam guiding system	$A = \pi \cdot r^2$
light - /signal velocity	c
Travelling time of the wavetrains for one circle	$T = L/c = 2\pi \cdot r/c$
rotation rate of the beam guiding system	ω
tangential velocity of the circular beam guiding system:	$v_{co} = r \cdot \Omega$

Let's first have a look on the beam which is travelling into the direction of rotation (Fig 2.3b, at the top). During the travelling time in the direction of rotation T_+ the beam guiding system itself moves a little bit (ΔL_+) and the overall pathlength of the beam increases to :

[2.1] $$L_+ = L + \Delta L_+ = L + v_{co} \cdot T_+ .$$

The necessary travelling time T_+ is given by the light beam velocity c and the total pathlength L_+ :

[2.2] $$T_+ = \frac{L_+}{c}$$

or:

[2.3] $$L_+ = T_+ \cdot c$$

Combining equation [2.1] and [2.3] we get the following expression for the travelling time T_+ :

$$T_+ \cdot c = L + v_{co} \cdot T_+ ,$$

or

[2.4] $$T_+ = \frac{L}{c - v_{co}} .$$ [3]

On the other hand the pathlength L_- for the counterpropagating wavetrain decreases, because the beam guiding system moves towards the direction of wavetrain movement (Fig 2.3b). In this case the travelling time T_- is given by:

[2.5] $$T_- = \frac{L}{c + v_{co}}$$

The travelling time difference between these two wavetrains is :

[2.6] $$\tau = T_+ - T_- = \frac{2 \cdot L \cdot v_{co}}{\left(c^2 - v_{co}{}^2\right)} = \frac{1}{\left(1 - \left(\frac{v_{co}}{c}\right)^2\right)} \cdot \frac{2 \cdot L \cdot v_{co}}{c^2} .$$

Due to the fact that the tangential velocity v_{co} is rather small compared with the velocity of light c, or ($\frac{v_{co}}{c} << 1$), we can simplify equation [2.6] to:

[2.7] $$\tau = \frac{2 \cdot L \cdot v_{co}}{c^2} = \frac{4 \cdot \pi \cdot r^2}{c^2} \cdot \omega = \frac{4 \cdot A}{c^2} \cdot \omega$$

with: $L = 2\pi \cdot r$ (circumference) and $A = \pi \cdot r^2$ (area of the ringinterferometer)

Equation [2.7] can easily be converted into an expression for the pathlength difference ΔL :

[2.8] $$\Delta L = c \cdot \tau = \frac{4 \cdot A}{c} \cdot \Omega$$

and for the phase difference, or phase shift $\Delta\varphi$:

[2.9] $$\Delta\varphi = 2\pi \cdot f \cdot \tau = \frac{8\pi \cdot A}{c \cdot \lambda} \cdot \Omega$$

with the light wave frequency : $f = c/\lambda$.

For practical purposes the relative shift Δz of the interference fringe pattern (Fig 2.3a) is much more interesting:

[2.10] $$\Delta z = \frac{\Delta\varphi}{2\pi} = \frac{4A}{c \cdot \lambda} \cdot \Omega .$$

The latter equation is probably the most important one, because it describes the actual observable effect.

The advantage of this simplified model for the Sagnac effect is, that the measurable effect is quite well described, - but the physical picture[4] behind is extremely incomplete and leaves a lot of important questions open. This approach fails completely if we are interested in the role of signal velocity: 'is it possible to use acoustic waves instead of light waves?'. Other questions, such as: 'why does this

[3] The term "$(c - v_{co})$" should not be interpreted as new kind of resulting signal velocity.
[4] it is something like a "race" between the wavetrain and a moving target.

effect only occur within rotating systems ?' , or: 'how important is the shape of the ringinterferometer and the position of the centre of rotation?' - are also not treated, although they are of particular interest for inertial navigation .

Anyhow - the simple Sagnac model discussed so far, serves quite well as an introduction and it gives a rather good feeling for the magnitude of this effect; here - as an example - the output data for a ringinterferometer with following characteristics;

resonator shape	circular
diameter	11,3 cm
area	100 cm^2
wavelength	6,33 10^{-5} cm
input rotation rate[5]	0,01 °/h = 2,8 10^{-6} °/sec

This type of ring interferometer delivers the following output values:

travelling time difference	2 10^{-26} sec
pathlength difference	6 10^{-16} cm
phase shift	2 10^{-10} rad
relative fringe shift [6]	1 10^{-11}

The most important value for the readout is the relative fringe shift, which is obviously too small to detect rotation rates in the order of 0,01°/h directly. The question behind all types of optical gyros, - laser gyro, glass fiber gyro, passive ringresonator gyro, - is the same: how to increase the sensitivity for small rotation rates. The approach of the two most common optical gyros is somewhat different : In case of the fiber gyro the enhancement is achieved by a kind of multiplication, or increase of the active area *A* (ref. to equation [2.4] - [2.8]). On the other hand: the idea behind the laser gyro is a transformation of the phase shift (eq. [2.6]) into a frequency shift. These different approaches result in different error mechanisms; while the fiber gyro suffers from an increase of basic ringinterferometer problems (noise, nonreciprocal phaseshifts caused by time and thermal dependent changes of the interferometer parameters), shows the laser gyro a completely new error behaviour, - the so called 'lock-in' effect, which is a typical problem of oscillating systems. (This will be discussed in detail in the next chapter by F. Aronowitz.)

2.1.3 Sagnac-Effect as a Relativistic Effect.

"The Sagnac effect is related to the propagation characteristics of a light wave within a rotating system."

Many authors use this, or a similar statement to justify their treatment of the (light)wave equation within the framework of General Relativity in order to describe the Sagnac effect. "The outcome of this, however, after some abstruse algebra /14/ , is a result which provides little insight into the physical situation and which yields an expression for the fringe-shift essentially identical to Eqn. .... " [2.10]. /14/

---

[5] Gyros for medium-quality inertial navigation systems must be able to measure rotation rates in the order of 0,01°/h !

[6] "relative fringe shift" is defined by the ratio of the absolute fringe shift to distance between two neighbouring fringes.

Due to the inherent acceleration of the rotating system, a treatment which is based on general relativity is - from a theoretical point of view - the only possible approach. But the use of special relativity - as an approximation ! - delivers the same numerical results and we achieve a much easier access to some basic mechanisms, especially the role of signal velocity, without highly sophisticated mathematics. It also seems likely to discuss the Sagnac effect on the basis of a (light)wave equation, because the effect is always realized by means of an optical ringinterferometer; but this approach prevents the reader from understanding that the effect is not only associated with the propagation of electromagnetic waves but with any kind of signal transmission. Therefore, the definition :

"The Sagnac effect is a runtime effect within a rotating systems."

represents a good compromise between physical correctness and the readers request for an understandable way of explanation.

The basic approach to evaluate the propagation time dt of a signal around any arbitrary contour is given with the following equation:

[2.11] $$dt = \frac{dL}{u}.$$

With u : "effective signal velocity",

dL : line element of the signal guiding contour.

For a further treatment of the problem we have taken into account, that most of the parameters are vectors; therefore we should rewrite equation [2.11] :

[2.12] $$dt = \frac{d\vec{L} \cdot \vec{e}}{\vec{u} \cdot \vec{e}}$$

with the scalar components $dL = d\vec{L} \cdot \vec{e}$, $u = \vec{u} \cdot \vec{e}$ and $\vec{e}$ as the unit vector directed along the signal guiding contour.

Figure 2.4 shows the basic (most relevant) elements of a Sagnac setup. The rotation rate of the whole system is $\vec{\Omega}$. The main problem for describing the Sagnac effect is related to the so called "effective signal velocity" $\vec{u}$ or the scalar component $u = \vec{u} \cdot \vec{e}$ as mentioned above. Just in terms of the classical picture of a "race" between the signal and a moving target, many authors combine the signal velocity $\vec{u}$ with the velocity $\vec{v}_{co}$ of the guiding system (relative to the observers position) to $\vec{u} \mp \vec{v}_{co}$.[7] In order to proceed cautiously we should also take into account, that the signal which is travelling into the clockwise direction ($\vec{u}_+$) might be different from the other one travelling counterclockwise ($\vec{u}_-$). Therefore the "effective signal velocity" $\vec{u}$ can be written as:

[2.23] $$\vec{u} = \vec{u}_\pm \mp \vec{v}_{co}$$

and the travelling time $dt_\pm$ becomes:

[2.14] $$dt_\pm = \frac{dL}{\vec{u} \cdot \vec{e}} = \frac{dL}{(\vec{u}_\pm \mp \vec{v}_{co}) \cdot \vec{e}} = \frac{dL}{u_\pm \mp v_{co}}$$

[7] This notation reflects only the picture of a race between the signal and the moving target. Even if we assume that the signal velocity $\vec{u}$ is identical with the light velocity $\vec{c}$, - this notation is **not** in contradiction to the basis of relativity !

with $u_{\pm}$ and v_{co} as the scalar components of $\vec{u}_{\pm}$ and $\vec{v}_{co}$ along the contour. The travelling time difference $d\tau$ is then:

[2.15] $$d\tau = dt_{+} - dt_{-} = \frac{dL}{u_{+} - v_{co}} - \frac{dL}{u_{-} + v_{co}}$$

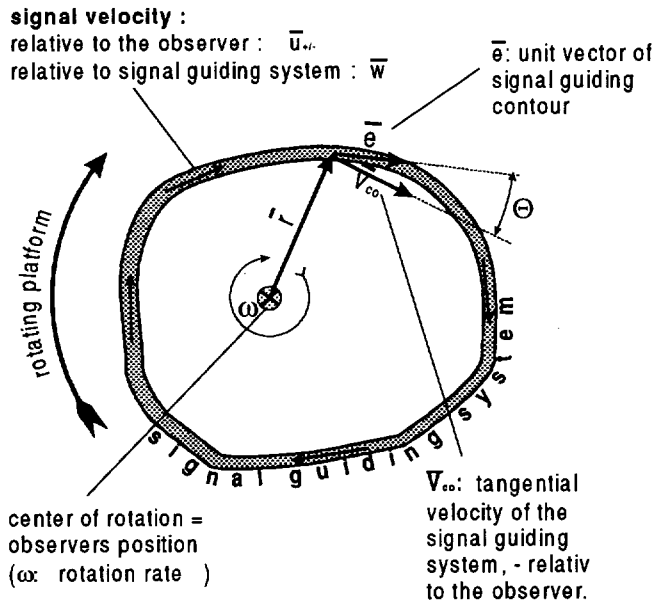


Fig. 2.4 : Sketch for the interpretation of the Sagnac effect.

If we assume for a moment that the velocity values of the two counterpropagating signals (u_{+} and u_{-}) are equal and moreover, if they are identical to the velocity of light ($u_{+} = u_{-} = c$) , we get:

[2.16] $$d\tau = dt_{+} - dt_{-} = \frac{dL}{c - v_{co}} - \frac{dL}{c + v_{co}} = \frac{2 \cdot v_{co}}{\left(c^2 - v_{co}^2\right)} \cdot dL = \frac{2 \cdot v_{co}}{c^2 \cdot \left(1 - \frac{v_{co}^2}{c^2}\right)} \cdot dL$$

and with the usual approximation $c >> v_{co}$[8], or: $\frac{v_{co}}{c} << 1$, the travelling time difference turns out to be:

[2.17] $$d\tau = dt_{+} - dt_{-} \approx \frac{2}{c^2} \cdot v_{co} \cdot dL,$$

or for a complete "round trip" we get for the total time difference τ :

[2.18] $$\tau = \frac{2}{c^2} \oint \vec{v}_{co} \cdot d\vec{L},$$

[8] signal velocity is large compared with the tangential velocity of the signal guiding system

which is exactly the same as we got with the "simplified" model above, - and actually: the only improvement we reached so far, is the use of a vector notation ($\vec{v}_{co} \cdot d\vec{s} = |\vec{v}_{co}| \cdot |d\vec{L}| \cdot \cos\Theta$ has to be interpreted as a scalar product), which allows us to calculate the Sagnac effect also for a noncircular signal guiding system and to answer the question how - and if so, how much - does the position of the rotation centre influence the Sagnac effect. We will postpone these problems to the following chapters, for the moment we should concentrate on another - may be much more important - question: What is the role of the signal ?

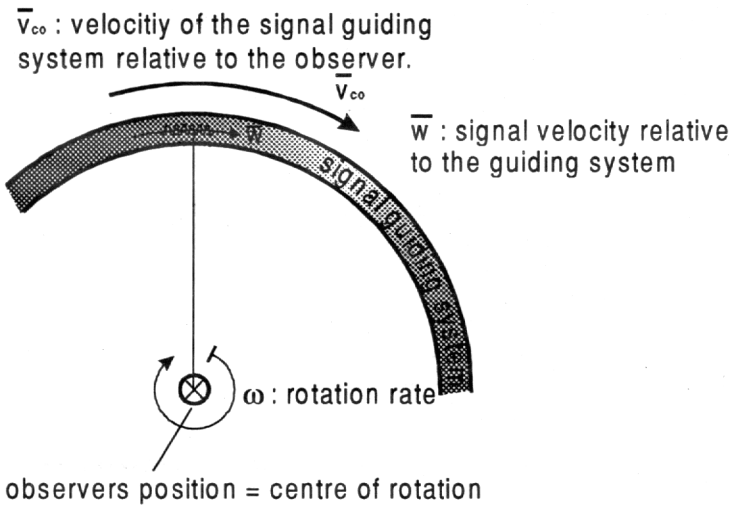


Fig. 2.5: Sketch for definition of signal velocity $u_{\pm}$.

Up to now, we did not use any information about the signal itself. The signal velocity *c* which is used in equation [2.8] and [2.16] could be the velocity of an electromagnetic wave, of an acoustic wave, or of any other kind of signal transmission. This leads to the following question:

What happens to the Sagnac effect if signal velocity is changed ?

At a glance it seems to be advantageous to reduce the signal velocity *c* and if we are able to replace the light wave with its vacuum velocity of $3*10^{10}$ cm/sec for instance by an acoustic wave with $3*10^{4}$ cm/sec it should be possible to increase the sensitivity of the Sagnac effect several orders of magnitude. [9]

But unfortunately: such a "Super Sagnac Effect" was never observed. On contrary: all experimental experiences only allow one interpretation, - the Sagnac effect is completely insensitive against any variation of the signal velocity !

At this point we have to leave the pure classical approach and try to a find an explanation of some general validity with out referring to a special kind of signal. Many authors /12,14/ discuss this phenomenon only regarding on the propagation characteristics of light waves; they show that the material contributions of refraction and dispersion within an optical ringinterferometer both cancel, but they leave their readers in the darkness whether the Sagnac effect is only related to electromagnetic waves, - or a fundamental effect within any rotating system.

Figure 2.5 defines some parameters which are necessary to handle this question.

We have derived the Sagnac effect in terms of a race between the signal (= light wave) and a moving target. We did that, by using the "effective" signal velocity $\vec{u} = \vec{u}_{\pm} \pm \vec{v}_{co}$, where $\vec{u}_{\pm}$ was the signal velocity and $\vec{v}_{co}$ was the velocity of the signal guiding system, both relative to the observer and -as already mentioned, - v_{co} is the scalar component along the guiding contour. The velocity of the signal relative to

[9] This consideration was probably the background for a patent specification which describes an acoustic Sagnac effect. (M. Böhm, "Einrichtung zur Messung von Drehgeschwindigkeiten", Offenlegungsschrift DE 33 26 038 A1)

guiding system is $\vec{w}$. In this case the values of $\vec{w}$ and the scalar component w along the contour are identical.

The problem now is: what is the relevant signal velocity $\vec{u}_{\pm}$ if we want to calculate the travelling time of the signal with respect to any resting or moving observer. Within the framework of classical mechanics we are tempted to use the relation :

[2.19] $$\vec{u}_{\pm/classic} = \vec{w} \pm \vec{v}_{co},$$

but since Einstein's very first paper about Special Relativity "Zur Elektrodynamik bewegter Körper" (Annalen der Physik 17, 1905, §5)) we know how to get the correct velocity of a signal within a moving reference system.

[2.20] $$u_{+/-} = \frac{\sqrt{w^2 + v_{co}{}^2 +/- 2wv_{co}\cos\theta - ((wv_{co}/c_0)\cdot\sin\theta)^2}}{1+/-(wv_{co}/c_0^2)\cdot\cos\theta}$$

with: $w = |\vec{w}|$ signal velocity relative to the guiding system; $v_{co} = |\vec{v}_{co}|$: velocity of the guiding system relative to the observer; Θ : angle between $\vec{w}$ and $\vec{v}_{co}$. (ref. to Fig. 2.4)

Equation [2.20] can easily be reduced to well known special cases: if the signal is identical with a vacuum light wave ($w = c_o$) the resulting velocity is (of course !):

[2.20a] $$u_{\pm} = c_0$$

and if the directions of movement for the reference system and the signal velocity are always identical ($\Theta = 0$), - which is the case for a circular signal guiding system with an observer in the centre (ref. to fig 2.3b) - , we get the well known theorem for the relativistic addition of velocities:

[2.20b] $$u_{\pm} = \frac{w \pm v_{co}}{1 \pm \frac{w \cdot v_{co}}{c_0^2}}$$

Putting expression [2.20b] for the resulting signal velocity $u_{\pm}$ into equation [2.15] for the travelling time difference $d\tau$, we get - without any approximation ! - a very interesting result :

[2.21] $$d\tau = \frac{1\cdot}{\left(1 - \left(\frac{v_{co}}{c_0}\right)^2\right)} \cdot \frac{2}{c_0^2} \cdot v_{co} \cdot dL,$$

or with the usual approximation $\frac{v_{co}}{c_0} << 1$:

[2.22] $$d\tau = \frac{2}{c_0^2} \cdot v_{co} \cdot dL,$$

and for a complete circle : (returning to vector notation)

[2.23] $$\tau = \frac{2}{c_0^2} \cdot \oint \vec{v}_{co} \cdot d\vec{L}$$

which is (nearly) the same expression as we already got (ref. to [2.17], [2.18]), but with an inconspicuous, but actually very important difference: we have lost the

signal velocity 'w', (or 'c' as we called it in equation [2.17]) ! Instead of this, the vacuum light velocity 'c_0' comes into the game, - but as a constant and **not** as the signal velocity!

In terms of special relativity , - equation [2.21] is valid for a "stationary observer" , which refers to an observer position outside the signal guiding system. For a moving observer we have to apply the well established expression for time dilation : $1/\sqrt{1-(v_{co}/c_0)^2}$ /11/.

Now we are able to answer the above raised question concerning the role of signal velocity: it is useless to manipulate the signal velocity in order to increase the travelling time difference, because the signal velocity w will not even appear among those parameters which influence the travelling time difference τ !

Before we turn our attention to some more practical aspects, we should stay one moment with another rather obvious question: If it is true, that only the relativistic approach can be used for the description of the Sagnac effect, - and if it is also true that signal velocity plays no part in this game , - if these conclusions are really correct, - why is the simplified model used in Chapter 2.1.2 able to deliver the right result ?

The answer is a little bit disappointing and somewhat trivial. Actually we made several mistakes which cancel each other out: without taking any notice of relativity, or the problem how to describe the signal velocity within a moving frame, we simply assumed, that the signal velocity $\vec{u}_\pm$ is identical with the light velocity c and c_0 respectively:

$$u_\pm = c_0 ,$$

and with this naive approach we already applied the results of special relativity: the assumption for the signal velocity $u_\pm = c_0$ is nothing else than a special case of the relativistic addition theorem [2.20a], if one of the velocities is identical with the vacuum light velocity c_0. If we had really used the approach of classical mechanics [2.19]: $\vec{u}_{\pm/classic} = \vec{w} \pm \vec{v}_{co}$ (with $\vec{w}$ = signal velocity relative to the guiding system; $\vec{v}_{co}$ = velocity of the guiding system relative to the observer) and if we had applied this to the expression [2.15] for the travelling time difference $d\tau_{classic}$, we would have obtained :

$$[2.23] \qquad d\tau_{classic} = dt_+ - dt_- = \frac{dL}{u_+ - v_{co}} - \frac{dL}{u_- + v_{co}} = \frac{dL}{w + v_{co} - v_{co}} - \frac{dL}{w - v_{co} + v_{co}} = 0$$

and that means, - within classical mechanics the Sagnac effect will vanish. Or in other words: the Sagnac effect is a pure relativistic effect !

2.1.4 Practical Aspects, - or: How to use the Sagnac Effect.

The travelling time difference τ between two counterpropagating signals within a rotating system does not depend on the type of signal (ref. to [2.22]) and for this reason, the effect itself is not touched by any parameter which influences the travelling characteristic of the signal, such as the refractive index of the lightwave guiding material. (At least within the framework of the approximations used in this paper).

But actually nobody is able to measure the Sagnac effect in terms of time units. A rotation rate of 10°/h for instance yields a travelling time difference of 2 10^{-23} sec (ref. to section 2.1.2) , - a value which is far beyond any possibility for time measurement techniques. If there is any hope at all, to measure such an extreme small effect, we have to transform the travelling time difference τ into a quantity which is large enough for todays measurement technology. If the signal of the Sagnac setup is realized by means of a electromagnetic wave we should rather detect the phaseshift and the phaseshift difference φ respectively instead of the travelling time difference τ :

[2.24] $$\varphi = \tau \cdot (2\pi \cdot f) = \frac{4\pi \cdot f}{c_0^2} \cdot \oint v_{co} \cdot dL ,$$

with f : lightwave frequency.

For a ring-shaped Sagnac setup the integral of equation [2.24] will be simplified to:

$$\oint v_{co} \cdot dL = v_{co} \cdot L = \omega \cdot r \cdot 2\pi \cdot r = 2 \cdot A \cdot \Omega$$

with Ω: rotation rate ; A: enclosed area; L : circumference. Now the expression for the phaseshift looks like:

$$\varphi = \frac{8\pi \cdot A \cdot f}{c_0^2} \cdot \Omega .$$

In practice it is rather easy to measure the relative phaseshift of a wave train by superposing it with a reference wave of the same frequency f (interferometer principle) and it is obviously very advantageous to use a frequency as high as possible.[10] Within the Sagnac ringinterferometer the reference wave is realized by a counter propagating wave train. The shift of the interference fringe pattern which occurs as a result of rotation is:

[2.25] $$z = \frac{\varphi}{2\pi} = 2 \cdot \frac{f}{c_0^2} \cdot \oint v_{co} \cdot dL$$

and for the special case of a ring-shaped Sagnac interferometer we get:

$$z = \frac{4A \cdot f}{c_0^2} \cdot \Omega .$$

In contrast to the travelling time difference τ (ref. to equation [2.23]) now a signal characteristic, - in terms of the light wave frequency f -, comes into the game and because of this, the Sagnac effect is no longer independent from the kind of signal. On the other hand: the frequency f of a light wave will not be changed if the refractive index of the carrier medium varies and therefore we can still state, that the Sagnac effect is independent of a comoving medium. If we replace f by the free-space wavelength λ_0 and the free-space light velocity c_0 using the relation $f = \frac{c_0}{\lambda_0}$ we get a much more familiar expression for the fringe shift z :

[2.26] $$z = \frac{2}{c_0 \cdot \lambda_0} \cdot \oint v_{co} \cdot dL ,$$

[10] Thats exactly in contrast to the approach of the "acoustic gyro" we discussed above.

or: $$z = \frac{4A}{c_0 \cdot \lambda_0} \cdot \Omega$$

for a ring-shaped interferometer

Two other very important questions are still open:

a.) How does a Sagnac setup react on an uniform velocity (translational motion).

b.) Does the Sagnac effect depend on the position of the centre of rotation ?

To answer the first question we can use a quite formal argument: for an uniform translational velocity field it is easy to show that the term $\oint \vec{v}_{co} \cdot dL$ in eq. [2.26] is always zero. In this case the fringe shift z will vanish and we have no Sagnac effect! This argument is very often used in the literature (see for example: /11/), but within the framework of special relativity, we have a much more plausible reason : $\vec{v}_{co}$ describes the relative velocity between the signal guiding system and the observers position:

$$\vec{v}_{co} = \vec{v}_{observer} - \vec{v}_{guidingsystem}$$

and within a uniform translational velocity field $\vec{v}_{observer}$ is always equal to $\vec{v}_{guidingsystem}$, which leads to :

$$\vec{v}_{co} = 0 \ .$$

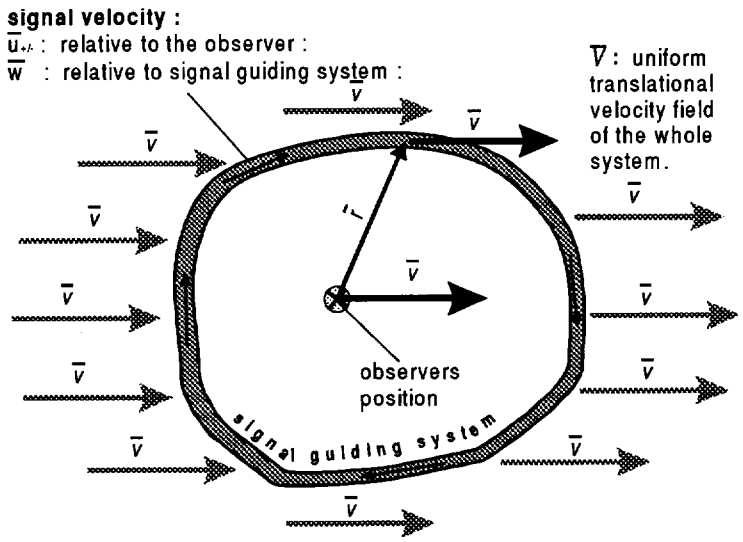


Fig.: 2.6: Sagnac setup within a uniform translational velocity field.

To answer the second question we should recall the basic equation for the Sagnac effect:

[2.23] $$\tau = \frac{2}{c_0^2} \cdot \oint \vec{v}_{co} \cdot d\vec{L} \ ,$$

which includes the relative velocity $\vec{v}_{co}$ between the observers position and the signal guiding system. In Figure 2.4 this velocity was identical with the relative velocity between the centre of rotation and the signal guiding system. Now we have to distinguish between the velocities of the signal guiding system relative to the

centre of rotation $\vec{v}_c$ and relative to the observer $\vec{v}_o$. Figure 2.7 shows the dependencies between these various velocities : the relative velocity between the centre of rotation and the signal guiding system $\vec{v}_c$ is:

[2.27] $$\vec{v}_c = (\vec{r} \times \vec{\omega});$$

the relative velocity between the centre of rotation and the observer $\vec{v}_{observer}$ is:

[2.28] $$\vec{v}_{observer} = (\vec{a} \times \vec{\omega})$$

and from this we can derive an expression for the relative velocity between the observer and the signal guiding system $\vec{v}_o$ (ref. to fig. 2.7):

[2.29] $$\vec{v}_o = \vec{v}_c - \vec{v}_{observer} = \vec{v}_c - (\vec{a} \times \vec{\omega}).$$

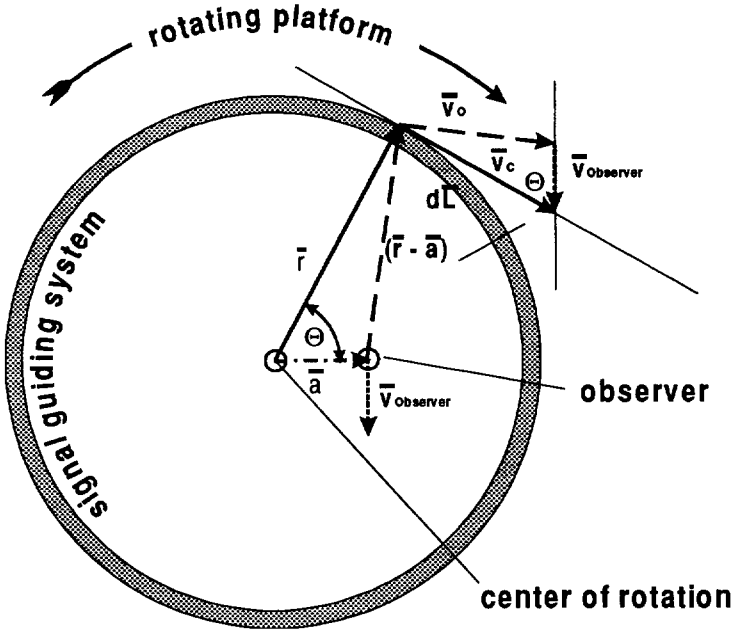


Fig. 2.7: Sagnac set up with different positions for the centre of rotation and the observer.

$\vec{r}$: vector from the centre of rotation to the signal guiding system;
$\vec{a}$: vector from the centre of rotation to the observer;
$(\vec{r} - \vec{a})$: vector from the observer to the signal guiding system;
$\vec{v}_C$: velocity of the signal guiding system relative to the centre of rotation
$\vec{v}_0$: velocity of the signal guiding system relative to the observer
$\vec{v}_{observer}$: velocity of the **observer** relative to the centre of rotation.

With the knowledge of the dependency between these relative velocities we are able to calculate the travelling time difference τ using equations [2.18, 2.21], but instead of the v_{co} we now have to use the relative velocity between the *observer and the signal guiding system* :

$$\vec{v}_o = \vec{v}_c - \vec{v}_{observer}$$

This leads to a travelling time difference τ_O for an observer who is not in the centre of rotation :

$$\tau_O = \frac{2}{c_0^{\,2}} \oint \vec{v}_o \cdot d\vec{L} = \frac{2}{c_0^2} \oint (\vec{v}_c - \vec{v}_{observer}) \cdot d\vec{L} \,,$$

or:

[2.30] $$\tau_O = \frac{2}{c_0^2} \cdot \oint \vec{v}_c \cdot d\vec{L} - \frac{2}{c_0^2} \cdot \oint \vec{v}_{observer} \cdot d\vec{L}$$

The first term in this equation : $\frac{2}{c_0^2} \oint \vec{v}_c \cdot d\vec{L}$ describes the Sagnac effect relative to the centre of rotation, - we are already familiar with this result. (Ref. to Eq. [2.23]) The second term $\oint \vec{v}_{observer} \cdot d\vec{L}$, which is related to the observers position , is zero. This can easily be examined if we take into account that $\vec{v}_{observer}$ is a constant, $d\vec{L}$ can be replaced by : "$\vec{r} \cdot d\Theta$" (ref. to fig. 2.7) and $\oint \cos\Theta \cdot d\Theta$ is zero.

$$\oint \vec{v}_{observer} \cdot d\vec{L} = \oint |\vec{v}_{observer}| \cdot |d\vec{L}| \cdot \cos\Theta = |\vec{v}_{observer}| \cdot r \cdot \oint \cos\Theta \cdot d\Theta = 0$$

This demonstrates, that the Sagnac effect is independent of the observers position

$$\tau_O = \frac{2}{c_0^2} \cdot \oint \vec{v}_c \cdot d\vec{L} - \frac{2}{c_0^2} \cdot \oint \vec{v}_{observer} \cdot d\vec{L} = \frac{2}{c_o^2} \oint \vec{v}_c \cdot d\vec{L} = \tau \,.$$

May be this is the most important feature of the Sagnac effect, because otherwise it would not be usable for inertial navigation.

2.1.5. References

/1/	Harreß, F.	"Die Geschwindigkeit des Lichtes in bewegten Körpern", Dissertation, Jena
/2/	Knopf,O.	"Die Versuche von F.Harreß über die Geschwindigkeit des Lichtes in bewegten Körpern." Annalen der Physik, IV Band 62, S.389 (1920)
/3/	Sagnac,G.	G.Compt.rend. 157, S.708 + 1410 (1913)
/4/	Sagnac, G.	G. Journ. d. Phys. (5), 4, S. 177, (1914)
/5/	Harzer, P.	"Über die Mitführung des Lichtes in Glas und die Aberration" Astronomische Nachrichten, Bd. 198, Nr. 4748, S. 378 (1914)
/6/	Einstein, A.	"Bemerkungen zu P.Harzers Abhandlung ' Über die Mitführung des Lichtes in Glas und die Aberration' ". Astronomische Nachrichten, Nr.4753, S. 7, (1914)
/7/	Harzer, P.	"Bemerkungen zu meinem Artikel in Nr. 4748 im Zusammenhang mit den vorstehenden Bemerkungen des Herrn Einstein." Astronomische Nachrichten, Nr. 4753, S. 9, (1914)
/8/	Laue, M.v.	"Zum Versuch von Harreß". Annalen der Physik, IV, Bd. 80, S. 217, (1920)
/9/	Pogany, B.	"Über die Wiederholung des Harress-Sagnacschen Versuchs."

Annalen der Physik, IV, Bd. 80, S. 217 (1926)

/10/ Pogany, B. "Über die Wiederholung des Harressschen Versuches." Annalen der Physik, IV Bd. 85, (1928)

/11/ Post, E.J. "Sagnac Effect." Review of modern Physics, Vol. 39, No.2, S. 475 (1967)

/12/ Post, E.J. "Interferometric Path-Lenghth Changes Due to Motion." Journ. o. Optical Soc. of America, Vol. 62, No.2, S. 234, (1972)

/13/ Einstein, A. "Zur Elektrodynamik bewegter Körper", Annalen der Physik 17, S. 891-921, (1905)

/14/ Wilkinson, J.R. "Ring Lasers" Progress in Quantum Electronics, Vol. 11, No.1, p. 1-103,(1987); Pergamon Press, Oxford, New York.

3. FUNDAMENTALS OF THE RING LASER GYRO

Dr. Frederick Aronowitz
11430 Manzanita Trail
Dewey, AZ 86327, U.S.A.

3.1 INTRODUCTION

One of the many successful uses of the laser is as a rotation sensor [1-5]. In the past, the usual approach in the design of gyroscopes was to use a spinning mass mounted on stable element, so that the gyroscope was fixed in inertial space and through the use of gimbals, isolated from the vehicle motion. With the advent of the laser, a clever implementation made use of a relatively little known physical mechanism[6] to allow the laser to sense rotation. The rotation sensing is possible in the frame of the laser, thus allowing the device to be measured directly to the vehicle, and avoiding the need for gimbals[7]. In that the laser was configured with a closed contour for the cavity, it has been commonly described as a ring laser gyro, or sometimes simply as a RLG.

The RLG was first operated[3] in 1962 and over the past 37 years, it has been investigated by many university and industrial laboratories. It was put into large scale production[8] in 1978 by Honeywell, as a strapdown inertial reference unit for aircraft applications. It is now considered a mature technology and there have been many publications describing the various form of its implementation. There have been two works which describe the development of the RLG from a historic standpoint[9-10]. Because of the large number of publications describing the RLG, it would not be possible to present a complete list of references in this review. There have been previously published review articles [11-13] on the RLG, which can be referenced for a more complete list of references.

It should be noted that although the RLG has been in production over the past 21 years by numerous companies, many techniques and processes have been maintained as proprietary, and details remain unpublished. Most or all, are known by each of the manufacturers. The theoretical basis for the operation of the RLG is well known worldwide. The basic theory is straight-forward and shared by all, although in many aspects, designs are based on empirical data. It is this data base which still remains protected.

There are still some areas which have not been adequately explained, and with the shift in emphasis in sensor development, may remain as such. In other areas, the description of the RLG is of such complexity, that the model tends to be used only to provide guidance in the design of the device. A major impediment to the further theoretical development of the RLG is that there is a scarcity of publications of data that has been obtained by the manufacturers in many of the areas of physics of the RLG. The second is the difficulty of universities obtaining this data first-hand, because of the cost and complexity of fabricating ring lasers having the capability of operating with the high level of performance required to obtain meaningful data.

In this review, a simple description of the operation of the RLG will be presented. The theoretical model, as it is known, will be presented, with a discussion on the distinction between what is known about the model and how it is typically used. Scale factor of the RLG and units will be covered. Typical errors will be discussed, with emphasis on coupling and the lock-in mechanism. The mechanical dither technique to avoid the lock-in problem will be discussed in detail. Other techniques used to avoid lock-in will be discussed, but with less detail. The fundamental limit of operation will be covered. A comparison of the various approaches to fabrication of the RLG will be made.

3.2 PRINCIPLE OF OPERATION

3.2.1 Summary Of Concept Of Operation

The concept of operation of the laser gyro can be summarized in simple form. In a linear laser, as represented by the schematic shown in Fig 2.1, the frequency of oscillation is determined by the condition that the cavity consists of an integral number of wavelengths. The cavity is defined as the position at which the beam repeats itself. Thus the linear laser cavity consists of a double pass of the distance between the mirrors. Since the beam must replicate itself for successive passes over the cavity length, there will be nodes at the two mirrors and the two oppositely directed beams comprise a standing wave.

When the cavity is of a ring configuration, as represented by the schematic shown in Fig 2.2, the two oppositely directed beams can be considered traveling waves, and the node constraint at the mirrors no longer exists. Thus they can be independent of each other and each can oscillate at a different frequency and a different amplitude. The oscillation frequency of each is determined by the optical, as opposed to geometric cavity path length. Thus any physical mechanism that causes the optical paths to be different for the two beams, results in the oscillation frequency being different.

REPRESENTATION OF LINEAR LASER OPERATION

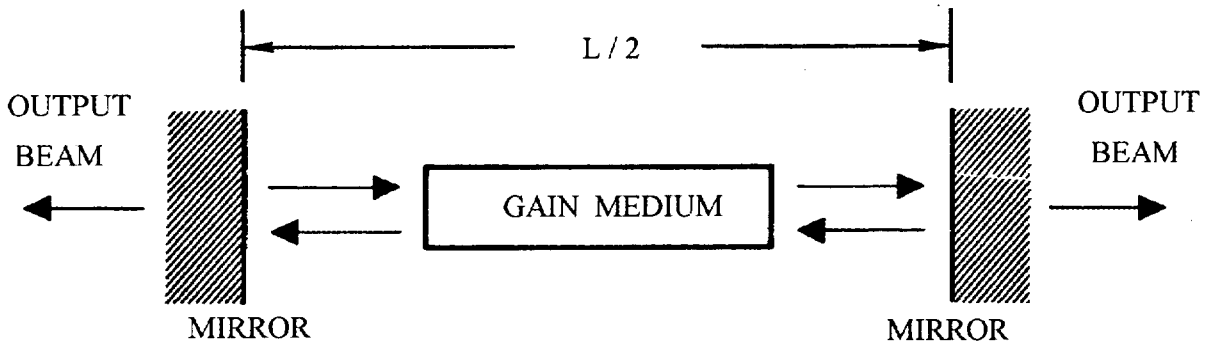


Figure 2.1. The figure shows a schematic representation of a simple linear laser. A medium, located between two mirrors, provides amplification of spontaneous emission. This causes a standing wave to build up in the cavity between the mirrors. A small fraction of the light passes through the glass substrates, upon which the mirrors are located. The transmitted light is the useful output of the laser.

REPRESENTATION OF RING LASER OPERATION

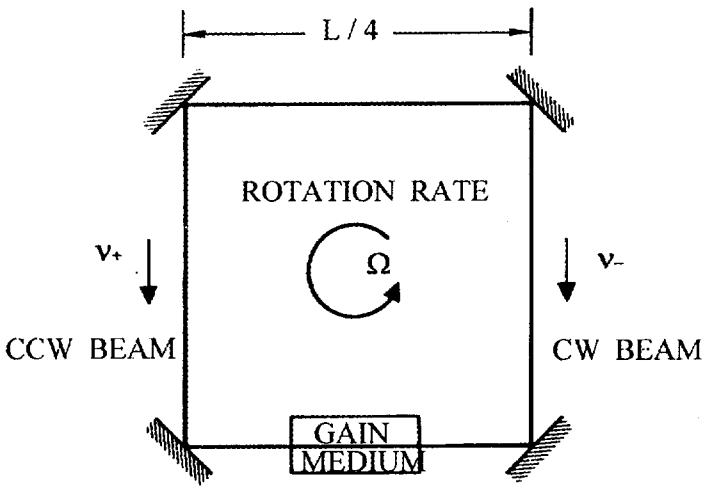


Figure 2.2. The figure shows a schematic representation of a simple ring laser of square configuration. The two oppositely directed beams, which are traveling waves, can operate independently of one another.

As demonstrated by Sagnac[6], when the ring cavity is rotated, the degeneracy in optical path is removed. The oppositely directed beams oscillate at different frequencies and the frequency difference is proportional to the rotation rate of the cavity. Thus by measuring the frequency difference, the rotation of the laser cavity (and any vehicle to which it is mounted) can be determined. The above enumerated principle of operation is conceptually simple. The analysis of the laser gyro to first order, is also straight-forward and simple. This treatment will be described in the next subsection. However, because of the very high accuracy potential of the RLG, and its corresponding sensitivity to many physical phenomena, error analysis can be complicated. In what follows, some of these mechanisms are discussed.

3.2.2 Ring Laser Interferometer

The oscillation conditions for a laser, either linear or of ring configuration, are two-fold. The first is that at the frequency of oscillation, the loss of the cavity is balanced by the gain of the oscillator. The second is a closure condition, which is that there are an integral number of wavelengths for the light making one cavity pass. For both the simple linear laser and ring laser, as illustrated in Figs 2.1 and 2.2, the closure condition is given by

(2.1) $m\lambda = L$

where **m** is an integer, λ is the vacuum wavelength and **L** is the optical path for the beam to return on itself. For the linear laser, the cavity modes consist of two oppositely directed traveling waves with equal amplitudes and frequencies, which then comprise a standing wave. In the ring laser, the oppositely directed beams (beams traveling in the clockwise and counter clockwise) can be independent in that each can oscillate with different amplitudes and frequencies. Then Eq (2.1) can be written for the ring laser in terms of the frequencies ($\nu_\pm$) of the two beams as

(2.2) $\nu_\pm = m c / L_\pm$

where the (+) and (-) refer to the counter clockwise and clockwise beams, respectively, and $L_\pm$ represents the optical path for each of the two respective beams. From Eq (2.2), the relative frequency difference can be written as

(2.3) $\Delta\nu / \nu = \Delta L / L$

Equation (2.3) shows that to operate a ring laser with the two oppositely directed beams having different frequencies, some mechanism must be provided such as to cause the optical paths for the two beams to be different.

3.2.3 Sagnac Effect

This subsection gives a simple description of the Sagnac effect, to provide continuity for this section. A more detailed description can be found in section 2 in this monograph, written by Dr. R. Rodloff.

The Sagnac effect can be shown to provide a removal of the path length degeneracy between the two beams. As can be shown[14], the effect is first order in **v/c**, and hence can be evaluated classically.

REPRESENTATION OF PASSIVE SAGNAC INTERFEROMETER

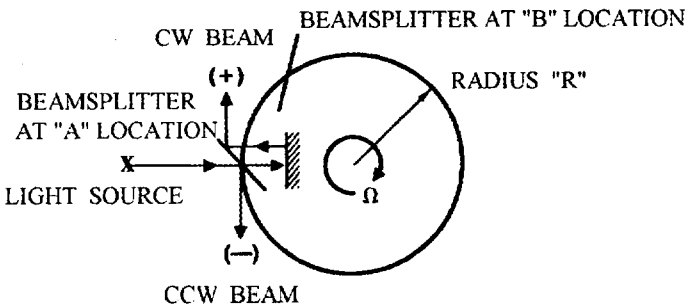


Figure 2.3. The figure shows a simple representation of the classic Sagnac Interferometer

Figure 2.3 shows a simple representation of the effect. It shows a source, located on a rotating frame, striking a beamsplitter at point A. The (+) beam travels through the beamsplitter and is retro-reflected back on itself. It then strikes the beamsplitter and is reflected, so that it can propagate along the interferometer (which in this illustrative case is circular for simplicity) in the clockwise direction. The (-) beam is reflected by the beamsplitter so that it can propagate along the interferometer in the clockwise direction. Since the interferometer is rotating, at the time in which the beam closes upon itself, point A has moved in inertial space to point B. From the geometry of Figure 2.3, it can be seen that the (+) beam traveling along the direction of rotation has traveled a greater distance than the (-) beam traveling against the direction of rotation. Since the velocity of light is constant, the transit time for the light traveling in the two directions of rotation can be shown to be

(2.4) $t_+ = 2\pi R / (c - R\Omega)$ $\qquad$ $t_- = 2\pi R / (c + R\Omega)$

From Eq (2.4), the transit time difference to first order, can be obtained as

(2.5) $\Delta t = t_+ - t_- = 4A\Omega / c^2$

where **A** is the enclosed area traversed by the two beams. The transit time difference Δt is the fundamental quantity to the operation of Sagnac interferometers. Although the expression has been simply derived using classical theory, a derivation using general relativity gives the same result,[12,15] and shows the effect to be independent of the location of the center of rotation.

The path difference, which is required for Eq (2.3), is obtained by multiplying Eq (2.5) by the velocity of light, giving

(2.6) $\Delta L = 4 A \Omega / c$

Equation (2.6) is the final expression for the path difference for light traveling in opposite directions around a closed contour enclosing an area **A**, and rotating at an angular rate Ω. To first order, it is independent of the index of refraction of the material being traversed. This assumes that optical path symmetry exists for light traveling in the two directions. If not, second order terms are introduced[16], which commonly result in large errors in practical devices.

The magnitudes of the differential time and path length, as given by Eqs (2.5, 2.6), are extremely small. For example, consider an ideal circular Sagnac interferometer of radius 10 cm, being used to measure earth's rotation rate. Only the rotation component normal to the area of the effective perimeter of the device is measured. For a placement on the surface of the earth with the sensitive axis being vertical, the vertical component of earth's rate is measured. At a latitude of approximately 45°, the vertical component of earth's rotation is 10 deg/hr (= 4.9×10^{-5} rad/sec). Then using Eq (2.4), the transit time for light making one pass around the interferometer is 2.1×10^{-9} sec. The differential transit time, as given by Eq (2.5) is 6.8×10^{-23} sec, which by most standards, is an extremely small time. Using Eq (2.6), this time difference is equivalent to a path difference between the two beams of 2.0×10^{-12} cm. At the helium-neon wavelength of 0.633 μm, which is the transition typically used in the RLG, the fraction of a wavelength that corresponds to the above path difference is 3.2×10^{-8}.

For typical aircraft navigation applications, an accuracy on the order of 1/1000 of earth's rate is required[17]. For these applications, the above quantities are reduced by the same factor, resulting in extremely stringent measurement requirements. In practice, an accurate RLG can readily provide measurements[18] in the range of 10^{-4} deg/hr, which corresponds to a measured path difference of approximately 1/1000 of the radius of the electron.

3.2.4 Ring Laser Gyro

Equation (2.3), which relates the frequency difference between the two beams, demonstrates one of the major advantages of using the ring as an active interferometer. As an active device, a small path length difference translates into a more readily measurable frequency difference. In that frequency, as shown by Eq (2.3), is a multiplication factor of path length difference, the advantage of using optical over microwave devices is apparent. Note that the relative frequency measurement still requires the same stringent accuracy.

Using Eq (2.6), Eq (2.3) gives the fundamental RLG equation as

(2.7) $\Delta \nu = 4A \Omega / (\lambda L)$

Equation (2.7) shows the RLG frequency difference, commonly called beat frequency, to be proportional to the product of the geometric area enclosed by the light beams and the angular rate of the cavity. The beat frequency is also inversely proportional to the product of the vacuum wave length of the laser and the optical path of the cavity. Optical path implies sources of scale factor error. Equation (2.7) has been obtained with the assumption of no other sources of optical path asymmetry between the two beams other than rotation. As will be discussed, there are numerous mechanisms that violate this assumption in a practical device.

As an illustration of the enhanced sensitivity of the active ring interferometer, consider the same numerical values that were used in section 3.2.3 to describe the Sagnac effect. With the passive device, a differential path length measurement of 2.0×10^{-12} cm was required The active RLG requires a measurement of 15 Hz, which is available without a complex hardware design. The relative frequency measurement is still the small value of 3.2×10^{-14}. Note that this requires the stability of the differential optical path (index of refraction) to be of this magnitude. This stability requirement places stringent restrictions on the construction of the rate sensor, and applies to the passive sensor as well as the active RLG.

For the above numerical example, the RLG scale factor, relating the beat frequency in Hz to the angular rate in either deg/hr or deg/sec, is given by

(2.8) $\Delta \nu_{Hz} = 1.53\, \Omega_{deg/hr} = 5.5 \times 10^{3}\, \Omega_{deg/sec}$

Thus for highly accurate gyros having performance capability on the order of 10^{-4} deg/hr, Eq (2.8) requires a beat frequency measurement on the order of 0.0002 Hz. It also shows that for high rotation applications[19] requiring input rates on the order of factors of 10^3 deg/sec, bandwidth requirement is on the order of tens of MHz, which is readily obtainable with today's technology.

The RLG has been commonly described as an integrating rate gyro, with a digital output. This can be seen using Eq (2.7), which when integrated gives

(2.9) $N = (4A / \lambda L)\,\theta$

where **N** represents the number of cycles of the beat frequency that is obtained when the RLG is rotated through an angle θ. With the numerical example previously used, Eq (2.9) shows that the gyro must be rotated through an angle of 3.2×10^{-6} rad to obtain an output of one count. This angle is commonly described as the inherent scale factor (SF) of the RLG and is usually expressed in units of arc-sec, which for this case, is 0.65 arc-sec/count. For systems usage, the inverse, or 1.5 counts/arc-sec, is typically used. For SF experiments, the gating angle is typically a full revolution. With the numerical values of the previous example, an output of 2.0×10^6 counts is obtained for one revolution of motion. The SF of the RLG is of sufficient accuracy that over one, ten and even one hundred revolutions, the SF is quantization limited by $\pm$ 1 count[20].

The above description of the SF, as derived from Eq (2.7) is straight forward, but the units can be confusing. It is useful to scale the SF for later usage. In so doing, the area in Eq (2.7) can be expressed in terms of a dimension-less parameter **H** by writing

(2.10) $A = H L^2$

For the two common cases of the perimeter being a square and an equilateral triangle, the values for **H** are 1/16 and $\sqrt{3}$ / 36, respectively. The unit of rad/sec for rotation is not commonly used. It is more usual to express rotation in units of deg/sec. Then using Eqs (2.7) and (2.10), the SF, which has units of Hz/deg/sec (or counts per degree), is

(2.11) $SF_{\text{Hz/deg/sec=counts/deg}} = \pi H L / (45\lambda)$

As an example, using a square cavity with perimeter of 20.00 cm, and for operation at a wavelength 0.6328×10^{-4} cm, the SF is 1379 counts/deg, or 2.610 arc-sec/count. It should be noted that it is usual to refer the SF to a nominal value, based on a nominal cavity dimension of a class of instrument. Due to the very high accuracy in the determination of SF in the laser gyro, it is typically expressed as a parts per million deviation from the nominal.

3.2.5 Ring Laser Gyro Design

There are numerous RLG configurations that are being used. In all designs, there are many common features. Figure 2.4 shows a simplified diagram of a conventional two-frequency RLG. It shows the laser cavity to be determined by the machining of a solid frame[21-22]. Holes are drilled to allow the containment of a gas, which is used to provide amplification for the laser transition. Gain for the atomic transition is provided by means of a dc discharge. Mirrors are mounted to the frame so that the two beams can be constrained to travel around the cavity. Some means is used to combine the two beams so that the beat frequency can be determined. Although not shown, there must be some means of mounting the RLG.

For all known devices, the gas is a mixture of helium (He) and neon (Ne). Although the infra-red transitions have been investigated, most production devices make use of the 0.6328 μm neon transition. Fill pressures are typically in the range of 5-10 Torr. The He-Ne mixture ratio is not designed for optimal power and is typically in the range of 5:1 to 20:1. To avoid problems of gain competition between the two beams, the neon is usually[23] a 50:50 mixture of the isotopes Ne^{20} and Ne^{22}. The wide range of the above mentioned parameters is due to design trade-off issues with respect to laser gain, gyro lifetime, gas discharge stability, laser intensity for detection sensitivity, single mode operation, thermal sensitivity and performance. Each RLG design requires a separate set of parameters and similar designs in different facilities typically have different design parameters.

The illustrative case shown in Figure 3, uses a square cavity configuration, which is similar to the first operating RLG[3]. In practice the geometry tends to vary. On the first major aircraft production program[8], the geometry was triangular. The operation of a five-sided RLG has been reported[24]. A trade-off study is typically made with respect to the geometry design. In all production cases, mirrors are used to turn the corners, although in early experiments at Honeywell, prisms were used. The use of prisms was quickly discontinued at Honeywell. The difficulty experienced was the lack of preservation of the differential path length stability to the previously discussed accuracy, when the beams traversed optical elements. The thought at that stage of development was difficulties with respect to magnetic sensitivity due to stress induced birefringence in the glass. Production projections at Honeywell also favored the use of mirrors over prisms. The mirrors are considered one of the most

critical components in the fabrication of the RLG, in that they directly affect the oscillation parameters of the two beams. In section 6 in this monograph, the use of total reflection prisms as an alternative to mirrors is discussed in detail.

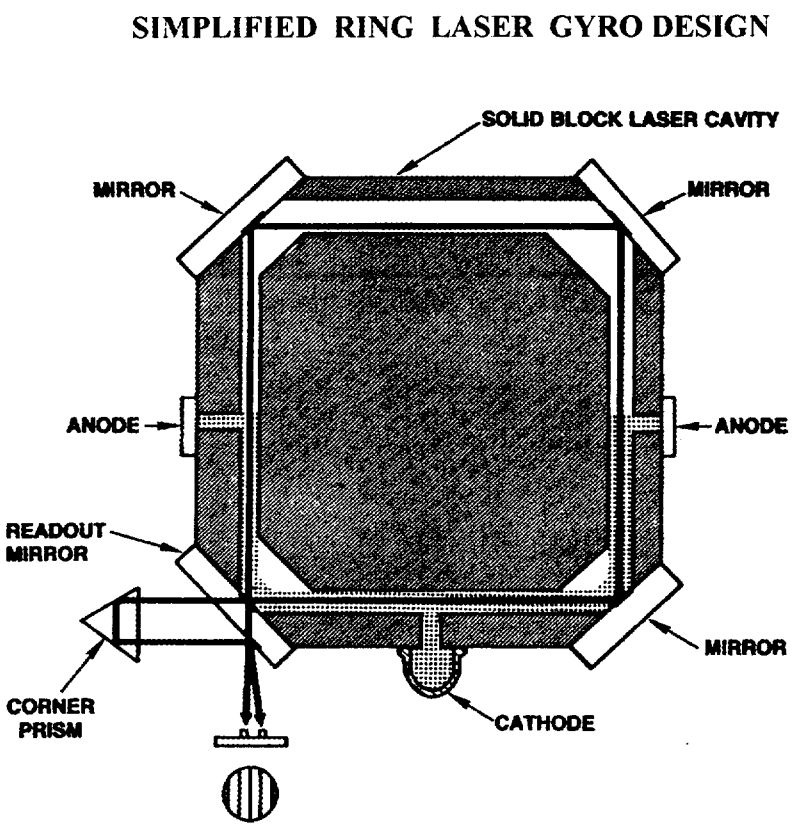


Figure 2.4. Schematic of a solid-block-construction RLG. In this example, a 4-sided geometry is shown, with a dc parallel discharge obtained with a common cathode and two anodes. The figure shows a typical readout in which the two oppositely directed output beams are made co-linear to form a fringe pattern.

The gas discharge always occurs over a portion of the total cavity. Figure 2.4, shows a parallel discharge generated by two anodes and one cathode. To obtain sufficient gain without operating with high discharge currents, the discharge typically passes around a corner for production sized devices. A high discharge current requires an increased input power requirement, which results in larger thermal gradients and turn-on transients that can cause performance degradation. It also requires a larger cathode to minimize gas cleanup, which shortens lifetime and causes performance variability. The use of two cathodes, with either one or two anodes can give rise to bias instability due to a combination of gas flow effects and the uneven burn-in effect from the two cathodes.

For larger sized gyros, all the electrodes can be placed on a single side of the cavity, such that the discharge does not turn a corner. This is desirable in that it tends to avoid interaction of the discharge with the reflecting surface of the mirrors and UV induced mirror degradation[25]. However, this tends not to be compatible with the gain requirements with the present trend to smaller sized gyros. It is critical that the discharge be of a parallel configuration because of gas flow effects[26], which must be balanced. This will be further discussed.

In the assembly of the gyro, alignment of the mirrors is critical. For stability considerations, it was found that it was preferable to fabricate the device with no provision for alignment of the mirror after attachment. The philosophy was that with no provision for alignment, there would also be no provision for misalignment. The alignment was mainly performed by

controlling tolerances in the fabrication of the frame. For cavity stability, one or two of the mirrors are curved. It is possible to compensate for misalignments in the faces of the frames by carefully positioning the center of curvature of the curved mirrors with respect to the frame. For a 4-mirror (even number of mirrors) RLG, requirements on the alignment are in the arc-sec range. For a 3-mirror (odd number of mirrors) RLG, requirements on the alignment out-of-plane are also in the arc-sec range. In the plane however, the requirements can be relaxed to the arc-min range. This is because it can be analytically shown that a closed path can always be found for this case[27] and hence there is inherent stability. However because of gas flow effects, there are stringent requirements on the location of the beam with respect to the bore. When performance is more important than production cost, in-plane tolerances are also maintained to the arc-sec range.

Once the mirrors are properly positioned, the conventional approach to sealing them onto the frame is the use of optical contact. By this it is meant that the two surfaces are "almost-perfectly" polished, cleaned and pressed together. Molecular attraction maintains a sufficient seal for all but the most severe environments. Adhesives tend not to be used because of the need for almost perfect isolation of the cavity from contaminants. With contamination, lifetime degradation will occur and the flow characteristics of the gas will change, affecting performance. It is possible to introduce virtual leaks, if because of improper cleaning and/or fabrication of the frame, contaminants from the frame, diffuse continuously into the cavity. Typically an activated "getter" is used to continuously interact and remove some of the more common contaminants.

For long lifetime and the maintaining of performance, it is required to maintain the gas mixture within a desirable range. As the laser is operated, a removal of the neon occurs as the discharge causes an imbedding of the neon atoms in the cathode. This gas cleanup will be the ultimate lifetime limitation of the gyro. To minimize cleanup, the dimensions of the cathode are carefully chosen, and correct processing to form an oxide coating on the cathode, is critical. The proper choice of discharge current and cathode material, is part of the tradeoff.

For performance considerations, it is necessary to use a material that maintains its shape over a temperature range, which is sometimes as wide as -55 °C to +85 °C. Fused silica materials, which tend to allow excellent polishing, and have a low thermal coefficient of expansion, have the property that allows the helium to diffuse from the cavity, and hence are unacceptable. Other low thermal coefficient of expansion ceramic-type materials having much lower diffusion properties for helium, were found to be acceptable. Cer-Vit was originally used, but was discarded due to it being no longer manufactured. Zerodur[28] is currently the material of choice.

Because of the cost of optical polishing, other approaches have been considered. A frame fabricated from a boro-silicate glass has been used[29]. It has a low diffusion coefficient for helium, is compatible with a frit seal for the mirrors, but suffers from the disadvantage of a relatively large thermal coefficient of expansion. Because of frame distortion, limited performance is obtained. Applications tend to be in the range of 1-10 deg/hr, as opposed to those in the range of 0.01 deg/hr. Another low-cost approach considered was the use of a fine-grind on the frame and an indium seal for the mirrors[30-31]. To maintain the necessary alignment and avoid the stringent tolerance on the thickness of the indium, the frame can be used for alignment determination. Long-term alignment stability is the issue with this approach.

Aluminum electrodes which tend to be chosen for cost considerations, are typically attached using an indium seal. This pressure weld has demonstrated success, although technique is critical with respect to maintaining a leak-free seal over the required temperature gradients. To avoid the cost of an additional seal, the anode typically doubles as the fill port. A tube built at the end of the anode is attached to a vacuum fill station. Once the gyro has been filled and taken through any preliminary burn-in process, fresh gas is added and the tube is pinched-sealed. Technique is important in obtaining a leak-free seal.

The basic approach to the readout of rotation in the RLG, is similar in all designs. Figure 2.5 illustrates the technique. To separate the rotation contribution from the high optical frequency, the two beams are first made approximately co-linear. Some means of retro-reflecting one of the beams onto the other is used. This results in a fringe pattern, with a spacing determined by the degree of co-linearity. When the two frequencies are equal, the fringes are stationary in time. When the frequencies are different, the fringes move across the field of view at the beat frequency rate. Thus by measuring the number of fringe moving across the field of view for a fixed time, rotation rate (using the SF to convert counts to angle) can be determined. By using each count in real time as a determination of the angle through which the sensor has been turned, the inherent integrating property of the RLG can be used. The direction in which the fringes move determines direction of rotation. To determine this direction, it is customary, to use two sensors positioned so that each observes a portion of the fringe in quadrature (an offset of 90° in phase) from the other. It should be noted that because of the larger relative accuracy in time measurement, reduced quantization can be obtained by measuring time for a fixed number of counts.

With this readout approach, the inherent SF of the RLG can also be reduced by obtaining a count of output for less than one full fringe of motion. Using 90° of phase motion, a readout of 4x enhancement is readily obtained. Factors up to approximately 1000 have been obtained[32]. Using a phase locked loop approach, a reduction factor of 128 was obtained, which reduced the quantization from 1.57 arc-sec to 0.012 arc-sec[20].

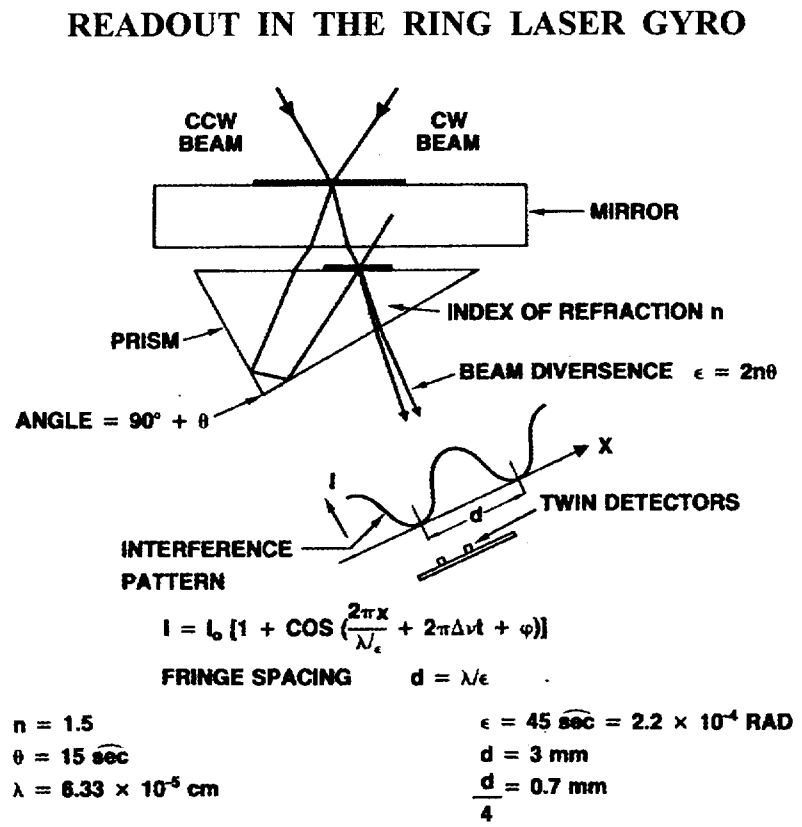


Figure 2.5. This figure shows an illustration of an early readout technique used in the RLG. A prism, mounted off the RLG is used to combine the two beams to form a fringe pattern. The pattern spacing is determined by the prism angle, as illustrated. A numerical example of the fringe spacing is shown. Because of the prism separation, which led to thermal effects causing bias instability, current designs have the prism (or corner cube for less sensitivity to misalignment) attached to the RLG.

3.3 FUNDAMENTAL LIMIT OF PERFORMANCE

3.3.1 Definition And Numerical Example

In describing the operation of the RLG, it is desirable to consider the ultimate operational capability of the instrument, commonly called " the fundamental limit". A number of papers have described the fundamental limit of the performance in the RLG[12,33-34]. A description is presented here that follows an intuitive classical derivation[35].

The laser beam is coherent in phase. Due to spontaneous emission, which causes incoherent photons with random phase, to be emitted into the direction of the two laser beams, there will be a degree of incoherence in the output beams. The fundamental limit occurs with the addition of one spontaneous emitted photon into the stimulated beam consisting of $\mathbf{N_p}$ photons. In the approach to be followed, the line width of each of the beam is determined. The line width of the beat frequency is the addition of the line widths of each of the two beams. The measurement accuracy is the variance of the random walk in phase. The diffusion constant is a measure of the beat frequency line width.

The phase uncertainty due to one spontaneous photon among $\mathbf{N_p}$ structured photons is

(3.1) $\delta\theta = 1 / N_p$

Since the photon has a lifetime in the cavity of $\mathbf{T_c}$, the phase uncertainty gives rise to a frequency uncertainty of

(3.2) $\delta\nu = \delta\theta \ / (2\pi\, T_c)$

The total energy dissipated from the cavity ($\mathbf{P_d}$), is

(3.3) $P_d = N_p\, h\nu \,/\, T_c$

where $\mathbf{h\nu}$ is the energy per photon, **h** is the Planck constant and **ν** is the light frequency. The cavity line width is defined from the photon lifetime as

(3.4) $\Delta\nu_c = 1 \,/\, (2\pi\, T_c\ )$

Using the above definitions, the laser line width can be found as

(3.5) $\delta\nu = 2\pi\, h\nu\, \Delta\nu_c^2\ (\ ½\, f\ ) \,/\, P_d$

In Eq (3.5), an ad hoc (½ **f**) factor has been added to give agreement with more rigorous derivations of laser line width[36]. The (½) factor is a field quantization effect. The f-factor is associated with the degeneracy **g** of the atomic transition and is given by

(3.6) $f = [\ 1 - (N_1 g_2\,) \,/\, (N_2 g_1\,)\ ] - 1$

where $\mathbf{N_i}$ are the number of atoms in the i'th laser transition level. For the low gain He-Ne transition, the levels are approximately equally populated. The degeneracy factors are $\mathbf{g_2} = 3$ and $\mathbf{g_1} = 5$, giving $\mathbf{f} = 2.5$. The variance of the laser line width growth in time is

(3.7) $\sigma^2 = 2\ (2\pi\ \delta\nu)\ \tau$

where the beat frequency line width is taken as twice the individual line widths, expressed in rad/sec.

Using Eq (3.5), the diffusion coefficient **W** (in units of rad/sec ^{½}) of Eq (3.7), is found as

(3.8) $W = \sigma \,/\, \sqrt{t}\ \ = 2\pi\ \ \Delta\nu_c\ \ (f\, h\nu \,/\, P_d)^{½}$

The cavity width can be express in terms of the measurable loss per pass Γ, the velocity of light **c**, and the cavity length **L**, as

(3.9) $\Delta\nu_c = c\, \Gamma \,/\, (2\pi\, L)$

Rather than express the fundamental limit in terms of the total dissipated power, which is difficult to measure, it is more convenient to use the power observed through an output mirror of known transmission loss **T**. The relation between the two is

(3.10) $P_o / P_d = T / \Gamma$

and using Eqs (3.9-10), Eq (3.8) gives

(3.11) $W = (c \,/\, L)\ [\ f\, h\nu\, \Gamma\, T \,/\, P_o\ ]^{½}$

It is customary to express the diffusion coefficient in inertial units. To do so, the RLG SF from Eq (2.11) can be used to give

(3.12) $SF_{\ rad/deg} = SF_{\ count/deg} \times 2\pi_{\ rad/count} = 2\pi^2\ \ H\, L \,/\, (45\lambda)$

Thus **W** must be divided by Eq (3.12) and also multiplied by 60 to convert $\sqrt{sec}$ to $\sqrt{hr}$, giving the diffusion coefficient in deg $/\sqrt{hr}$ as.

(3.13) $W\ = 1350\ [\ c\, \lambda \,/\, (H\, \pi^2\, L^2)\ ]\ [\ f\, h\nu\, \Gamma\, T \,/\, P_o\ ]^{½}$

The use of Eq (3.13) can be illustrated with the following numerical example. Consider a RLG with a square cavity of length 20 cm, operating at the He-Ne 0.633 μm transition. Also consider a cavity design such that the total cavity loss is 360 ppm and with the laser having an output power of 30 μ W through an output coupler having a transmission of 60 parts per million (ppm). Then the parameters of Eq (3.13), with their units are

$L = 20$ cm, $c = 3 \times 10^{10}$ cm/sec, $\lambda = 0.633 \times 10^{-4}$ cm, $H = 1/16$, $h = 6.63 \times 10^{-27}$ erg-sec,
$\nu = 4.74 \times 10^{14}$ /sec, $\Gamma = 360 \times 10^{-6}$, $T = 60 \times 10^{-6}$, $P_o = 300$ erg/sec (30 μ W), $f = 2.5$.

Substituting the above parameters into Eq (3.13), the fundamental noise limit, usually called angular random walk due to the form of Eq (3.7), is found as 247 μdeg/$\sqrt{hr}$. To avoid the complexity of dealing with the units of the various parameters that determine the angular random walk coefficient, it is simpler to use the expression with scaled parameters. Using the above numerical example, Eq (3.13) can be put in the form

$$(3.14) \quad W = (15.4 / H)\ (20 / L)^2\ (\Gamma / 360)^{1/2}\ (T / 60)^{1/2}\ (30 / P_o)^{1/2}$$

where the units are: **L** in cm; **T**, Γ in ppm,; $\mathbf{P_o}$ in μW; **W** in μdeg/$\sqrt{hr}$. For the two cavity configurations that are predominately in usage, **H** = 1/16 for the square and $\sqrt{3}$ /36 for the triangle.

3.3.2 Comparison Of Square and Triangular Geometry

From the early days of the development of the RLG, an issue has been the relative merits of using the square vs the triangular geometry. Hence it is of interest to compare the fundamental limit for both the square and triangular geometry. The form of Eq (3.14) cannot be directly used, in that there are different design parameters used for the two geometries. Thus further scaling is needed for these parameters. As will be discussed in more detail in Section 4, the measured power (in μw) can be written in terms of the output mirror transmission **T** (in ppm), the laser gain **G** (in ppm), the cavity loss Γ (in ppm), and a design parameter $\mathbf{P_a}$ (in μw), as

$$(3.15) \quad P_o = T\, P_a (G / \Gamma - 1)$$

The $\mathbf{P_a}$ parameter will depend on the geometry and gas fill, and experience indicates that it can be approximated for all classes of instruments as 0.1 μw. Since the fundamental limit, as given by Eq (3.14), scales with the square root of power, small differences in P_a will not cause a large discrepancy.

The output coupler is also a design parameter, but as can be seen from Eq (3.14), the explicit dependence cancels. Although the gain/loss ratio will be different for the various designs, reasonable estimates can be made. A RLG with a triangular geometry can be estimated as having gain/loss ratios of 2-3. Because of the odd number of mirrors, a larger value of gain tends to result in multi-mode oscillation in the region midway between the lowest order longitudinal mode. A RLG with a square configuration can be designed with a much higher gain/loss ratio, because the even number of mirrors result in all higher order transverse modes being close to the lowest order longitudinal mode. These higher order transverse modes tend not to occur because of gain saturation due to the presence of the lowest order mode. Gain/loss values of 10 are not uncommon. For scaling purposes, a gain/loss value of 3 and 8, will be used for the triangular and square cavity, respectively. Mirror technology has improved over the years and although mirrors from each manufacturer have different characteristics, a common loss value will be used for all classes of instruments.

The basic difference between the square and triangular gyro, is the number of mirrors. It will be assumed that the triangular RLG will have only one of the three mirrors as an output coupler, which has an added transmission loss. The square will be assumed to have two of its four mirrors as output couplers. Note that with these assumptions, the reduced mirrors loss factor (1.22) of the triangle almost exactly cancels the geometry advantage (1.30) of the square. An advantage factor of 1.87 for the square geometry remains because of the higher power capability. It should be noted that the mission application and resultant packaging, will also play a role in the geometry choice. Efficient packaging can allow the use of a larger sized instrument, with a lower angular random walk.

Figure 3.1 shows a plot comparing the fundamental limit for the square and the triangle, using the above described assumptions. The mirror and transmission losses used for the tradeoff are shown in the figure. The square configuration shows a factor of two reduction for comparable cavity length. For larger sized instruments, the absolute difference between the two is small, although these sized instruments are not in favor in today's market. With smaller sized instruments, the difference between the two curves becomes very pronounced with a large advantage going to the square geometry. Much more detail than was presented here would go into an actual trade study, including packaging issues.

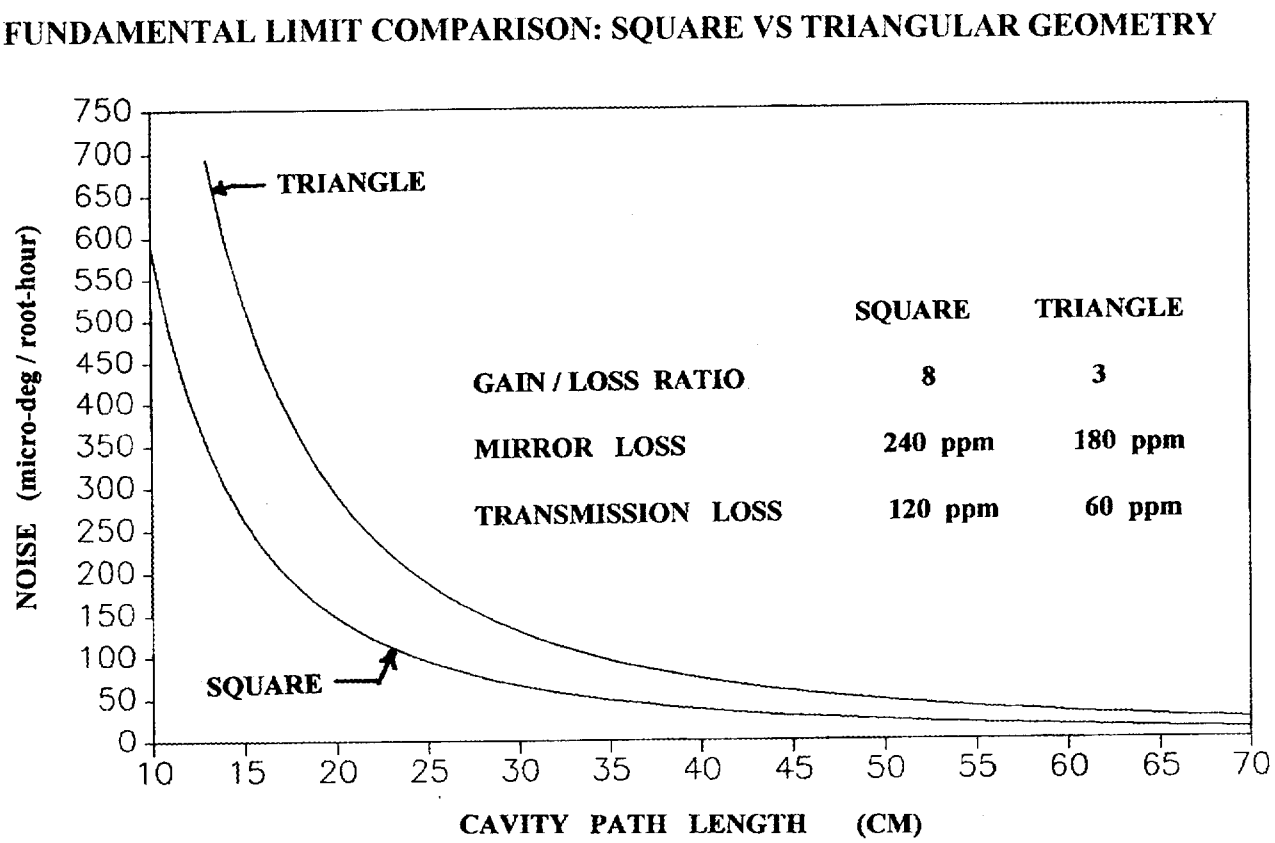


Figure 3.1. This figure shows a comparison of the fundamental limit (angular random walk or "noise") of performance for the square and triangular geometries. The upper and lower curves represent the triangular and square geometries, respectively. The design parameters used in the comparison are shown in the figure.

3.4 RING LASER GYRO THEORECTICAL MODEL

3.4.1 The Semi-Classical Lamb Formalism

A theoretical model of the RLG was developed[23,37] soon after the demonstration of the device, and well before production usage. As the RLG went into production, the basic formalism of the model remained unchanged. Refinements were made to account for observed phenomena and to better match the model with current designs. The mathematics used in the derivation of the model are complex and tedious and in that they have been presented in previous publications[11,13], they will not be repeated here. A summary of the various treatments and the assumptions used, will be presented.

The model that is typically found in the literature, makes use of the Lamb formalism[38] that was developed for the linear laser and generalized[23,37] for the ring laser. The technique is considered semi-classical and uses a self consistent approach. It assumes an electromagnetic field to exist in the cavity. The electric component of the field polarizes the electric dipoles of the individual atoms, which results in a macroscopic polarization. This polarization is then used as a source term in Maxwell's equations, which determines a reaction electric field. For self consistency, the resultant electric field must equal the original assumed field. The self consistency gives rise to a set of equations which describes the amplitudes and frequencies of the oppositely directed beams. The formalism is considered semi-classical in that Maxwell's equations are used for the generation of the electric field and quantum theory is used for the determination of the polarization.

To account for coupling between the two beams, numerous approaches have been taken[39], which have led to identical terms in the amplitude-frequency equations. In an early approach[40] a fictitious current density was introduced, which resulted in an additional source term in Maxwell's equations. This then modified the set of amplitude-intensity equations with additional terms

that provided a qualitative, although incomplete description of the effects of coupling. A more general approach is covered in the Lock-in Section of this review.

In Lamb's original formalism[39] for the linear laser, the electric field in the cavity consisted of two oppositely directed traveling waves that summed to a standing wave. In the ring laser, the two traveling waves are independent in that they can have different frequencies and amplitudes. This results in a set of four equations, describing the amplitudes and frequencies of each of the oppositely directed traveling waves. For more complex RLG configurations, the polarization of the two beams at a given location, need not be the same[41]. For these cases the number of equations describing the beams increases to eight, with a considerable increase in complexity. To address the case of magnetic field effects, the general polarization treatment must be used. To avoid this increased complexity in this limited review, only solutions of a single polarization will be considered.

The amplitude and frequency equations have been shown to describe the observed characteristics of the ring laser[37]. Characteristics described by the amplitude equations include: conditions on threshold operation; output intensity as a function of oscillation frequency tuning, including the "Lamb dip" phenomena; competition effects between the oppositely directed beams as a function of the relative concentration of the neon isotopes and the oscillation frequency; modulation of the intensity of each of the beams at the beat frequency rate; a dc intensity competition that occurs at small rotation rates; jump in intensity with frequency tuning as mode goes above threshold and no jump with frequency tuning in the reverse direction. Characteristics described by the frequency equations include: a frequency splitting between the two beams when the ring laser is operated on a rotating frame; coupling terms leading to frequency synchronization at low rotation rates; distorted wave forms at rotation rates slightly above the lock-in threshold; a scale factor non-linearity varying inversely with the square of the rotation rate; gas dispersion effects leading to frequency shifts that vary with oscillation frequency; frequency shifts caused by motion of the gas atoms that contribute to the laser gain.

3.4.2 RLG Amplitude And Frequency Equations

The starting point in the development of the RLG model, is the writing of Maxwell's equations in one dimension on a rotating frame. The term arising from the rotation gives rise to an asymmetry in the index of refraction for light traveling with and against the direction of rotation. The wave equation is of the form[23,42]

$$(4.1) \qquad -\frac{1}{\varepsilon_o\mu_o}\frac{\partial^2 E}{\partial z^2} + \frac{\omega}{Q}\frac{\partial E}{\partial t} + \frac{\partial^2 E}{\partial t^2} + \frac{a}{\varepsilon_o\mu_o}\frac{\partial^2 E}{\partial z \partial t} = \frac{\omega^2}{\varepsilon_o}P - \frac{\sigma_s}{\varepsilon_o}\frac{\partial E_s}{\partial t}$$

where **E** is the electric field intensity, **P** is the polarization, **Q** is the passive quality factor of the cavity, the term in **a** is proportional to the rotation, $\mathbf{E_s}$ represents the back scattering field, and σ_s is a fictitious conductivity. The complexity in the use of Eq (4.1) to determine the amplitude-frequency equations of the RLG, is the quantum mechanical determination of the polarization. In the Lamb appoach[38], a perturbation treatment was used, where the density matrix was expanded in powers of the interaction between the radiation field and the atomic system.

In the amplitude and frequency equations, the first order terms determined the threshold conditions of oscillation and the unsaturated dispersion effect of the index of refraction (frequency pulling). Third order terms were necessary to obtain the effects of gain saturation on both the amplitudes of the electric field and the dispersion effects (frequency pushing). The second order terms gave a description of the saturated population inversion.

The amplitude-frequency equations for the ring configuration with coupling, were shown to be of the form[40]

$$(4.2a) \qquad (2L/c)\,(dE_1/dt) = E_1\,[\,\alpha_1 - \beta_1 I_1 - \theta_{12} I_2\,] - 2r_2 E_2 \cos(\psi + \varepsilon_2)$$

$$(4.2b) \qquad (2L/c)\,(dE_2/dt) = E_2\,[\,\alpha_2 - \beta_2 I_2 - \theta_{21} I_1\,] - 2r_1 E_1 \cos(\psi - \varepsilon_1)$$

$$(4.3a) \qquad \omega_1 + (d\phi_1/dt) = \Omega_1 + \sigma_1 + \rho_1 I_1 + \tau_{12} I_2 - (c/L)\, r_2 (E_2/E_1) \sin(\psi + \varepsilon_2)$$

$$(4.3b) \qquad \omega_2 + (d\phi_2/dt) = \Omega_2 + \sigma_2 + \rho_2 I_2 + \tau_{21} I_1 + (c/L)\, r_1 (E_1/E_2) \sin(\psi - \varepsilon_1)$$

where (1,2) represent the two oppositely directed beams; $\mathbf{I_i}$ (i=1,2) the dimension-less laser intensity (square of $\mathbf{E_i}$); ω_i the oscillation frequency; ϕ_i an arbitrary phase angle of each beam; Ω_i the natural frequency of the cavity (which changes linearly with rotation rate); ψ the instantaneous phase difference between frequency 2 and 1; $\mathbf{r_i}$ the scattering coupling of each of the beams into the other; ε_i scattering phase angles; α_i, β_i, θ_{ij}, σ_i, ρ_i, τ_{ij}, the Lamb coefficients that describe the atomic

transition[43]. Assuming a He-Ne RLG, these coefficients are functions of the Ne_{20}/Ne_{22} isotope ratio, the oscillation frequencies of each of the traveling waves and the total gas pressure.

The above formalism was developed for lasers of the He-Ne class, where the atomic structure is predominately Doppler broadened. This led to a Gaussian line-shape. A Lorentzian contribution to line-shape broadening occurred due to a combination of natural broadening , hard collisions (excited state lifetime reduction due to collisions), and soft collisions (velocity shift due to collision that does not result in state change) was included[44]. The combined Gaussian-Lorentzian line-shape resulted in complicated expressions for the Lamb coefficients.

It is customary to replace the individual frequency expressions with one that describes the frequency difference, which is the output of the RLG. Then using the measurable laser intensities, instead of the field amplitudes, and writing the frequency difference as the time derivative of the instantaneous phase difference between the two beams, Eqs (4.3a) and (4.3b) become

$$(4.4)\quad (d\psi/dt) = (\Omega_2 - \Omega_1) + (\sigma_2 - \sigma_1) + (\tau_{21} I_1 - \tau_{12} I_2) + (c/L)\,[\, r_1 (I_1/I_2)^{1/2} \sin(\psi - \varepsilon_1) - r_2 (I_2/I_1)^{1/2} \sin(\psi + \varepsilon_2)\,]$$

The units of Eq (4.4) are those of laser frequency. As shown from Eq (2.7), the frequency can be related to rotation rate through a scale factor.

Correspondingly, the two amplitude equations (4.2a) and (4.2b) can be written in terms of the laser intensities as

$$(4.5a)\quad (L/c)\,(dI_1/dt) = I_1\,[\, \alpha_1 - \beta_1 I_1 - \theta_{12} I_2\,] - 2r_2 (I_1 I_2)^{1/2} \cos(\psi + \varepsilon_2)$$

$$(4.5b)\quad (L/c)\,(dI_2/dt) = I_2\,[\, \alpha_2 - \beta_2 I_2 - \theta_{21} I_1\,] - 2r_1 (I_1 I_2)^{1/2} \cos(\psi - \varepsilon_1)$$

In Eqs (4.5), the α_i term represents the gain-minus-loss, or the threshold condition (α=0) for each beam. When the RLG is rotated, the individual cavity frequencies are different and the coefficients are then evaluated at these slightly different frequencies. In most cases there is little change in the value of the coefficients. A physical effect only shows when a difference between the coefficients is obtained. Then, as shown by the first order term in a Taylor series expansion, the difference results in a term proportional to the product of the derivative of the coefficient and the rotation rate[45].

3.4.3 High Gain Intensity Equation

The β_i and θ_{ij} terms in Eqs (4.5) tend to saturate the gain (reduce α). Neglecting the scattering (terms in **r**), it can be seen from Eq (4.5), that (assuming symmetry) the intensity steady state condition for each beam is given by

$$(4.6)\quad I = \alpha / (\beta + \theta)$$

In the determination of the Lamb coefficients, the saturation terms β, θ were found to be proportional to the transition gain **G**. Since the α term is excess gain, it is of the form **G**-Γ, where Γ is the total cavity loss. Then Eq (4.6) is of the form (1- Γ/**G**), showing an asymptote for a large value of gain. For typical gas fill parameters, the laser intensity shows, in contradiction to this asymptote, a linear dependence on the discharge excitation current. This linear dependence holds for values of the gain being much larger than the cavity loss. This discrepancy raises an issue with respect to the general usage of the intensity equations as given by Eqs (4.5). They are only valid for usage when the transition gain is close to the cavity loss, i.e. **G**/Γ ratios of less than 1.2 (approximately). Since the **G**/Γ ratio for most practical gyros are on the order of 2 or greater, the intensity equations given by Eqs (4.5) do not adequately describe the operation of the gyro.

The problem is due to the use of third order perturbation theory, which is most valid close to threshold. Some derivations have included fifth order terms[46], but because of slow convergence, are also inadequate for usage in a practical RLG. High signal theories have been developed[47], which are valid for general usage. Unfortunately, the analysis is of such complexity, that closed form solutions are not obtainable. Computer analysis of the solutions are required and physical insight is difficult.

A viable approach, is the usage of the low-gain intensity equations for understanding and guidance, and for modeling purposes, the intensity is approximated by an expression of the form

$$(4.7)\quad P = P_c\,(G / \Gamma - 1)$$

The loss can be determined by measurement of the cavity quality factor **Q** and the gain. The $\mathbf{P_c}$ parameter can be determined from empirical data. Note that Eq (4.7) increases with linearly with gain at high ratios of **G**/Γ, and reduces to Eqs (4.5) near

threshold, when $\mathbf{G} \approx \Gamma$. The expression for α and its functional dependence, as derived using the low gain model, is still valid, since it is independent of saturation effects.

3.4.4 Frequency Coupling In The Intensity Equation

Since the third order theory is descriptive for low gain, the coupling terms (which tend to be small compared to α) provide an accurate description of experimental data. As can be seen from Eq (4.4), when gas dispersion and coupling effects are neglected, the integration of the phase difference between the two beams can be performed, to show the phase increasing linearly with time. The time derivative of the phase is the beat frequency of the RLG and is proportional to the rotation rate ($\Omega_2 - \Omega_1$). Hence, the coupling terms in Eqs (4.5) show an intensity modulation at the beat frequency rate.

For the case when the RLG is frequency synchronized, the phase difference between the beams is constant in time, and the coupling terms introduce a small change in excess gain. This is most dramatically shown[37] for the case of small excess gain and when the cavity mode is set between operating modes so that no laser oscillation occurs. Then as the cavity frequency is moved towards the Doppler center frequency, the laser threshold will be reached. At this frequency, there will be a jump in laser intensity. If the frequency tuning is reversed, the intensity will drop more gradually, showing a hysteresis pattern. The explanation is thought to be that at threshold, the scattered energy is phased so that it reduces the cavity loss and changes the threshold condition. Thus the tuning-into-threshold condition requires a higher gain to initiate laser oscillation.

When the laser is rotated, but still locked, Eq (4.4) shows the phase difference to vary with input rate. This can be more clearly seen using the sum and difference formalism, as presented in section 3.5, where lock-in is described This causes the effective scattering-gain (or loss correction) to act in opposition for the two beams and cause one to increase and the other to decrease from the no-coupling value. Thus as the laser is rotated, the intensity difference increases with increasing rotation rate, up to the lock-in threshold, where there is an abrupt return to approximately the zero-rate values. The intensity difference reverses sign with a reversal in the rotation direction.

3.4.5 Beat Frequency Equation

The frequency difference, as given by Eq (4.4) is basic to the operation of the RLG. The first term shows the beat frequency to be proportional to the rotation rate, and is considered the ideal operation. A measure of "goodness" of the device is determined by the closeness of the operation to ideal. The last term in Eq (4.4) causes a coupling between the two beams and leads to frequency synchronization. This effect will be discussed more fully in section 3.5, where lock-in is described. The middle terms give rise to bias (output in the absence of an input rate) and scale factor corrections.

3.4.6 Scale Factor Error

The scale factor error can be most readily understood using the second term in Eq (4.4). As discussed with respect to the amplitude equations, the coefficients are evaluated at the individual frequencies of the two beams. The difference, as shown by a Taylor expansion, gives rise to a term proportional to the derivative of the coefficient and the rotation rate, which by definition is a scale factor contribution[45]. In that the dispersion coefficient σ is odd in frequency, the derivative is even, and has (approximately) a maximum at the frequency at which the path length control operates. The scale factor correction due the dispersion can vary on the order of tens of ppm for the low gain He-Ne 0.633 μm transition to thousands of ppm for the higher gain He-Ne 1.15 μm transition. This variation is one of the major reasons for the requirement of path length control. The lack of perfect coincidence between the maximum of the differential dispersion and maximum gain (which determines the discriminant for path length control), is due to pressure broadening[48] induced line shape asymmetry. Except for applications requiring extremely high scale factor stability over long periods of operation and/or dormancy, this deviation can be ignored.

The scale factor correction contribution from the third term in Eq (4.4) is only valid close to threshold, based on the use of the previously discussed third order perturbation approach. In keeping with the use of Eq (4.7) for the empirical expression for the laser intensity, a corresponding expression for the scale factor correction can be written as

$$(4.8) \qquad (\Delta S/S) = (\Delta S/S)_N - a\,G / (1 + x\,P_o) + NL(\Omega)$$

where $\Delta S/S$ is the total scale factor correction from an arbitrary nominal S_N, expressed in ppm; $(\Delta S/S)_N$ is a constant scale factor correction (independent of dispersion and scattering) from the same arbitrary nominal, in ppm; the product $\mathbf{aG}$ is the first order differential dispersion contribution; $\mathbf{G}$ is the laser gain; the product $\mathbf{xP_o}$ is the intensity saturation contribution; $\mathbf{P_o}$ is the measured laser intensity through an output mirror; the quantity $\mathbf{NL(\Omega)}$ is a nonlinear contribution arising from the coupling. The nonlinear term will be considered more fully in section 3.5, where lock-in is described. For most RLG designs, a plot of scale factor correction vs. discharge current, shows a linear relation, which tended to validate the low gain model. However, experience shows that Eq (4.8) is also reasonably linear over the discharge current operating range of most ring lasers, and it is difficult to separate the two models based on discharge current linearity.

Once the scale factor correction is known, the actual scale factor of the RLG is determined from the expression

(4.9) $SF = S_N [1 + (\Delta S/S) \times 10^{-6}]$

A problem that sometimes arises in the use of Eq (4.9) is that component engineers tend to express scale factor (**S**) in the measurable unit of counts/revolution, or counts/arc-sec, while system engineers tend to express the scale factor ($\mathbf{S_k}$) in analytic units of arc-sec/count. When the scale factor is expressed in ppm deviation from a nominal, the ppm values are not simply the reciprocal of one another. They are related by

(4.10) $S_N [1 + (\Delta S/S)] \times S_{kN} [1 + (\Delta S_k/S_k)] = 1$

where $(\Delta \mathbf{S_k}/\mathbf{S_k})$ is the deviation of the scale factor as expressed in arc-sec/ct from the nominal $\mathbf{S_{kN}}$. When the scale factor deviation from nominal is on the order of 2000 ppm, which could be a typical value, the discrepancy between the two approaches is 4 ppm. This discrepancy is considerable for most applications.

The dispersion correction of Eq (4.8) can be estimated, based on typical values of the design parameters. It must be noted that each instrument will have different parameters. Using units of ppm, a 0.633 μm He-Ne RLG with a square cavity design, with the laser intensity given by Eq (4.7), can have parameters of **a** = 0.09 ppm/ppm (at 6.5 torr with an equal mixture of Ne_{20}-Ne_{22}), **G** = 2000 ppm, **x** = 0.08 /μw, Γ = 400 ppm, $\mathbf{P_c}$ = 7 μw, $\mathbf{P_o}$ = 28 μw. With parameters of these values, the dispersion part of the scale factor correction is 56 ppm.

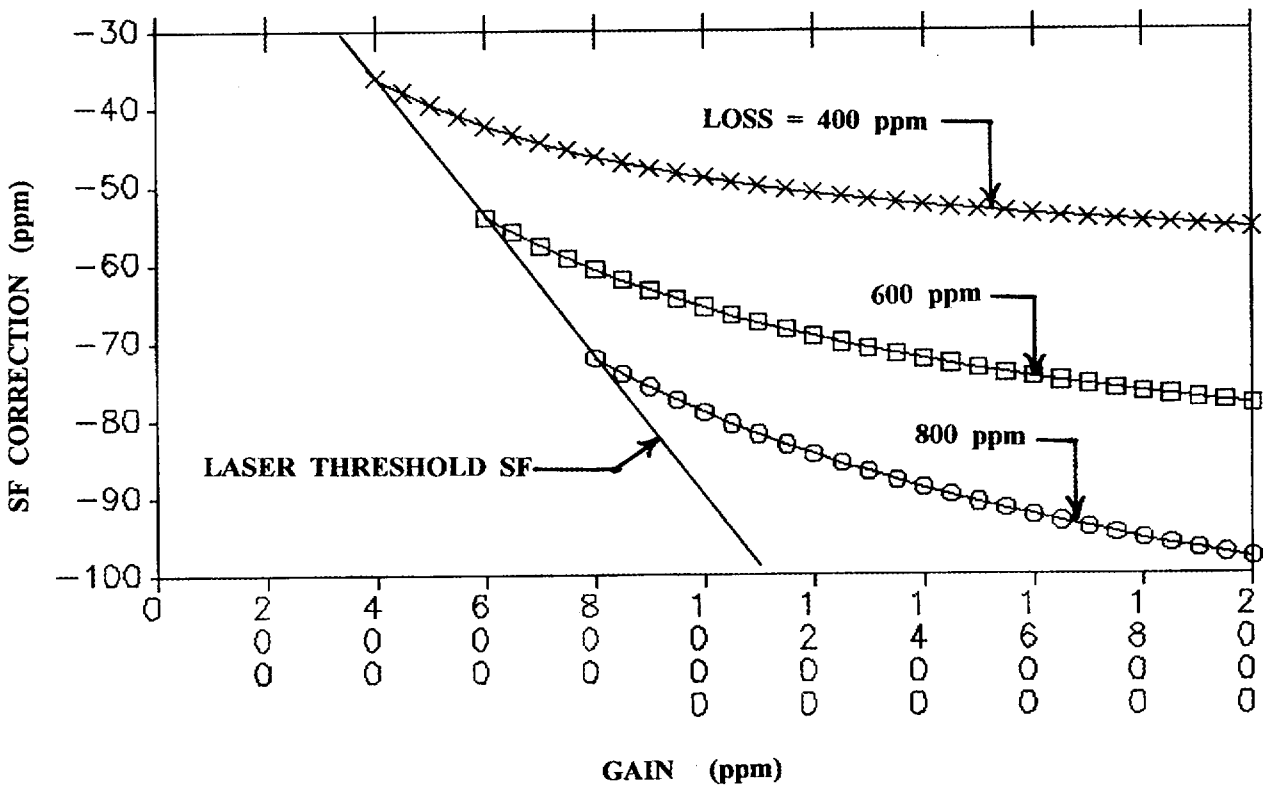


Figure 4.1. This shows the SF correction from nominal due to dispersion of the gas medium. The "laser threshold SF" shows the SF correction at threshold (gain = loss), for increasing values of gain. Once threshold is reached, then for increasing gain at fixed values of loss, the SF correction changes differently from how it changes for the threshold case. The SF sensitivity to gain is shown for three different values of loss. The SF sensitivity to loss at fixed values of gain, is obtained by determining the SF change from the above curves at vertical gain lines. Parameters used for these plots are: $\mathbf{P_c}$ = 7 μw, **a** = 0.09, **x** = 0.08 / μw

3.4.7 Scale Factor Sensitivity To Gain And Loss

Figure 4.1 shows a plot of the dispersion contribution to the scale factor correction, as given by Eq (4.8), plotted against the laser gain for various total cavity loss. I t shows the correction to increase negatively with increasing value of gain. The solid line represents the threshold value of the scale factor correction. Once laser oscillation occurs, laser intensity saturation reduces the slope with gain (discharge current). What is of interest is the stability of the scale factor correction, as opposed to the magnitude. To determine stability, the intensity equation, as given by Eq (4.7) can be substituted into the dispersion part of the scale factor correction. Then the partial derivatives for the cases of constant loss and gain, can be obtained as

$$(4.11) \qquad \left.\frac{\partial S}{\partial G}\right|_{loss} = -a\,(1 - x\,P_c) / (1 + x\,P_o)^2$$

$$(4.12) \qquad \left.\frac{\partial S}{\partial \Gamma}\right|_{gain} = -a\,x\,P_c\,(G / \Gamma)^2 / (1 + x\,P_o)^2$$

The sensitivity equations plus the laser intensity equation allow a trade to determine the optimum design parameters for maximum scale factor stability. Note that all three equations depend only on the gain/loss ratio, as opposed to the individual values of the gain and loss. They show that for the (**a**, **x**, $\mathbf{P_c}$) values as given above, and for gain/loss ratios such that operation is well above threshold, the loss sensitivity in on the order of -0.1 ppm/ppm. The corresponding gain sensitivity is on the order of -0.01 ppm/ppm, an order of magnitude smaller. For operation close to threshold, the two sensitivities tend towards the value of -.05 ppm/ppm. The RLG is usually not operated close to threshold in that at that region, insufficient output power is obtained for the usual low-cost electronics design. Thus the higher loss sensitivity is more typical and hence mirror loss stability is especially critical for applications requiring a high degree of scale factor stability.

3.4.8 Experimental Determination Of Gain And Loss

As has been demonstrated, the non-linearity contribution to the scale factor correction, as given by Eq (4.8), can be readily separated by determining the high rate asymptotic value. Then using the resultant expression and Eq (4.8) for the laser intensity, it is simple to invert the two equations to solve for the laser gain and cavity loss as

$$(4.13) \qquad G = [\,(\Delta S/S)_N - (\Delta S/S)\,]\,[\,1 + x\,P_o\,] / a$$

$$(4.14) \qquad \Gamma = [\,(\Delta S/S)_N - (\Delta S/S)\,]\,[\,1 + x\,P_o\,] / [\,a\,(P_o / P_c + 1)\,]$$

Once the gyro has been calibrated for the determination of the design parameters, a measurement of the non-linearity corrected scale factor and the laser intensity, allows the calculation of gain and loss. The ability to quantifiably determine the gain and loss is highly significant. A scale factor and laser intensity measurement could be used to determine the predominant cause of any observed change. As can be seen from Eq (4.7) and Figure 4.1, if the scale factor and laser intensity change in the same direction (increase or decrease), the effect is loss driven. If they change in opposite directions, the effect is gain driven.

The calibration of the parameters of Eqs (4.13) and (4.14) is non-trivial, but an approach has been demonstrated. The $\mathbf{P_c}$ coefficient is a basic design parameter and many techniques are available for its determination. Techniques to determine cavity loss have been reported[49]. Simultaneously, the laser intensity $\mathbf{P_o}$ is determined. Because of the probability of loss changes, for maximum accuracy, a scale factor calibration should be performed as soon as possible. The calibration allows the removal of the non-linearity term. The calibration is repeated for varying values of discharge current. From this data, the slope of the nonlinear scale factor vs. laser intensity ($\mathbf{S_p}$) can be determined. By working with the derivative, the unknown nominal scale factor can be eliminated from Eq (4.8), and the "x" parameter can be determined as

$$(4.15) \qquad x = \{[\,(K_{SF}\,P_c)^2 + 2K\,P_{av}\,(P_c + P_{av})\,]^{1/2} - (P_{av} + K\,P_c)\} / P_{av}^2$$

where $\mathbf{P_{av}}$ is the average intensity over the discharge current range and the quantity $\mathbf{K_{SF}}$ is defined as

$$(4.16) \qquad K_{SF} = -a\,\Gamma / (2S_p\,P_c)$$

As previously discussed **a** can be determined theoretically and it has a small pressure dependence. Once **x** is determined, Eq (4.8) can be used to determine the geometric scale factor correction $\mathbf{(\Delta S/S)_N}$. The rationale for the technique of using current variation, is an attempt to vary the laser gain without a corresponding loss variation. A temperature change will vary both. It should be noted that although inaccuracies in the determination of the design parameters will affect the absolute values of the gain and loss, relative values can still be determined to a high degree of accuracy.

As an example of the ability to determine gain and loss, Figure 4.2 shows a plot of the turn-on characteristics of a 20 cm square perimeter RLG. The data is described in the caption under the figure. This data shows generic characteristics of the RLG. A temperature turn-on transient always occurs. The magnitude is strongly dependent on the gyro design. Typically, the laser intensity and SF also show a characteristic turn-on transient. Then using the approach described in subsection 3.4.8, the cavity gain and loss can be calculated during the turn-on period. Relative loss accuracy is less than 0.1 ppm, which because of the capability of the RLG to obtain very stable scale factor and intensity data, is probably the most accurate obtained to date. There have been no other reported experiments in which gain has been measured to an accuracy of under 1 ppm. It can be observed that the gain exhibits a quick transient with a time constant less than that of the temperature and intensity. This is possibly due to the establishment of the discharge characteristics. Of more interest is the continuous growth of the cavity loss. This data illustrates a common problem with the RLG, which is the tendency of mirror loss growth[50], with operation of the laser. The effect is thought to be associated with the bombardment of the mirror coating with UV generated by the gas discharge.

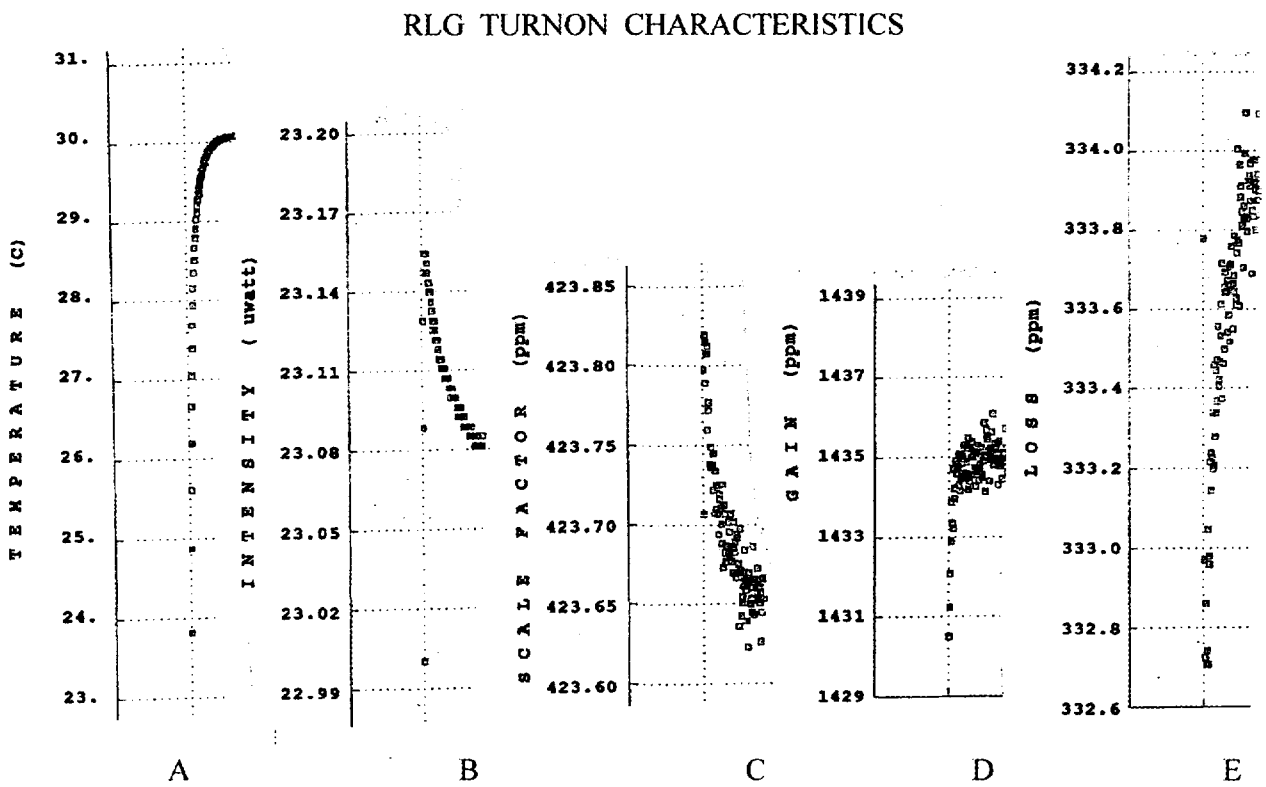


Figure 4.2. These plots show the determination of the RLG cavity gain and loss from a cold-start measurement of laser intensity and SF for a 20 cm square RLG. SF data was obtained from turn-on over a sector or 10 revolutions, for 11 rates ranging from 600 to 700 deg/sec with steps of 10 deg/sec. For each rate, data was obtained for rotations of both the clockwise and counter clockwise directions. Each data point shown in the figure encompassed approximately two minutes and the entire run encompassed four hours. Plot A shows the measured temperature. Plot B shows the measured laser intensity. Plot C shows the asymptotic value of SF (deviation from an arbitrary nominal), determined by removing the rate non-linearity for each sequence of 22 data points. Plot D shows the calculated of laser cavity gain. Plot C shows the calculated value of laser cavity loss.

3.4.9 Gas Flow Induce Bias

Gas flow bias[26] is one of the critical error mechanisms in the RLG. It tends to get very little attention in the public domain, as opposed to the more dramatic lock-in effect. The RLG manufacturers have recognized the criticality of this error mechanism. As a result, they have accumulated a very large reservoir of data, which has remained proprietary and unpublished.

Because of the complexity of gas discharge mechanisms and the very high sensitivity of the RLG to small physical phenomena, there is still, after 35 years, limited understanding of the gas flow drivers. The usual approach is empirical,

which has resulted in the large amount of data. Fortunately, the mechanisms are highly stable and repeatable, so that a systematic design can be obtained that is producible. For these reasons, a somewhat detailed description will be given of the generic bias mechanism, with a more vague description of the gas discharge mechanism.

In the description of the scale factor correction arising from the dispersion of the atomic transition, it was shown that the frequency of each beam is pulled by an amount proportional to the distance between it and the peak gain frequency. When the laser is rotated, a splitting between the two frequencies occurs. The rotational splitting is corrected by the frequency pulling. Since the dispersion is odd in frequency space, the correction is proportional to the product of the rotation rate and derivative of the dispersion coefficient. In that the correction is proportional to rotation rate, it results in a scale factor correction.

In an equivalent analysis, it was shown[26] that when the gas that provides laser gain, has a non-zero velocity flow with respect to the path along which the laser beams travel, a zero rate offset (bias) occurs. This non-zero flow can occur due to thermal gradient effects, a component flow due to the use of a dc discharge to provide excitation, or an intentional flow of the gas. The analysis to obtain the velocity bias is identical, irrespective of the source of the flow.

The bias is obtained from the same term that results in the above described scale factor correction. For the case of gas flow, the dispersion curve is shifted in opposite directions with respect to the oscillating frequency of each beam. The frequency pulling is similar to what occurred with rotational splitting, but now the rotation term is replaced with the gas flow velocity. Considering only the first order dispersion correction, the beat frequency (Ω_b), obtained from Eq (4.4), was shown to be[51]

$$(4.17) \qquad \Omega_b = \Omega (1 - A) - 2K V A$$

where **K** is the wave number ($2\pi/\lambda$), **V** the velocity flow around the cavity perimeter, and A represents the derivative of the σ-dispersion curve and is proportional to laser gain. A is a positive quantity and in frequency space, is bell shaped and opening downwards. The sign convention is that for a positive bias ($\omega_2 > \omega_1$), the beam defined as frequency (ω_2) travels in opposition to the rotation rate. Positive rotation is in the direction of the rotation, so that when the velocity is in the direction of frequency (ω_2), the velocity is negative and the bias is positive. As has been demonstrated, the frequency tuning characteristics of the dispersion corrected scale factor and the velocity induced bias are identical.

Avoiding gas flow induced bias error is critical in the design of a practical RLG. In a typical design, a sensitivity of 4 deg/hr/cm/sec was found[51]. For aircraft navigation grade gyros, where a bias stability on the order of milli-deg/hr is required, the above sensitivity puts extremely stringent requirements on the gyro design. Effort must be put on the thermal design to avoid gradients and turn-on transients must not occur.

In addition to an externally driven gas flow, it has been shown that the presence of a dc discharge is sufficient to result in gas flow. The common solution is the use of a parallel discharge to cancel the flow effects. Sensitivities to current unbalance of 0.007 deg/hr/μA have been[51], indicating the need for a discharge current circuit stability of < 1 μA. The actual bias sensitivity to current unbalance is a strong function of the gyro design. For example, Eq (4.17) can be expressed in terms of the gain and velocity flow of the individual legs (1,2) of a parallel discharge as

$$(4.18) \qquad \Omega_b = \text{const} \times (G_2 V_2 - G_1 V_1)$$

The above defined sign convention for gas flow has been used, so that a flow along the direction of the beam of frequency (ω_2), results in a positive bias. In a typical RLG design, the gain is proportional to discharge current (with coefficient **m**). If the gas flow velocity is also proportional to discharge current (with coefficient **n**), then the bias, as given by Eq (4.18) can be written as

$$(4.19) \qquad \Omega_b = \text{const} \times \frac{\left[\overline{m}\,\overline{n}\,\Delta i\,\overline{I} + \frac{1}{2} \times (\overline{m}\,\Delta n + \overline{n}\,\Delta m)\,\overline{I}^2 \right]}{1 + y\,\overline{I}}$$

where the discharge current coefficients and the discharge currents ($\mathbf{I_c}$) of each leg are written as the average and difference between the two legs. The denominator factor representing gain saturation has been included in the same form as used for the scale factor equation, given by Eq (4.18). A linear factor (**y**) relating laser power to discharge current has been used.

With the above assumptions, the gas flow bias consists of two terms. The first is proportional to the product of the average discharge current and the current unbalance and represents the effect of discharge current unbalance. The second represents

the effect of asymmetry between the two legs. It shows that even with a balanced discharge current, there must be gain as well as gas flow symmetry. The current balanced asymmetry bias is expected to increase with the square of the discharge current.

Based on early experiments, when the discharge currents were unbalanced for the $\Delta i > 0$ case, a positive bias was found. This required a gas flow mechanism such that the flow was from the anode to the cathode. The Langmuir flow mechanism[52] was thought to be the source of this flow. Based on Eq (4.19), the gas flow bias sensitivity (deg/hr/mA current unbalance) is expected to increase with increasing discharge current. Unpublished experiments indicated the reverse occurs. With the example[53] shown in Figure 4.3, for higher values of discharge current, a reversal of the bias sensitivity is shown to occur. Figure 4.4 shows bias data for the same gyro, obtained at different values of discharge current (sum of the two legs). The data was obtained for the case of an current unbalance (difference between the two legs) of ± 0.1mA and for the case of the currents being balanced. A crossover in sensitivity can clearly be seen. Also a saturation is found at high discharge currents, which is consistent with the high power saturation term shown in Eq (4.19). The experiments indicated that the constituency of the gas used for the gain mixture and geometry determined the exact form of the sensitivity. For example, a constant pressure gas mixture consisting of higher He:Ne ratios resulted in a downwards displacement of the sensitivity curve. For ratios approaching 20:1, a cathode to anode flow was found at all discharge current levels. Based on data of this form, it is difficult to understand how Langmuir flow can be the dominant source of the observed gas flow bias in the RLG.

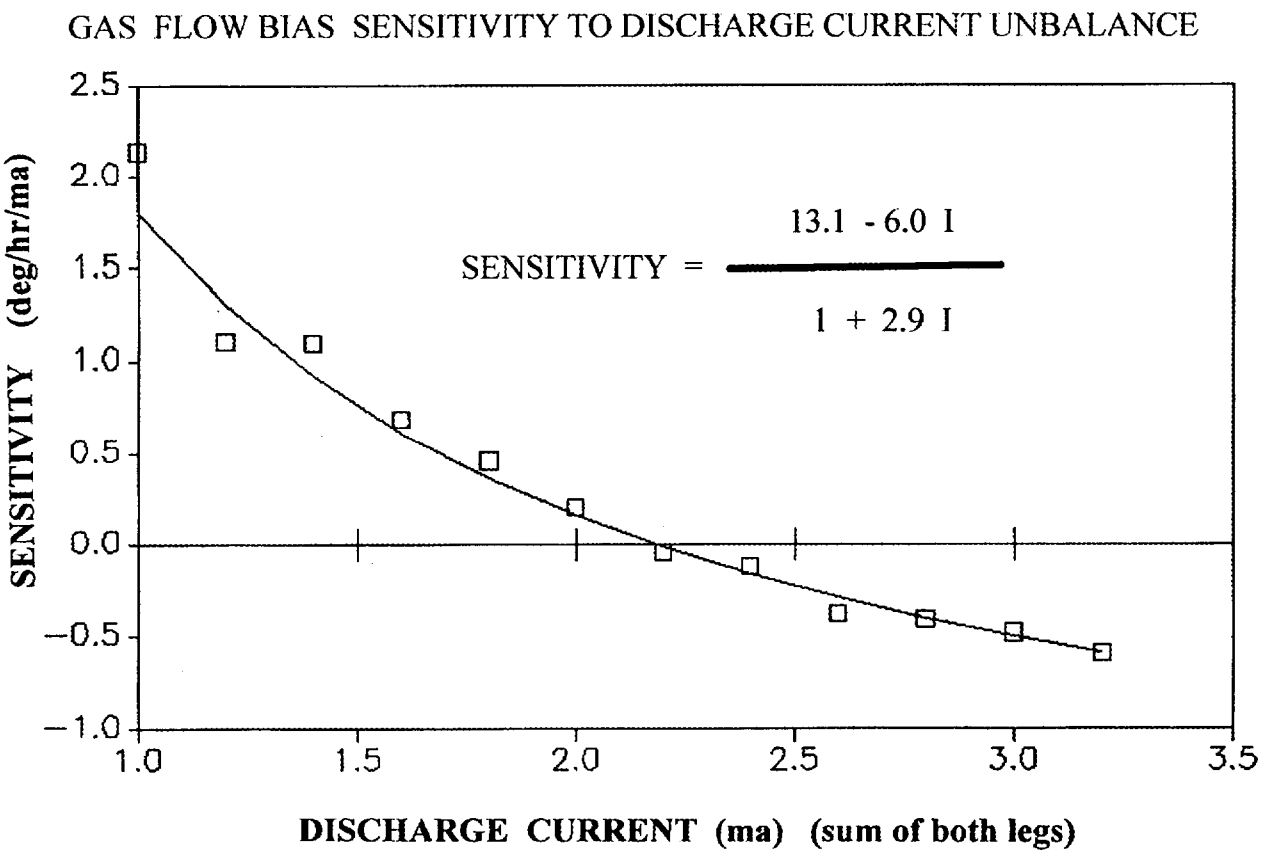


Figure 4.3. Bias sensitivity to total discharge current, as a function of total discharge current. The sensitivity is shown to change sign at higher value of the discharge current. A curve fit of the sensitivity is shown in the figure.

In an attempt to understand the source of the gas flow bias, this author has taken an alternate approach to allow a fit of the experimentally observed data. The approach was to assume the presence of two competing mechanisms. The first must be discharge current dependent, but result in a flow from the cathode to the anode. This will result in a flow having the negative bias sensitivity characteristic. The second term requires a discharge current independent contribution to the gas flow, which is in the anode to cathode direction.

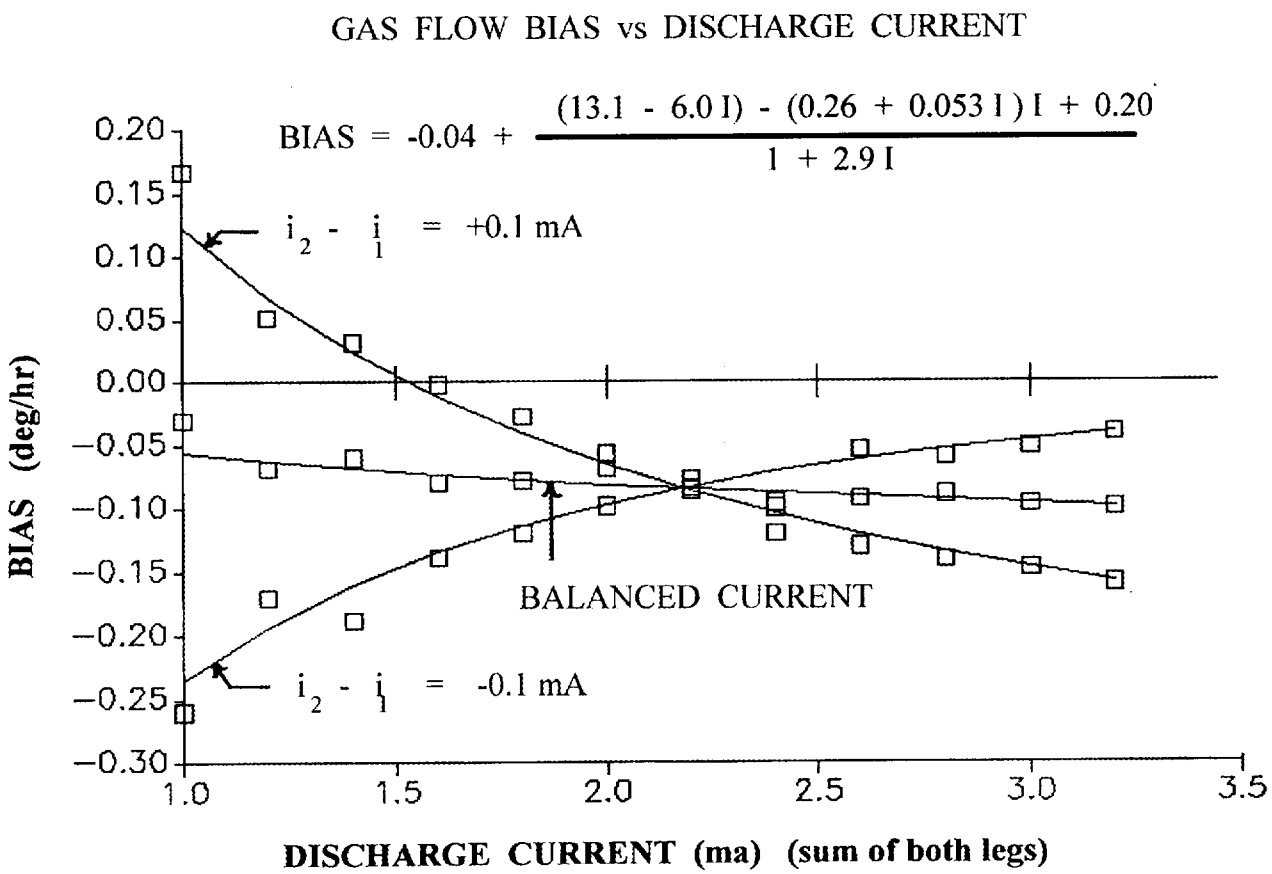


Figure 4.4. These curves show the gyro bias for both balanced and unbalanced discharge current between each side of the gyro. The fit of the data is shown in the figure. Comparing the data to that shown in Fig 4.3, the case of no bias sensitivity to discharge current unbalance, at a total discharge current of 2.2 mA, is clearly seen. This data was obtained with a RLG of 38 cm cavity length and square geometry. The RLG was built by Rockwell, Inc., now Boeing North American.

A possible mechanism for the discharge current independent contribution to gas flow, which is used only as an example to illustrate the requirements for the observed gas flow, might be dissociative recombination. In this mechanism ionized neon molecules (Ne_2^+) are moved by the electric field of the discharge towards the cathode. During this flow, recombination occurs, and the two neutral atoms, remain in excited states corresponding to the levels ($\mathbf{n_b}$, $\mathbf{n_a}$, $\mathbf{E_b} > \mathbf{E_a}$) between which the atomic transition occurs. It is assumed that the excited atoms populated by the dissociative recombination can move with different velocities $\mathbf{v_b}$, $\mathbf{v_a}$ along the beam directions. Then using the approach of replacing the frequencies ω_2, ω_1 with $\omega_2 + \omega_{b,a}$ and $\omega_1 - \omega_{b,a}$ the first order dispersion coefficients become

(4.20a) $\sigma_2 \equiv \sigma(\omega_2) \rightarrow n_b F(\omega_2 + v_b) - n_a F(\omega_2 + v_a)$

(4.20b) $\sigma_1 \equiv \sigma(\omega_1) \rightarrow n_b F(\omega_1 - v_b) - n_a F(\omega_1 - v_a)$

where **F** is the σ term with the gain removed. The gain contribution is proportional to the population difference $\mathbf{n_b} - \mathbf{n_a}$ and for this example, each level is considered to have a different velocity flow. The beat frequency can be obtained in manner similar to that used to obtain Eq (4.17), which, considering only the gas flow terms, results in a contribution of the form

(4.21) $\Omega_b = \Omega - \text{const} \times [\, (n_b - n_a)\, \overline{V}\, F_v + (n_b + n_a)\, \Delta v\, F_v \,]$

where $\overline{V}$ and Δ v represent the sum and difference of the $\mathbf{v_a}$ and $\mathbf{v_b}$, and $\mathbf{F_v}$ represents the derivative of the dispersion curve. Since, for the low gain He-Ne visible transition the population difference is extremely small, it can be expected that $\mathbf{n_b - n_a} << \mathbf{n_b + n_a}$. Thus even if $\mathbf{v} << \mathbf{V}$, either term could give rise to the required gas flow mechanism, if the velocities are independent of discharge current. Then, for a two leg discharge and writing the gas flow bias model in terms of discharge current, the gas flow contribution to the bias Ω_b, can be written as

(4.22) $\Omega_b = E_1 + [(E_2 - E_3 I)\,\Delta i_c + (E_4 + E_5 I)\,I_c + E_6] / (1 + E_7 I)$

where: $\mathbf{I_c}$ is the sum of the discharge currents for each leg;
$\Delta \mathbf{i_c}$ is the discharge current unbalance between the two legs;
$\mathbf{E_1}$ is the non-gas flow contribution to zero rate offset;
$\mathbf{E_2}$ is the discharge current independent gas flow term;
$\mathbf{E_3}$ is the discharge current dependent gas flow term;
$\mathbf{E_4}$ is the asymmetric current independent gas flow term;
$\mathbf{E_5}$ is the asymmetric current dependent gas flow term;
$\mathbf{E_6}$ ad hoc term introduced to provide a better fit to data;
$\mathbf{E_7}$ is the high gain saturation term.

The solid curves in Figures 7,8 represent a fit of Eq (4.22) to the data and the coefficients are shown in the figures.

Gas flow bias remains one of the areas in which further understanding of the mechanism is needed.

3.5 LOCK-IN

3.5.1 Description Of Lock-In

As previously described, the RLG senses rotation by a measurement of the frequency difference between the oppositely directed traveling waves. As is well known in the theory of oscillators, coupling between the two can lead to frequency synchronization, resulting in the two waves oscillating at the same frequency[54]. This frequency synchronization, commonly described as "lock-in", typically results in the RLG not being capable of sensing rotation and hence not usable as a practical device.

Lock-in in the RLG was reported shortly after the operation of the first RLG[55]. The magnitude of the lock-in band was large compared to the desired range of operation. Thus, for the successful development of a practical gyro, some approach to eliminating or avoiding lock-in was necessary. This section deals with a mathematical description of the lock-in mechanism. Lock-in is not only the most dramatic and most publicized of the characteristics of the RLG, but it is also the most important driver with respect to the design of a practical device. Hence a detailed review of the mechanism is presented.

An early publication[40] described lock-in as arising from mutual coupling between the two beams. The coupling was considered to be scattering of energy from each of the beams into the direction of the other. A simple description of the lock-in mechanism can be described for the special case of scattering of one of the beams into the direction of the other, using a vector phase-space approach. In Fig 5.1 , $\mathbf{E_1}$ and $\mathbf{E_2}$ represent the amplitudes of the two beams and ψ, the instantaneous phase difference between the two. The phase vectors are represented on a set of rotating axis such that the phase of field $\mathbf{E_1}$ is stationary. The additional phase shift per pass due to scattering from $\mathbf{E_2}$ onto $\mathbf{E_1}$, is represented by the vector $\mathbf{r_2 E_2}$. The phase β is a constant arising from the scattering process. Due to the additional phase shift per pass from scattering, the phase shift per pass from stimulated emission, will be shifted such that the total phase shift $\Delta\phi$, is still equal to an integral multiple of 2π. As seen from Fig 5.1, the shift in phase is given by

(5.1) $\alpha = r_2 (E_2 / E_1) \sin(\psi + \varepsilon_2) = \Delta\phi$ / pass = $-\Delta\omega$ x (time / pass)

The time/pass is just the inverse of the longitudinal mode spacing, or (**L/c**). Then the frequency shift of beam one, due to back scattering from beam two, is

(5.2) $\Delta\omega = -r_2 (c / L)(E2 / E1) \sin(\psi + \varepsilon_2)$

Equation (5.2) is a correction to the ideal beat frequency, resulting in an equation of the form

(5.3) $(d\psi/dt) = \Omega - \Omega_L \sin(\psi + \chi)$

where χ is a scattering induced phase angle, and

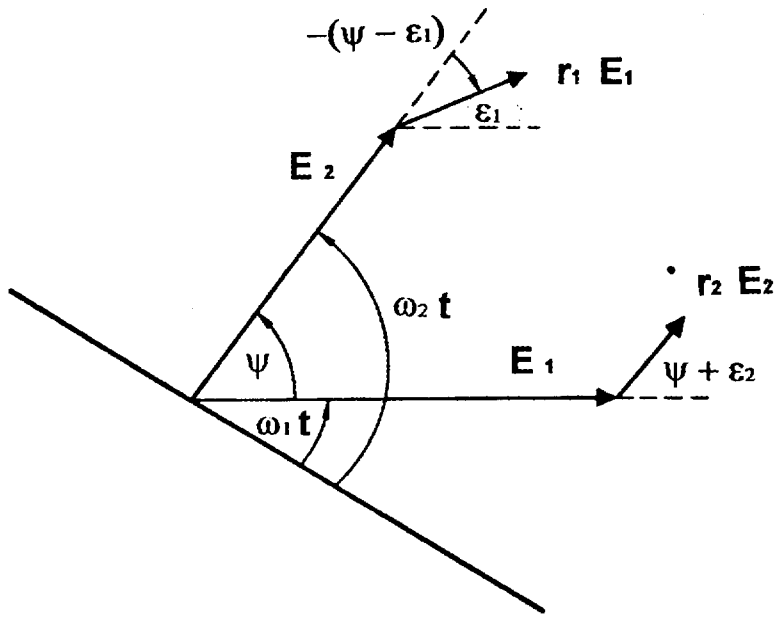


Figure 5.1. Vector phase space representation of the amplitudes of each of the oppositely directed beams. The scattering from each of the beams is shown added (with phase shift) to the beam traveling in the same direction as the scattering component.

(5.4) $\Omega_L = r_2 (c / L) (E_2 / E_1)$

If mutual back scattering was considered, the analysis would be similar, and Eq (5.3) would still apply. However, the expressions for the terms equivalent to χ and Ω_L would be more complicated.

Equation (5.3) describes lock-in and the characteristics of the equation are well understood. There are two types of solutions. For rotation rates less than the value Ω_L, a steady state solution exists, of the form

(5.5) $\psi + \chi = \sin^{-1} (\Omega / \Omega_L)$

Thus the phase difference between the two beams is constant, indicating both to be operating at the same frequency. Note that from Eq (5.5), as the rotation rate traverses the lock-in band, the phase shift changes by π.

For rotations greater than the threshold Ω_L, the equation has a time varying solution of the form

(5.6) $\tan^{1/2} (\psi + \chi - \frac{1}{2}\pi) = [(K_L + 1) / (K_L - 1)]^{1/2} \tan [\frac{1}{2} \Omega_L (K_L^2 - 1)^{1/2} t]$

where $K_L = \Omega / \Omega_L > 1$, and $\psi = \frac{1}{2}\pi - \chi$ @ t=0

For input rates much greater than the threshold, or $\mathbf{K_L} >> 1$, Eq (5.6) reduces to the form

(5.7) $\psi = \Omega t + \frac{1}{2}\pi - \chi$

Equation (5.7) shows the phase difference between the oppositely directed beams to increase linearly with time at the beat frequency rate, which is the expected result for the ideal case of no back scattering coupling. From Eq (5.3), it can be seen that

the beat frequency is not constant in time, giving rise to nonlinear wave forms[23]. These nonlinear wave forms are readily observed by displaying the beat frequency on an oscilloscope. As can be found from Eq (5.6), the solution is periodic and the period of time over which the phase changes by 2π is

(5.8) $$T = 2\pi \left[\Omega_L (K^2 - 1)^{1/2} \right]^{-1}$$

This period can be used to define the beat frequency Ω_b of the two beams as

(5.9) $$\Omega_b = \left[\Omega^2 - \Omega_L^2 \right]^{1/2}$$

Equation (5.9) shows the beat frequency between the two beams, which has been defined in terms of the periodicity of the instantaneous phase difference, to be reduced from the no-coupling value. As the rotation rate is reduced to the lock-in value Ω_L,, the beat frequency reduces to zero. As the beat frequency approaches zero, the wave form becomes more nonlinear. From the form of Eqs (5.3) and (5.5), irrespective of how high the input rate, the beat frequency is always modulated by the lock-in rate.

3.5.2 Formal Description Of Lock-In

The effects of back scattering have been formally described using the semi-classical model with back scattering fields represented by a fictitious current density[37]. The back scattering electromagnetic fields were represented by a fractional ($\mathbf{r_1}$, $\mathbf{r_2}$) scattering of each of the fields ($\mathbf{E_1}$, $\mathbf{E_2}$) into the direction of the other. The scattering process was assumed to have introduced an additional phase angle (ε_1, ε_2). With the large arrows representing the direction of the traveling waves, and the small arrows representing the back scattering fields, the vector representation of these fields were given as

$$E_2(\omega_2 t + \chi_2) \longleftarrow \qquad \longrightarrow E_1(\omega_1 t + \chi_1)$$

$$r_1 E_1(\omega_1 t + \chi_1 + \varepsilon_1) \longleftarrow \qquad \longrightarrow r_2 E_2(\omega_2 t + \chi_2 + \varepsilon_2)$$

$$\psi = (\omega_2 t + \chi_2) - (\omega_1 t + \chi_1)$$

With the addition of back scattering coupling, the equations describing the intensities of the two beams and the frequency difference, were found as

(5.10) $$(dI_1/dt) = (c / L)\, I_1 \left[\alpha - \beta I_1 - \theta I_2 - 2r_2 (I_2 / I_1)^{1/2} \cos(\psi + \varepsilon_2) \right]$$

(5.11) $$(dI_2/dt) = (c / L)\, I_2 \left[\alpha - \beta I_2 - \theta I_1 - 2r_1 (I_1 / I_2)^{1/2} \cos(\psi - \varepsilon_1) \right]$$

(5.12) $$(d\psi/dt) = \Omega + (c / L) \left[r_2 (I_2 / I_1)^{1/2} \sin(\psi + \varepsilon_2) + r_1 (I_1 / I_2)^{1/2} \sin(\psi - \varepsilon_1) \right]$$

For simplicity in the writing of Eqs (5.10) - (5.12) which the treatment of back scattering, the Lamb coefficients have been simplified by neglecting any asymmetric contribution. Corrections to the beat frequency due to "frequency-pulling" (which is intensity independent) and "frequency-pushing" (which is intensity dependent) have also been neglected. The "frequency-pulling" results in a gain-dependent scale factor correction and is not of interest for a description of lock-in. The "frequency-pushing" can affect lock-in, but has been shown to be an odd function[43] about the normal frequency of operation of the RLG. Hence, it gives zero contribution for normal operation, although usage of this term can provide insight into details affecting the design of the RLG[13].

Equations (5.10) - (5.12) can be simplified by working with a phase difference defined as[56]

(5.13) $$\phi = \psi - \pi + \frac{1}{2} (\varepsilon_2 - \varepsilon_1)$$

and the average of the scattering phase angle, given by

(5.14) $$\varepsilon = \frac{1}{2} (\varepsilon_2 + \varepsilon_1)$$

Then Eqs (5.10) - (5.12) become

(5.15) $(dI_1/dt) = (c / L) I_1 [\alpha - \beta I_1 - \theta I_2 + 2r_2 (I_2 / I_1)^{1/2} \cos(\psi + \varepsilon)]$

(5.16) $(dI_2/dt) = (c / L) I_2 [\alpha - \beta I_2 - \theta I_1 - 2r_1 (I_1 / I_2)^{1/2} \cos(\psi - \varepsilon)]$

(5.17) $(d\phi/dt) = \Omega - (c / L) [r_2 (I_2 / I_1)^{1/2} \sin(\phi + \varepsilon) + r_1 (I_1 / I_2)^{1/2} \sin(\phi - \varepsilon)]$

These are the basic equations that describe the operation of the RLG with the presence of back scattering coupling. Note that these equations do not show a dependence on the phasing of the scattering from each direction. They depend only on the average scattering. Because of the time varying intensity in the lock-in coefficient, an exact closed form solution has never been obtained. Computer solutions of the equations are easy to obtain, but they do not provide the insight of closed form solutions. One problem with a computer solution, is that it is not straight-forward to determine the rotation value at which the solutions change from steady-state to oscillatory. This difficulty occurs because there are multiple solutions which give rise to a hysteresis. By this it is meant that the lock-in threshold is higher when it is approached with increasing rotation, as opposed to a decreasing rotation.

Usage of the third order theory is not a major limitation, in that the scattering phenomena can be adequately described by the small coupling terms. Also, to obtain closed form solutions, a linear approximation is the usual approach. As can be seen from the two intensity equations, there is a modulation of each due to the back scattering coupling. The coupling term in the beat frequency equation, is similar to the lock-in equation given by Eq (5.3), with a more complicated form of the lock-in threshold.

3.5.3 Analysis Of The Intensity And Beat Frequency Equations

Equations (5.15) - (5.17) are simpler to analyze and more understanding can be obtained by expressing both the intensities and scattering coefficients in terms of sums and differences, or

(5.18) $I = \frac{1}{2} (I_2 + I_1) \quad i = \frac{1}{2} (I_2 - I_1)$

(5.19) $R = \frac{1}{2} (r_2 + r_1) \quad r = \frac{1}{2} (r_2 - r_1)$

To simplify the writing of the equations, the **(c/L)** factor can be incorporated in the definition of the scattering average and difference. It is also more convenient to normalize all the Lamb coefficients in terms of rotation units by defining

(5.20) $\Omega_\alpha = (c / L) \alpha \qquad \Omega_\pm = (c / L) (\beta \pm \theta) \qquad \Omega_\beta = (c / L) \beta$

(5.21) $\Omega_G = \Omega_\alpha (\beta - \theta) / (\beta + \theta) \qquad \Omega_R = (c / L) R \qquad \Omega_r = (c / L) r$

In the analysis, certain combinations of the scattering occur, and in a formalism developed by the author, the equations can be further simplified by defining the quantities

(5.22) $\mathbf{A} = \Omega_R \cos\varepsilon \quad \mathbf{B} = \Omega_r \sin\varepsilon \quad \mathbf{a} = \Omega_R \sin\varepsilon \quad \mathbf{b} = \Omega_r \cos\varepsilon$

With these definitions, the beat frequency and intensity equations can now be written as

(5.23) $(dI/dt) = \Omega_\alpha I - \Omega_+ I^2 - \Omega_- i^2 + 2 (I^2 - i^2)^{1/2} (\mathbf{A} \cos\varphi - \mathbf{B} \sin\varphi)$

(5.24) $(di/dt) = \Omega_\alpha i - 2 \Omega_\beta i I + 2(I^2 - i^2)^{1/2} (\mathbf{a} \sin\varphi - \mathbf{b} \cos\varphi)$

(5.25) $(d\varphi/dt) = \Omega - [2 / (I^2 - i^2)^{1/2}][(\mathbf{A} I + \mathbf{b} i) \sin\varphi + (\mathbf{B} I + \mathbf{a} i) \cos\varphi]$

Equations (5.23) - (5.25) are difficult to simultaneously analyze in the non-steady state region. Computer solutions show that well above the lock-in threshold, the ac contributions to the intensities are small. Since the coefficients for the two beams have been treated as symmetric (no asymmetry in laser gain or loss), there is no dc term in the differential intensity. Then, the complex term $(\mathbf{I}^2 - \mathbf{i}^2)^{1/2}$ can be simply replaced by **I**, as a first order approximation. With this approximation, the equations can be shown to give closed form solutions in a straight forward manner, although care must be taken in the approach. A successful technique has been one of iteration[57]. The solutions, which are most valid well above the lock-in threshold, are taken to be of the form

(5.26) $I = I_o + I_1 \sin(\varphi + \eta)$

(5.27) $i = i_1 \sin(\varphi + \theta)$

(5.28) $\varphi = \omega t$

where, as shown by Eqs (5.26) and (5.27), η and θ are the phases of the total and differential intensities, respectively, with respect to the beat frequency.

The total intensity can be determined by substituting Eqs (5.26) - (5.28) into Eq (5.23) and equating coefficients of sin φ, cos φ and the constant term. The solutions are then found as

(5.29) $I_o = [\,\Omega_\alpha - \frac{1}{2}\Omega_+ I_o (I_1/I_o)^2 + (\mathbf{A} \sin\eta - \mathbf{B} \cos\eta)(I_1/I_o) - \frac{1}{2}\Omega_- I_o (i_1/I_o)^2\,] / \Omega_+$

(5.30) $(I_1/I_o) \sin\eta = 2(\mathbf{A}\ \Omega_\alpha + \mathbf{B}\ \Omega) / (\Omega^2 + \Omega_\alpha^2)$

(5.31) $(I_1/I_o) \cos\eta = 2(\mathbf{A}\ \Omega - \mathbf{B}\ \Omega_\alpha) / (\Omega^2 + \Omega_\alpha^2)$

When Eqs (5.30) and (5.31) are substituted into Eq (5.29) and the zero order solution for $\mathbf{I_o}$ is used to simplify Eq (5.29), the terms in $(\mathbf{I_1}/\mathbf{I_o})$ cancel. Because of the term in $\mathbf{i_1}$ in Eq (5.29), the solution of the differential intensity equation is required to obtain an explicit dependence of $\mathbf{I_o}$ on the gyro parameters. Equations (5.30) and (5.31) can be solved to give

(5.32) $I_1 = 2 I_o [\,(\mathbf{A}^2 + \mathbf{B}^2) / (\Omega^2 + \Omega_\alpha^2)\,]^{1/2}$

(5.33) $\tan\eta = (\mathbf{A}\ \Omega_\alpha + \mathbf{B}\ \Omega) / (\mathbf{A}\ \Omega - \mathbf{B}\ \Omega_\alpha)$

Equations (5.26), (5.29) and (5.32) show the sum of the intensities of the two beams to have a modulation at the beat frequency rate, with a proportionality coefficient $(\mathbf{A}^2 + \mathbf{B}^2)$. The modulation decreases inversely with the rotation rate. Equation (5.33) defines the phase of the total intensity modulation with respect to the beat frequency.

Now consider the solution of the differential intensity term. In the product term $\mathbf{i\,I}$, the term in $\mathbf{I_1}$ can be dropped, since it is second order in scattering. The quadratic term in intensity can be replaced by $\mathbf{I_o}$. From Eq (5.29), $\mathbf{I_o}$ can be expressed in terms of the non-scattering solution. With these simplifications, and with the definitions in Eqs (5.20) and (5.21), the differential intensity equation, as given by Eq (5.24), can be written as

(5.34) $(di/dt) = -\Omega_G\, i + 2 I_o (\mathbf{a} \sin\varphi - \mathbf{b} \cos\varphi)$

Then by writing the solution of the form

(5.35) $i = i_1 \sin(\varphi + \theta)$

and as before, substituting into Eq (5.34) and equating coefficients of sin φ and cos φ the solution is found as

(5.36) $i_1 \cos\theta = 2 I_o (\mathbf{a}\ \Omega_G - \mathbf{b}\ \Omega) / (\Omega^2 + \Omega_G^2)$

(5.37) $i_1 \sin\theta = -2 I_o (\mathbf{a}\ \Omega + \mathbf{b}\ \Omega_G) / (\Omega^2 + \Omega_G^2)$

Equations (5.36) and (5.37) can be solved to give

(5.38) $\tan\theta = -(\mathbf{a}\ \Omega + \mathbf{b}\ \Omega_G) / (\mathbf{a}\ \Omega_G - \mathbf{b}\ \Omega)$

(5.39) $i_1 = 2 I_o [\,(\mathbf{a}^2 + \mathbf{b}^2) / (\Omega^2 + \Omega_G^2)\,]^{1/2}$

Equations (5.35), (5.38) and (5.39) show the intensity difference between the two beams to have a modulation at the beat frequency rate, with the proportionality coefficient $(\mathbf{a}^2 + \mathbf{b}^2)$. The modulation can be seen to decrease inversely with the rotation rate. Equation (5.38) defines the phase, referenced to the phase of the beat frequency as defined from Eq (5.38), of the total intensity modulation.

Now using the solution for the differential intensity, the expression for the total intensity, as given by Eq (5.29) becomes

(5.40) $I_o = [\, \Omega_\alpha - 2\, \Omega_G\, (\mathbf{a}^2 + \mathbf{b}^2) / (\Omega^2 + \Omega_G^2)\,] / \Omega_+$

The scattering term gives little change to the total intensity and in what follows, it will be ignored.

Now that there are expressions for the intensities in the region well above lock-in, it is possible to put the beat frequency equation given by Eq (5.25), in a more recognizable form. Neglecting the term in $\mathbf{i}^2$ and replacing the total intensity by $\mathbf{I_o}$, Eq (5.25) becomes

(5.41) $(d\phi/dt) = \Omega - 2\, (\mathbf{A} \sin\phi + \mathbf{B} \cos\phi) - 2\, (i / I_o)(\, \mathbf{b} \sin\phi + \mathbf{a} \cos\phi)$

Using the solution for the differential intensity, as given by Eqs (5.38) and (5.39), Eq (5.41) can be put in the form

(5.42) $(d\phi/dt) = \Omega\, [1 + 2\, (\mathbf{a}^2 + \mathbf{b}^2) / (\Omega^2 + \Omega_G^2)\,] - \Omega_{L1} \sin(\phi + \beta_1) - \Omega_{L2} \sin(2\, \phi + \beta_2)$

where

(5.43) $\Omega_{L1} = 2(\mathbf{A}^2 + \mathbf{B}^2)^{1/2} \qquad \tan\beta_1 = \mathbf{A} / \mathbf{B}$

(5.44) $\Omega_{L2} = 2\, (\mathbf{a}^2 + \mathbf{b}^2) / (\Omega^2 + \Omega_G^2)^{1/2}$

(5.45) $\tan\beta_2 = [\, (\mathbf{a}^2 + \mathbf{b}^2)\, \Omega + 2\, \mathbf{a}\, \mathbf{b}\, \Omega_G\,] / [\, 2\, \mathbf{a}\, \mathbf{b}\, \Omega - (\mathbf{a}^2 - \mathbf{b}^2)\, \Omega_G\,]$

Equation (5.42) is the generalized expression for the beat frequency in the RLG. It is similar in form to the previously discussed lock-in equation, as given by Eq (5.3). It has been formally derived using mutual back scattering from one beam into the other. It clarifies many of the diverse characteristics experimentally demonstrated with various types of gyros. It shows the positive scale factor correction previously reported[57]. It also shows a first order (linear in scattering) and second order (quadratic in scattering) lock-in contribution. A negative scale factor correction contribution can be obtained from the first order lock-in term by assuming the second order term to be smaller than the first and writing the phase ϕ, as

(5.46) $\phi = \omega\, t + \delta \sin(\phi + \lambda)$

Then by substituting Eq (5.46) into Eq (5.42), neglecting the second order term, and equating powers of sin φ and cos φ, the constant term gives

(5.47) $\omega = \Omega\, [\, 1 - 2\, (\mathbf{A}^2 + \mathbf{B}^2) / \Omega^2 + 2\, (\mathbf{a}^2 + \mathbf{b}^2) / (\Omega^2 + \Omega_G^2)\,]$

The scale factor correction is simply the deviation of the bracketed term from unity. This completes the determination of the solutions of the equations defining the intensities and beat frequency of the two beam in the RLG. The next step is the analysis of the solutions to gain insight into the operation of the gyro.

3.5.4 Lock-In For The Case Of Symmetric Scattering

For the first case, consider the simplest situation, which is when the scattering is symmetric, or from Eq (5.19), (5.21) and (5.22), $\mathbf{r} = 0$, and $\mathbf{B} = \mathbf{b} = 0$. Then it can be expected that since no energy is lost from the cavity, the system will pump energy from one beam to another. Thus the intensity sum will be constant. As can be seen from Eq (5.32), this would only be possible if the previously arbitrary phases of the individual traveling waves adjust such that $\mathbf{A} = 0$. From Eq (5.22) this puts a constraint on the average scattering phase to be $\varepsilon = \pi/2$. From Eq (5.24), it can be seen that with no energy pulsation, the differential intensity is modulated at the beat frequency rate. Hence energy is periodically transferred between the two beams. This concept has been previously discussed and has been sometimes described as a conservative system with hermitian coupling[58]. With this type of coupling, the intensity equations have simple solutions, which are of the form

(5.48) $I = I_o = \Omega_\alpha / \Omega_+$

(5.49) $i = i_1 \sin(\phi + \theta)$

(5.50) $i_1 = 2\,\Omega_R\, I_o / (\Omega^2 + \Omega_G^2)^{1/2} \qquad \tan\theta = -\Omega / \Omega_G$

The lock-in equation, given by Eq. (5.42) shows the interesting result that the first order lock-in threshold value is zero. The second order lock-in threshold value, given by Eq. (5.44) reduces to the value

(5.51) $\Omega_{L2} = 2\,\Omega_R^2 / \Omega_G$

In writing Eq (5.51), the term in Ω in the denominator of Eq (5.44) was dropped, since at threshold the lock-in value is much smaller than the Ω_G value[57]. Since the scattering rate is also much smaller than the Ω_G value, symmetric scattering results in a lock-in threshold much smaller than would be expected from simple back scattering. The phase vector model of scattering shows the rationale for this. Equation (5.43) shows the first order lock-in threshold to scale with $(\mathbf{A}^2 + \mathbf{B}^2)$ and using the definitions of $\mathbf{A}$, $\mathbf{B}$ from Eq (5.22), the lock-in threshold can be written in terms of the individual scattering coefficients, as

(5.52) $\Omega_{L1}^2 \rightarrow r_1^2 + r_2^2 + 2\, r_1 r_2 \cos(2\,\varepsilon)$

For symmetric scattering and for $\varepsilon = \pi/2$, Eq (5.52) shows that the scattering coefficients act in opposition to one another and cancel. For symmetric scattering, the scale factor correction, as shown from Eq (5.42), is shown to be positive, scales with a quadratic dependence on the scattering and contains a Lorentzian function of both the rotation and the excess gain. By considering a third order (second order gives no contribution) correction term in Eq (5.46) a negative scale factor correction that dominates close the lock-in threshold can be found. It has the value $(1/2)\,\Omega_{L2}^2/\Omega^2$, where Ω_{L2} is given by Eq (5.44). As expected, the lock-in threshold, given by Eq (5.51), scales inversely with the excess gain.

3.5.5 Lock-In For The Case Of Asymmetric Scattering

It should be noted that although most of the time the RLG shows a positive scale factor correction[57], negative corrections do occur. In fact, it is possible to observe both a positive and negative scale factor correction on the same gyro as it is taken over temperature. This is not predicted with the assumption of simple symmetric scattering.

The next step is to consider asymmetric scattering. Referring to Eq (5.32), it can be seen that the modulation of the total intensity can no longer be set to zero. This mean that a measurement of the sum of the intensities would show a modulation at the beat frequency rate. No such modulation has been reported. At best, the modulation coefficient, which is of the form

(5.53) $(\mathbf{A}^2 + \mathbf{B}^2) = (\Omega_R^2 \cos^2\varepsilon + \Omega_r^2 \sin^2\varepsilon)$

can be minimized by again choosing $\varepsilon = \pi/2$. In this case, although the scattering coefficients are in opposition, they do not completely cancel. Hence energy couples unequally from one beam to the other, with a resultant pulsation. Irrespective of the choice of ε, the modulation of the intensity sum is small compared to the total intensity.

For the case of asymmetric scattering, the form of the lock-in threshold is different. The second order contribution with its inverse gain dependence, as given by Eq (5.44), remains the same. But now the first order contribution, as given by Eq (5.43), is non-zero and using Eq (5.22), is

(5.54) $\Omega_{L1} = 2\,\mathbf{B} = 2\,\Omega r$

Comparing this value to the one in Eq (5.51), it can be seen that it may or may not be the dominant contribution. Based on experience, differential scattering is small, and this term is not expected to dominate. From Eq (5.47), this means that the scale factor correction will remain positive, making it difficult to correlate the theory with the experimental data. Thus, the assumption of differential scattering being the sole source of lock-in goes against RLG laboratory experience. Another mechanism capable of contributing to lock-in must be considered.

F. LOSS AS A CAUSE OF LOCK-IN

A number of authors have considered a generalized model [13,58-60] for the RLG. They treat conservative and non-conservative loss coupling, which have sometimes been called hermitian and skew hermitian coupling, respectively. The treatment[59] presented here is the one that follows the formalism used for the pure scattering case.

The concept of loss being a cause lock-in can be presented from the following ideal scatter-free case. Consider all the cavity loss to be represented by a uniformly distributed loss plus a delta function loss which is located at a fixed point in the cavity. When the RLG is stationary and for a minimum loss condition, the two beams form a standing wave with a node at the delta function loss. As the RLG is rotated, the frequencies would tend to shift to provide a frequency difference proportional to rotation. This would result in a change from a standing wave to a pair of oppositely directed traveling waves, which would result in the nodes traveling across the loss point at the beat frequency rate. Hence energy would be lost at the delta function loss and this would result in a higher loss system than would occur for a standing wave. The laser oscillator system would tend to resist this higher loss condition and remain frequency locked with the node at the delta function loss. However, this frequency locking of the rotating RLG would tend to result in a higher loss condition than for a stationary RLG. This occurs since both traveling waves are oscillating at a frequency other than the one which results in a minimum Fabry-Perot cavity loss. As the rotation increases, the mismatch in cavity frequency for the frequency locked laser would result in an increasing cavity loss. At the lock-in threshold, which is defined as the rotation rate where the mismatched cavity loss equals the delta function loss, the oscillators would unlock.

To analyze the case of loss-induced lock-in, non-uniform loss along the cavity path was treated[59] by considering the conductivity to have a spatial dependence. This resulted in the equations representing the intensities of the two beams and the beat frequency, as previously given by Eqs (5.10) - (5.12), to be modified to be of the form

(5.55) $(dI_1/dt) = (c / L)\, I_1 [\, \alpha - \beta I_1 - \theta I_2 - 2r_2 (I_2 / I_1)^{1/2} \cos(\psi + \varepsilon_2) + 2s\, (I_2 / I_1)^{1/2} \cos(\psi + \delta)\,]$

(5.56) $(dI_2/dt) = (c / L)\, I_2 [\, \alpha - \beta I_2 - \theta I_1 - 2r_1 (I_1 / I_2)^{1/2} \cos(\psi - \varepsilon_1) + 2s\, (I_1 / I_2)^{1/2} \cos(\psi + \delta)\,]$

(5.57) $(d\psi/dt) = \Omega + (c / L) [\, r_2 (I_2 / I_1)^{1/2} \sin(\psi + \varepsilon_2) + r_1 (I_1 / I_2)^{1/2} \sin(\psi - \varepsilon_1) - s\, [\, (I_2 / I_1)^{1/2} + (I_1 / I_2)^{1/2}\,] \sin(\psi + \delta)\,]$

The loss coefficient and phase angle, respectively were defined in terms of the non-uniform conductivity as

(5.58) $$s = \frac{1}{2\varepsilon_o L} \sqrt{\int_0^L \int_0^L \sigma(z)\sigma(z') \cos 2K_\lambda (z - z')\, dz\, dz'}$$

(5.59) $$\tan \delta = \int_0^L \sigma(z) \sin 2K_\lambda z\, dz \Big/ \int_0^L \sigma(z) \cos 2K_\lambda z\, dz$$

where the wave number $(2\pi / \lambda)$ is given by K_λ. Note that in Eqs (5.55) - (5.57), the functional dependence of the non-uniform loss terms are similar to the scattering terms, and hence can be expected to contribute to lock-in. The form of Eq (5.58) shows that large contributions occur when the spatial loss variation has a periodicity on the order of the laser wavelength. The spatial loss variation does not have to be a bulk effect. It can arise from surface scattering, when the scattered light is removed from the cavity. The effect is most pronounced when the surface scattering is due to a grating that gets "burnt" onto the mirror by the standing wave pattern of the two beams[25,50]. Equations (5.55) - (5.57) are exactly of the form of the previous intensity and beat frequency equations, as given by Eqs (5.23) - (5.25), with the non-uniform loss phase and amplitude, and the scattering coefficients as given by Eq (5.22), redefined as

(5.60) $\xi = \delta + \frac{1}{2} (\varepsilon_1 - \varepsilon_2) \qquad \Omega_s = (c / L)\, s$

(5.61) $\mathbf{A} = \Omega_R \cos\varepsilon - \Omega_s \cos\xi$

(5.62) $\mathbf{B} = \Omega_r \sin\varepsilon - \Omega_s \sin\xi$

With this definition, the solutions previously found for the intensity and beat frequency equations, do not change with the addition of non-uniform loss. However, the interpretation and conclusions obtained from the solutions are different.

First consider the case of no total intensity modulation. From Eq (5.32), the condition to satisfy this case is $\mathbf{A} = \mathbf{B} = 0$. Using the definition of the $\mathbf{A}$, $\mathbf{B}$ coefficients, as given in Eqs (5.61) and (5.62), conditions on the phase angles are found as

(5.63) $\cos^2 \varepsilon = (\Omega_s^2 - \Omega_r^2) / (\Omega_R^2 - \Omega_r^2)$

(5.64) $\tan \xi = (\Omega_r / \Omega_R) [(\Omega_R^2 - \Omega_s^2) / (\Omega_s^2 - \Omega_r^2)]^{1/2}$

Since $\cos \varepsilon \leq 1$ and is real, Eq (5.63) can only be satisfied for

(5.65) $\Omega_r \leq \Omega_s \leq \Omega_R$

Based on the described model, the analysis shows that when Eq (5.65) is not satisfied, it is possible for the total intensity to have a modulation in time[25]. Also, Eqs (5.63) and (5.64) shows that for the case of no intensity modulation, as the loss rate increases both the scattering and loss phase angles vary from $\pi/2$ to 0. Using Eq (5.63) to determine the value of the average scattering phase angle and from Eq (5.22), the scattering parameter $(\mathbf{a}^2 + \mathbf{b}^2)$ can be found as

(5.66) $\mathbf{a}^2 + \mathbf{b}^2 = \Omega_R^2 + \Omega_r^2 - \Omega_s^2$

From Eq (5.65), which describes the range of allowable values of Ω_s for the case of no total intensity modulation, Eq (5.66) shows that as the loss increases, the range of the scattering parameter to varies from Ω_R^2 to Ω_r^2. Since the differential intensity, lock-in and scale factor non-linearity all vary directly with the scattering parameter, this shows that a portion of the non-uniform loss can cancel the effects of scattering.

For the case when the non-uniform loss is larger than the average scattering, the analysis predicts that it is not possible to find a case when the total intensity does not have a modulation at the beat frequency rate. In the general case, it is difficult to determine constraints. However, in practice, differential scattering in the RLG is typically small, and it is a reasonable approximation to neglect it. Then it is possible to find constraints such as to minimize the amplitude of the total intensity modulation. The rationale for this is that the system will operate such as to maximize the total intensity. It will tend to avoid any condition which produces a minimum total intensity, and hence minimize the amplitude of the modulation. As seen from Eq (5.32) and (5.61), for the case of no differential scattering, the total intensity modulation coefficient is of the form

(5.67) $(\mathbf{A}^2 + \mathbf{B}^2)^{1/2} = [\Omega_s^2 + (\Omega_R \cos \varepsilon)^2 - 2\, \Omega_s (\Omega_R \cos \varepsilon) \cos \xi]^{1/2}$

By inspection, this term is minimized when $\xi = 0$. With this constraint, Eq (5.67) becomes

(5.68) $(\mathbf{A}^2 + \mathbf{B}^2)^{1/2} = (\Omega_s - \Omega_R \cos \varepsilon)$

Again, by inspection, this term is minimized when $\varepsilon = 0$ and the term becomes

(5.69) $(\mathbf{A}^2 + \mathbf{B}^2)^{1/2} = \Omega_s - \Omega_R$

As seen from Eqs (5.42) and (5.43), the first order lock-in contribution now becomes the major contributor and can be much larger than the second order contribution. Also, as seen from Eq (5.47), the scale factor correction now transitions from positive to negative.

If differential scattering was considered, the scattering phase angle would not differ greatly from zero. Using Eq (5.22), the scattering parameter $(\mathbf{a}^2 + \mathbf{b}^2)$, which determines differential intensity, the positive scale factor correction and the second order lock-in threshold becomes

(5.70) $(\mathbf{a}^2 + \mathbf{b}^2) = \Omega_r^2$

Thus the second order contributions tend to become negligibly small.

3.5.6 Summary Of Scattering And Loss Lock-In

The results of the above analysis have been summarized and are shown in Table 5.1. The results are separated for the case of "no" or "some" total intensity modulation. For the case of "some" total intensity modulation, it is assumed that differential scattering is zero, or at most a small fraction of the average scattering. For the case of "no" intensity modulation, the trending of the parameters are shown as the loss rotation coefficient ranges between the differential scattering rate and the average scattering rate.

Table 5.1 SUMMARY OF COUPLING-INDUCED PARAMETER VARIATION

PARAMETER	NO INTENSITY MODULATION	INTENSITY MODULATION
Loss Rate - (Ω_s)	for variation from Ω_r to Ω_R	for $\Omega_s > \Omega_R$, $\Omega_r \ll \Omega_s$
Loss Phase - (ξ)	for variation from $\frac{1}{2}\pi$ to 0	0
Scattering Phase - (ε)	for variation from $\frac{1}{2}\pi$ to 0	0
Total Intensity Modulation - (I_1)	0	$\frac{2 I_o(\Omega_s - \Omega_R)}{\sqrt{\Omega^2 - \Omega_\alpha^2}}$
Differential Intensity Modulation - (i_1)	$\frac{2I_o(\Omega_R^2 + \Omega_r^2 - \Omega_s^2)}{\sqrt{\Omega^2 - \Omega_G^2}}$	$\frac{2I_o\Omega_r}{\sqrt{\Omega^2 + \Omega_G^2}}$
for $\Omega \ll \Omega_G$	$\frac{2I_o\Omega_R}{\Omega_G}$ to $\frac{2I_o\Omega_r}{\Omega_G}$	$\frac{2I_o\Omega_r}{\Omega_G}$
Lock-In Threshold - (Ω_L)	$\frac{2(\Omega_R^2 + \Omega_r^2 - \Omega_s^2)}{\Omega_G}$	$2(\Omega_s - \Omega_R)$
as Ω_s goes from Ω_r to Ω_R	$\frac{2\Omega_R^2}{\Omega_G}$ to $\frac{2\Omega_r^2}{\Omega_G}$	
Scale Factor Correction	$\frac{2(\Omega_R^2 + \Omega_r^2 - \Omega_s^2)}{\Omega^2 + \Omega_G^2}$	$\frac{-2(\Omega_s - \Omega_R)^2}{\Omega^2}$
as Ω_s goes from Ω_r to Ω_R	$\frac{2\Omega_R^2}{\Omega_G}$ to $\frac{2\Omega_r^2}{\Omega_G}$	
$\Omega \gg \Omega_G$		

Ring laser gyro designers discovered that by adjusting the relative position of the mirrors that form the cavity that support the laser, the effective back scattering of the gyro could be changed[61-63]. This was done by moving two of the three mirrors of a three-mirror ring laser in a push-pull manner, so that the total path length remained constant. Since the scattering occurred at the individual mirrors, the net coupling to the beam was the vector superposition of the individual sources of scattering. Analysis[64] showed the spacing between the mirrors to be critical. The scattering parameters that were used in the above analysis can be considered to be the vector sum of all the scattering for each direction. Since it is possible to vary the magnitude of the scattering by push-pull motion of the mirrors, it is possible to consider the variation of the measurable quantities in Table 5.1, as a function of mirror tuning.

Analysis[64] has shown that a reasonable approximation of the back scattering coefficient as a function of mirror push-pull motion is of the form (1+cosx), where x is the normalized mirror push-pull motion. Then by specifying values for the parameters used

in Table 5.1, the measurable quantities can be calculated. Note that depending on the relative magnitude of the loss compared to the scattering contribution to lock-in, different formulas are used.

3.6 - DITHERED SCALE FACTOR

As previously discussed, it is difficult to accurately determine the lock-in threshold by reducing the rotation of the gyro and measuring the value at which the gyro locks. Information on the characteristics of the gyro, including lock-in threshold, can be obtained however, by measuring the scale factor at rates well above the lock-in threshold, and correlating the data with the mathematical model. In this section, a description of the types of data that can be expected, will be presented.

An expression, given by Eqs (5.22), (5.47) and (5.61), was presented that describes the scale factor in the presence of back scattering and spatially varying loss. Expressions, given by Eqs (5.43) and (5.44), were also presented that describe the lock-in thresholds, due to either back scattering or loss being dominant. The equations can be combined to give the scale factor deviation (measured as a relative deviation from the asymptotic value, commonly called scale factor) in terms of the lock-in thresholds as

$$(6.1) \qquad S = -\tfrac{1}{2}\,\Omega_{L1}^2/\Omega^2 + \Omega_{L2}\Omega_G/(\Omega^2+\Omega_G^{\,2})$$

Normally the defined scale factor deviation is multiplied by 10^6 and expressed in parts per million. In calculations to determine value, this will be done, but for brevity, the factor of 10^6 will be ignored. Note that when the back scattering contribution is much greater than the loss contribution, the term in Ω_{L1} will be zero and the scale factor correction at high rates will only consist of a positive term. At low rates and when loss dominates, as previously shown, a negative contribution will exist.. In what follows, the analysis will include both contributions, and depending on conditions, either or both will determine scale factor.

When the gyro is operated at a constant input rate, Eq (6.1) shows that a measurement of the scale factor can be used to determine the asymptote and hence the absolute value of the scale factor. Experiments of scale factor will be further discussed. A plot of the curvature of the scale factor, as a function of the rotation rate can determine the coefficients of Eq (6.1). The sign of the scale factor change immediately determines whether loss or scattering dominates.

A more interesting structure occurs when the gyro is operated with a time varying dither[5] and a constant rotation rate, or

$$(6.2) \qquad \Omega = \Omega_{Rate} + \Omega_D \cos\omega_d t$$

As can be seen from Eq (6.1), the gyro scale factor, which because of the dependence on input rate, becomes time varying. To determine the average scale factor over many dither cycles, it can be defined as the apparent angle through which the gyro is turned, divided by the true angle, or using Eq (6.1) for the scale factor and Eq (6.2) for the inertial rate,

$$(6.3) \qquad \underline{S} = \int_0^T s\,\Omega\,dt \Big/ \int_0^T \Omega\,dt$$

Although Eqs (6.1) - (6.3) are simple in form, the usual procedure is to numerically evaluate the convolution of scale factor/rotation rate. This approach is time consuming and does not show parameter dependence as one would obtain using a closed form solution. The above integral was shown[65] to be able to be evaluated in closed form, but since the procedure is involved, only the results will be presented here. As seen from Eq (6.1), the scale factor consists of two terms. To avoid the singularity at zero rate in the negative term, a term of magnitude Ω_{L1}^2 is added to the denominator, such at to make it similar in form to the positive term. Then the integral can be evaluated for one of the terms and the result applied to the other. The solution of Eq (6.3) can be expressed as follows, with the subscripts signifying which term is being considered.

$$(6.4) \qquad S\pm = [\,\Omega\pm/(2\Omega)\,]\,[\,U_1\pm + (U_1\pm^2 + U_2\pm^2)^{1/2}\,]^{1/2} / [U_1\pm^2 + U_2\pm^2]^{1/2}$$

$$(6.5) \qquad U_{1+,-} = \Omega^2 - \Omega_D^2 - \Omega_{G,L1}^2 \qquad\qquad U_{2+,-} = 2\,\Omega\,\Omega_{G,L1}$$

$$(6.6) \qquad \Omega_+^2 = \Omega_{L2}\,\Omega_G \qquad\qquad \Omega_-^2 = -\tfrac{1}{2}\,\Omega_{L1}^2$$

In the form shown, the dithered scale factor is complex and except for special cases must still be numerically evaluated to observe the characteristics. However, with the closed form solution, it is easier to evaluate the effects of parameter variation.

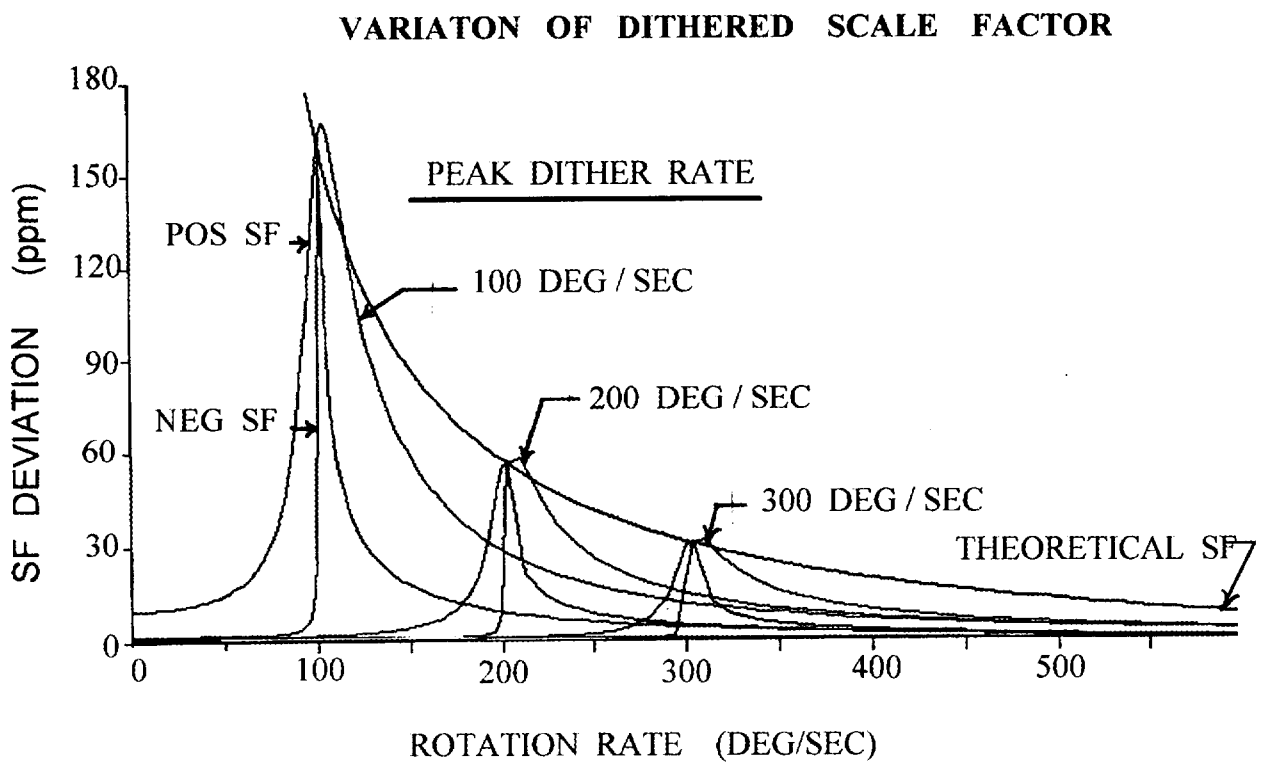


Figure 6.1. These theoretical plots show the scale factor deviation from an arbitrary nominal, as obtained from Eq (6.4), for three peak dither rates. The parameters have been adjusted, as per the text, so as to overlap the positive and negative SF corrections. The negative SF correction shows a smaller width in rotation space, than the positive correction, agreeing with experimental results. The calculated expression, as given by Eq (6.9), for the peak SF correction when the rotation rate is equal to the dither rate is also shown. Note that it agrees well with the values of the peaks shown on the plot.

Figure 6.1 shows a plot of the dithered scale factor correction, expressed in parts per million, for the case of Ω_G = 10 deg/sec, Ω_{L2} = 0.1 deg/sec, Ω_{L1} = 0.74 deg/sec and for dither rates of 100, 200, 300 deg/sec. Both the positive and negative contributions are plotted on the same scale. The negative term has been multiplied by -1, so that it can be scaled relative to the positive term. The value of Ω_{L1} was chosen so that the maximum values of the two contributions are approximately equal.

The results show that the scale factor correction peaks at the dither rate. This is reasonable in that when the rotation is equal to the peak dither rate, the gyro is passing through the lock-in band during the turn around and thus spends maximum time in this region. The maximum non-linearity is observed to decrease with increasing dither rate. The width of the non-linearity peak is much narrower for the negative contribution, due to the half width of the Lorentzian shape being determined by the lock-in, as opposed to the gain rate in the positive term. At low rates the scale factor is predicted to asymptote to a constant value, in agreement with experimental results. At high rates the correction tends to fall off more slowly and becomes independent of dither rate. The major characteristics of Figure 6.1 can be quantified by considering special cases for the scale factor correction, as given by Eq (6.4). For the case of zero rate and for the usual case of $\Omega_D >> \Omega_G$, Eq (6.4) reduces to

(6.7) $S_+ = \Omega_{L2}\,\Omega_G^2 / \Omega_D^3 \qquad S_- = -\frac{1}{2}\,\Omega_{L1}^3 / \Omega_D^3$

Since $\Omega_D >> \Omega_G$ and Ω_L, the SF correction at zero input rate is small. Since $\Omega_G >> \Omega_L$, the negative SF term shows a much smaller correction at zero input rate.

At $\Omega >> \Omega_D$, Eq (6.4) simplifies to

(6.8) $S_+ = \Omega_{L2}\,\Omega_G / \Omega^2$ $\qquad S_- = -\frac{1}{2}\,\Omega_{L1}^2 / \Omega^2$

Equation (6.8) shows that both terms follow the non-dithered inverse-square-of-rate scale factor correction. These characteristics are shown in Figure 6.1. For the case of the input rate being equal to the dither rate, Eq (6.4) reduces to

(6.9) $S_+ = \frac{1}{2}\,\Omega_{L2}\,\Omega_G^{1/2} / \Omega_D^{3/2}$ $\qquad S_- = -\frac{1}{4}(\Omega_{L1} / \Omega_D)^{3/2}$

Equation (6.9) is plotted as an overlay in Figure 6.1 for the parameters in the paragraph after Eq (6.6), and shows excellent agreement with the exact solution given by Eq (6.4). For equal lock-in for the positive and negative contributions, Eq (6.9) shows the SF non-linearity correction to be smaller for the negative contribution than for the positive contribution.

3.7 MECHANICAL DITHER

3.7.1 Introduction

The analysis describing the effects of mechanical dither was developed [66-67] soon after demonstration of the feasibility of the concept. This dither approach was used to avoid operation of the RLG in the lock-in band. It was initially thought that by rapidly moving the gyro through the lock-in band, the time constant of the system would keep the gyro from locking. It was recognized that a phase error would occur each time the gyro traversed the lock-in band, but that due to natural instabilities, the phase errors would occur in a random fashion, and hence not build up. However, with the inherent stability of a RLG fabricated from a solid glass frame that has no optical elements in the cavity, this phase instability was not sufficient to avoid error buildup that rectified in time. The effect was that that when the gyro was operated with a sinusoidal input rate, and for certain ranges of constant input rates, the phase pattern was found to follow the dither input rate and not be responsive to the inertial rates. This effective lock-in, sometimes called "dynamic lock-in", was caused by the existence of multiple steady-state solutions to the lock-in equation and the gyro rapidly find these solutions[68].

To avoid the gyro stabilizing onto a steady-state solution, it was found that on a dither cycle basis, it was sufficient to vary the initial phase condition[69]. If this was done on a truly random basis, no rectification, or introduction of a bias, occurred. The randomization was performed by the introduction of noise onto the dither amplitude. The cost of avoiding the dynamic lock-in was that randomization introduced a random drift to the output of the gyro.

In that mechanical dither is commonly considered the baseline approach to the avoidance of the lock-in problem, details of the randomized mechanical dither will be presented. Cases considered will be sinusoidal dither, square wave dither, and a constant rotation that periodically reverses after a fixed number of revolutions. For the analysis to be applicable to actual gyros, it will be necessary to numerically evaluate the results, so careful attention will be paid to proper units. This is especially important in that tradition has evolved mixed units.

3.7.2 Mechanical Dither With Sinusoidal Motion

The equation for the beat frequency in the presence of coupling has been described in section 3.4.5. To consider the effects of mechanical dither, the inertial input rate will be considered to consist of a constant input rate plus a sinusoidally varying rate. The beat frequency equation given by Eq (4.4) can be simplified by ignoring the frequency pulling and pushing terms. The coupling (lock-in) term can be combined in terms of amplitude and phase terms. Then Eq (4.4) can be written as

(7.1) $(d\psi/dt) = (2\pi / S_k)\,[\Omega_{Rate} + \Omega_D \sin \omega_d t + \Omega_L \sin(\psi + \chi)$

where ψ is the instantaneous phase difference between the two beams in radians; $\mathbf{d\psi/dt}$ is the beat frequency, or gyro output in rad/sec; S_k is the gyro scale factor in arc-sec/count; Ω_{Rate} is the constant input rate, Ω_D is the peak dither rate and Ω_L is the lock-in threshold, all in deg/hr (arc-sec/sec); ω_d is the dither frequency in rad/sec. It should be noted that it is assumed that the output consists of one count for each cycle of fringe motion. An output consisting of a larger number of counts per fringe cycle could be implemented[20].

With the implementation of mechanical dither, what is of interest is the calculation of the error angle buildup in time. Hence the fringe motion is separated into inertial motion (ψ_a) and error motion (ψ_e). The equations for the inertial motion and fringe error are

(7.2) $(d\psi_a/dt) = (2\pi / S_k)\,[\,\Omega + \Omega_D \sin \omega_d t\,]$

(7.3) $(d\psi_e/dt) = (2\pi / S_k)\,[\,\Omega_L \sin(\psi_a + \psi_e + \chi\,)$

The fringe error equation is deceptively simple. No exact closed from solution has been found. Computer solutions have been obtained, which give guidance to simplifying approximations that allow determination of closed form solutions. The major simplification is the recognition that the phase error buildup per dither cycle is small compared to the inertial angle. Hence the phase error angle can be neglected in Eq (7.3). In the integration of Eq (7.2) for the determination of the inertial angle, the contribution from the constant input rate is usually small, and can be neglected. The solution, with ψ_o being an initial condition, is

(7.4) $\psi_a(t) = \psi_o + [\,2\pi\,\Omega_D / (\omega_d S_k)]\,[1 - \cos \omega_d t\,]$

Computer analysis has shown that the phase error occurs as the gyro goes through the lock-in region, which occurs during the turnaround in the dither cycle. Then the sine wave dither can be approximated as a parabola and Eq (7.4) can be written as

(7.5) $\psi_a(t) = \psi_o + \pi\,\omega_d\,\Omega_D\,t^2 / S_k$

In integrating Eq (7.3), the limits can be taken to infinity, since the error contribution occurs close to zero. Then using Eq (7.5), the phase error per dither turnaround $\Delta\psi_e$ is obtained from Eq (7.2) as

(7.6) $\Delta\psi_e = [4\pi\,\Omega_L^2 / (\,S_k\,\omega_d\,\Omega_D\,)\,]^{1/2} \int_{-\infty}^{+\infty} \sin(x^2 + \alpha)\,dx \qquad x = [\,\pi\,\omega_d\,\Omega_D\,/\,S_k\,]^{1/2}\,t$

where α is the combined laser and dither phase angle. Equation (7.6) integrates to

(7.7) $\Delta\psi_e = \pi\,\Omega_L\,[\,2 / (\,S_k\,\omega_d\,\Omega_D\,)\,]^{1/2}\,(\sin\alpha + \cos\alpha\,)$

Equation (7.7) shows the phase error per dither turnaround to be a sinusoidal function of the arbitrary phase angle, with a peak value

(7.8) $\Delta\psi_m = 2\pi\,\Omega_L / [\,S_k\,\omega_d\,\Omega_D\,]^{1/2}$

3.7.3 Concept Of Dynamic Lock-In

The above result shows what was once an unexpected result in the initial design of the RLG. It shows that if the gyro is operated with a pure sinusoidal dither, it is possible for the gyro to remain dynamically locked. By this, it is meant that the gyro fringe pattern follows the mechanical dither by sinusoidally oscillating and is not responsive to the input rate. This can be seen more clearly, by considering Eq (7.1), which describes the total fringe motion. Since the dither is sinusoidal, it gives no net contribution per half dither cycle. The input rate contribution per half dither cycle ($T_d/2$) is

(7.9) $\Delta\psi_a = (\,2\pi / S_k\,)\,\Omega\,(T_d / 2)$

The error contribution from the lock-in has been calculated as Eq (7.7). The sum of the inertial and error phase angle accumulation per dither turnaround is approximately shown in Fig 7.1, as a function of the arbitrary dither-laser phase. It is approximately a sinusoidal function of the arbitrary dither-laser phase angle, with an average accumulation due to the input rate. For an arbitrary initial phase angle, there is some total phase accumulated over the half dither cycle. This total phase can be considered as a new initial phase condition for the next half dither cycle. Since Fig 7.1 shows the existence of zero-crossings, it is possible that no accumulated phase will occur at some arbitrary dither cycle. When this happens, the new initial phase angle for the next half dither will be the old initial phase angle. Hence no accumulated output phase will occur from this time on. All that happens is that the fringe pattern phase oscillates with the mechanical dither and is not responsive to the input rate. This result has been shown[66] with computer solutions of Eq (7.1) and verified experimentally[66]. Unlike what one might expect, it does not take many dither cycles before the gyro finds the "steady-state" solution. The necessary condition for steady state solutions is the existence of a zero crossing. If the input rate was of a magnitude so that the input rate accumulated phase per half dither cycle was greater than the maximum phase error, no dynamic lock-in would occur. From Eq (7.8), it can be seen that to minimize the dynamic lock-in threshold, it is desirable to minimize the lock-in threshold and maximize the product of the peak dither rate and dither frequency.

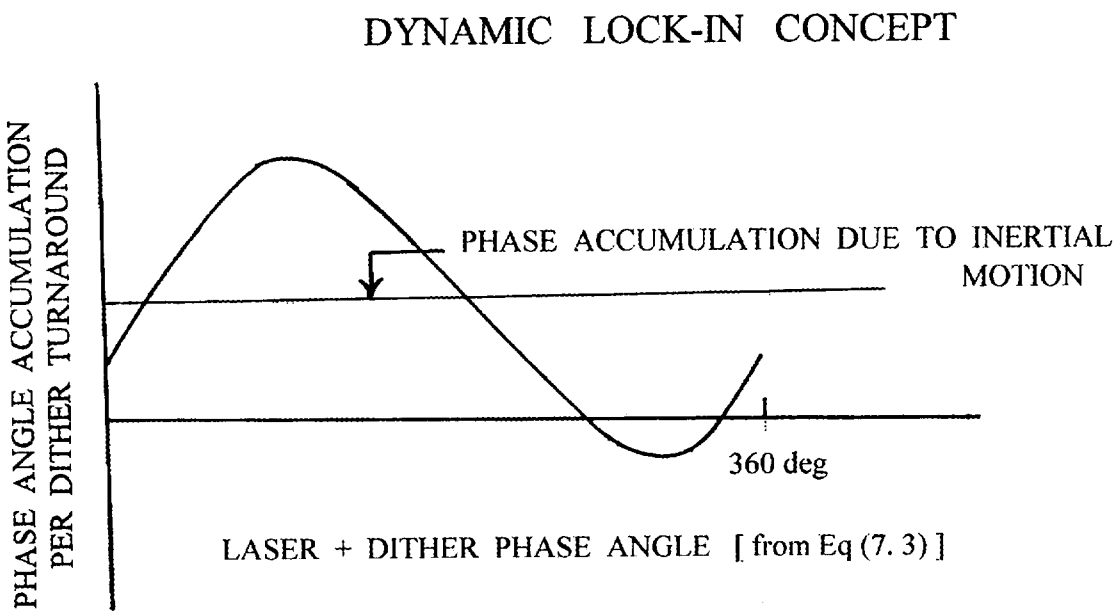


Figure 7.1. This plot shows the justification for the dynamic lock-in concept. The sine curve represents the phase accumulation per dither cycle for pure sinusoidal dither, showing a periodicity of 2π. It is superimposed on the horizontal line, which represents the phase accumulation per dither cycle due to a constant inertial rate. The inertial rate phase is independent of the phase angle.

By equating Eqs (7.8) and (7.9), and writing the peak dither amplitude θ_D as Ω_D / ω_d, the dynamic lock-in threshold can be found as

(7.10) $\Omega_L^* = [\, S_k / (\pi^2 \Omega_D)\,]^{1/2}\, \Omega_L$

Since the scale factor units are arc-sec/count and the dither amplitude units are arc-sec, the dynamic lock-in threshold can be express as

(7.11) $\Omega_L^* = \Omega_L / [\pi \sqrt{C}\,]$

where **C** is the dither amplitude in counts. In that the peak dither amplitude is typically on the order of hundreds of counts, a lock-in reduction of 50 can be readily obtained with pure sinusoidal dither. Since lock-in thresholds can be on the order of 0.1 deg/sec, this reduction still gives a residual lock-in threshold on the order of earth's rate. For most applications, a lock-in threshold of this magnitude is still unacceptably high.

With pure sinusoidal dither, another surprising situation occurs. As can be seen from Fig 7.1, for large input rates, it is possible for the accumulated phase to cross the 2π value. When this occurs, another steady state solution is found to exist. In this case however, each dither crossing results in exactly one output count. As the input rate continues to increase, new dynamic lock-in regions are found, with integral multiple of counts per half dither cycle. Thus, by attempting to avoid the lock-in region by introducing pure sinusoidal dither, the input output characteristic is change from the single lock-in to a series of reduced magnitude lock-ins at increasing input rates. In terms of the dither frequency $\mathbf{f_d} = 1 / \mathbf{T_d}$, the magnitude of the rotation rate for the higher order lock-in bands are

(7.12) $\Omega_N = S_k\, f_d\, N \qquad N = \pm 1, \pm 2, \ldots\ldots\ldots$

For a typical dither frequency of 500 Hz, and a scale factor of 2 arc-sec/count, Eq (7.12) predicts lock-in bands every 0.3 deg/sec.

3.7.4 Random Walk Noise In The Output Of The Ring Laser Gyro

The technique for the elimination of the multiple lock-in bands follows from the description of the dynamic lock-in concept. As can be seen from Fig 7.1, the phase accumulation from the true inertial input rate is just the total phase, averaged over all possible laser-dither phase angles. Thus, all that is needed is the randomization of the dither phase angle, so that over successive dither cycles, the laser never locks onto the steady-state solution. As seen from Eqs (7.2) and (7.3), this can be implemented by adding amplitude noise to the pure sine dither, which is equivalent to phase noise. If the noise is not fully randomized, the average phase accumulated per dither crossing will not be exactly that arising from the true inertial input. This would give rise to a rectification, or bias error. The required noise amplitude is one count per dither cycle. Since mechanically dithered gyros are traditionally high Q devices, it is not a trivial matter to provide sufficient noise to achieve pure randomization. A simple check as to whether the dither noise is sufficient, is to start at low noise level and increase the noise amplitude to watch the bias asymptote to a constant value.

Since there is no distinction between the dither phase angle and the arbitrary laser phase, it might be thought that inherent instabilities in the laser would be sufficient to eliminate the dynamic lock-in. However, with the solid block laser cavity construction and no optical elements in the cavity, the laser phase has been found to be highly stable and externally added noise is required. It should be noted that a high vibration-type environment provides input noise. Although added noise does not provide any improved bias stability benefit, gyros designed with marginal noise tend to operate with improved performance in a high dynamic environment.

The detriment to adding random noise to the gyro dither amplitude is that a random phase error occurs in the output of the gyro. The magnitude of the phase error can be calculated by recognizing that all values of dither phase are equally probable and the RMS phase error per dither crossing is $1/\sqrt{2}$ times the maximum phase error. The growth of the phase error is the RMS phase error times the square root of the number of dither crossing in a time **t**, times $\mathbf{S_k} / 2\pi$ (to express phase in the inertial unit of arc-sec), or

$$(7.13) \quad \Delta\psi_e(t) = (S_k / 2\pi)(1/\sqrt{2})[2\pi \Omega_L / (K \omega_d \Omega_D)^{1/2}][2 f_d t]^{1/2}$$

The diffusion coefficient of Eq (7.13), with units of arc-sec / sec $^{1/2}$, is

$$(7.14) \quad W = [S_k / (2\pi \Omega_D)]^{1/2} \Omega_L$$

Recall that with reference to Eq (7.1), units of Ω_L and Ω_D are arc-sec/sec or deg/hr. It is customary to express the angular random walk coefficient in units of deg / hr $^{1/2}$, so Eq (7.14) must be divided by 60.

As an illustration of the usage of the angular random walk expression as given by Eq (7.14), consider a square RLG with a 20 cm perimeter, that is mechanically dithered. Based on Eq (2.7), the scale factor is 2.6 arc-sec/count. A typical lock-in for this type gyro is 0.05 deg/sec, or 180 deg/hr. A reasonable dither design might be a dither frequency of 500 Hz, with a peak amplitude of 200 arc-sec. This corresponds to a peak dither rate of 6.28×10^5 deg/hr, or 174 deg/sec. Using Eq (7.14) with the rates expressed in deg/hr and using the factor of 1/60 to convert from deg/sec $^{1/2}$ to deg/hr $^{1/2}$, the value of 0.0020 deg/hr $^{1/2}$ is found. This is typical of values required for 1 nmph systems with a 2 minute alignment time[17].

If pure sine dither was used, the lock-in reduction factor, as given by Eq (7.11) is 34. With a lock-in value of 180 deg/hr, the reduced lock-in becomes 5 deg/hr. This is still a factor of 103 greater than what would be required for aircraft navigation applications[17], indicating the need for randomized dither. A summary of the effects of mechanical dither is shown in Table 7.1. Because of coupling and lock-in, the RLG has a rate threshold with no random noise in the output.

In summary, all ring laser gyros have demonstrated a minimum rate threshold, below which the gyro is not responsive to input rate motion. To avoid this lock-in threshold, some means must be employed. The concept of mechanical dither, which used a rapidly varying and alternating input rate was one of the early successes, leading to production of the ring laser gyro for commercial aircraft applications. However, with the addition of a pure sinusoidal dither, it was found that although the rate threshold was reduced, it was not reduced to zero. Hence a residual lock-in threshold still existed. To reduce the lock-in threshold to zero, it was found that it was necessary to add randomization to the dither amplitude. If the randomization was not of sufficient magnitude or if the randomization was not pure "white", a bias instability occurred. However, with the addition of randomization to the amplitude of the dither, an angular white noise was found in the output of the gyro. Thus with the randomized mechanical dither technique, the lock-in problem was overcome to the extent that the RLG was able to become a useful commercial instrument. However, the price was added cost due to added complexity in the design of the gyro, noise picked up by the accelerometers and noise in the output of the gyros which is a performance limitation in certain applications..

Table 7.1 EFFECT OF MECHANICAL DITHER ON THRESHOLD AND NOISE

TYPE OF DITHER	RATE THRESHOLD	OUTPUT NOISE
None	Ω_L	0
Sinusoidal	$\frac{\Omega_L}{\pi\sqrt{C}}$	0
Randomized Sinusoidal	0	$\Omega_L \sqrt{\frac{S_k}{2\pi\,\Omega_D}}$

3.7.5 Lock-In Avoidance Using An Alternating Rate Bias

The lock-in error, occurring as the gyro is dithered through the lock-in band, has been found to decrease with minimal time spent in the lock-in region. It might be thought that with a non-mechanical square wave type dither[70], sometimes called "Maytag" after the washing machine, with which the gyro spends zero time in the lock-in region, there would be no phase error buildup. As will be shown, although the error can be minimized, an error buildup will still occur. It arises due to a phase change in the output of the gyro as the gyro passes through the lock-in region.

For square wave dither, the approach used in treating sine wave dither analysis can be followed. Over a lock-in region for one-half dither cycle (dither period of $\mathbf{T_d}$), the inertial dither phase can be written (again neglecting the small contribution from the input phase) as

(7.15) $\psi_a = (2\pi / S_k)\,\Omega_D\, t \qquad T_d/4 \le t \le T_d/2$

(7.16) $\psi_a = (2\pi / S_k)\,\Omega_D\,(T_d - t) \qquad T_d/2 \le t \le 3T_d/4$

Using Eqs (7.15-16), the phase error angle, given by Eq (7.3) can be integrated over the above limits to give

(7.17) $\psi_a = 4\,(\Omega_L / \Omega_D)\,[\sin[(2\pi / S_k)\,\Omega_D\, T_d/8]]\,[\sin[\alpha + 3\,(2\pi / S_k)\,\Omega_D\, T_d/8]]$

where α is a combined dither-laser phase angle.

The analysis previously used for sine wave dither can be carried over, resulting in the angular random walk coefficient (units of deg/sec $^{1/2}$) as

(7.18) $W = (S_k\,\Omega_L / \pi)\,[2/(\pi\,\Omega_D\,\theta_D)]^{1/2}$

Again, units of Ω_D and Ω_L are deg/hr, the dither amplitude θ_D is arc-sec and the scale factor unit is arc-sec/count.

The advantage of square to sine dither for equal scale factor, can be determined by considering equal dither amplitude and frequency and dividing Eq (7.18) by Eq (7.14), to give the random walk coefficient reduction as

(7.19) $W_{square} / W_{sine} = (2/\pi) / \sqrt{C}$

where C is the dither amplitude in counts. Larger improvements than that shown in Eq (7.19) can usually be obtained for square wave dither, since for the non-mechanical cases where it is used, much higher dither parameters can be obtained.

In the description of sine-wave dither, the phase error is dependent on the number of zero crossings of the dither. In an attempt to reduce the number of crossings, a technique has been developed where the gyro is rotated relatively slowly (tens of deg/sec) over multiple revolutions[18]. Then to cancel any buildup of rotation induced bias instability, the direction of rotation is quickly

reversed, and the rotation cycle repeated. The tradeoff with respect to longer times before reversal is lack of exact bias error cancellation vs. higher random drift with more traversals through the lock-in region.

Constant rotation with reversals has sometimes been called slow-square-wave-dither, in that the phase error analysis follow that of the standard dither. The terms carousel and Maytag (after the washing machine) dither have been applied to it. In that mechanical systems cannot be reversed in zero time, the turnaround can again be approximated as a parabola in time, as in Eq (7.5). Then the maximum phase error per dither turnaround is given by Eq (7.8), where the product $\omega_d \Omega_D$ can be represented as an angular acceleration. Again the phase error is divided by $2^{1/2}$ to convert to RMS and multiplied by the square root of the number of zero crossing. For the gyro being turned at a constant rate Ω_{rate} and reversing after a number (**N**) of revolutions, the number of reversal in a time **t** is the total time **t** divided by the time per 1/2 cycle. The time per 1/2 cycle is 360 **N** / Ω_{rate}, where the rotation rate is expressed in deg/sec. To change the angular acceleration (A_c) from arc-sec/sec^2 to deg/sec^2, a factor of 3600 is included in the denominator of the square root. Then the angular random walk (units of arc-sec/sec«) can be found as

$$(7.20) \qquad W_{Rate} = \Omega_L [\, 5\, K\, \Omega_{Rate} / (\, A_c N\,)\,]^{1/2}$$

where Ω_L and Ω_{Rate} are in deg/sec, the scale factor is arc-sec/count and the angular acceleration is deg/sec^2. Equation (7.20) indicates that with a fast turn-around, low rotation rates and a large number of revolutions before turn around, are desirable. The number of revolutions before a turn-around must be consistent with gyro stability.

Using the previous example of a 20 cm square gyro with a scale factor of 2.6 arc-sec/count, a lock-in of 0.05 deg/sec, a rotation of 30 deg/sec, an angular acceleration of 100 deg/sec^2 and with 10 revolutions between turnaround, Eq (7.20) gives an angular random walk of 0.031 deg/sec$^{1/2}$, or 0.00052 deg/hr$^{1/2}$. This is a factor of 4 less than the value found using the mechanical dither technique. A tradeoff must be made as to whether the improved factor is worth the added size, weight, complexity and more sensitivity to an environment with high levels of shock. The scaling relation between the two approaches can be found from Eqs (7.14,20) as

$$(7.21) \qquad W_{Rate} / W_{Dither} = (1/60)\, [\, 10\, \pi\, \Omega_{Rate}\, \Omega_D / (\, A_c\, N\,)]^{1/2}$$

where rotation rate terms have units of deg/sec and the angular acceleration deg/sec^2. Note that the improvement is independent of the gyro parameters of lock-in and scale factor.

3.8 LOCK-IN CIRCUMVENTION

Because of the magnitude of the lock-in band, some technique must be implemented to allow the usage of the RLG for a practical device. In the early years of development of the RLG, many approaches were pursued. Because of the high sensitivity of the RLG to differential index of refraction effects, production capable approaches were difficult to implement. Mechanical dither, as described in the previous section, which provided an almost bias-free compensation, was the technique that ultimately dominated. This section describes alternate approaches that were attempted, one that is currently in use, and the various techniques that are being used to implement mechanical dither.

3.8.1 Spacial Separation Of The Two Beams In The Cavity

In that coupling was described as being the source of lock-in, a number of approaches have been proposed which reduce, or avoid the coupling. One that has never been implemented is the spatial separation of the two beams so that scattering from one beam into the other is not possible. The problem with this approach is that each beam has a different cavity length. Thus it is not possible to common mode cavity length fluctuations to the accuracy required to allow the measurement of the Sagnac effect for practical rotation rate values.

3.8.2 Phase-Locked Laser Gyros

Another approach to the avoidance of coupling and still being able to maintain identical cavity lengths, is to phase lock the two beams with an intra-cavity loss modulation[71]. This results in the beams being pulsed with a width much smaller than the cavity length. The laser can be designed such that the pulses will not overlap in the region of the mirrors, modulator, and the active medium. Thus, this should result in no coupling. Experiments demonstrated improvement over the free running case. However, lock-in threshold with the pulsed RLG was still high, compared to values obtained with comparable sized ring lasers with no optical elements in the cavity. With this approach, it still remains to be seen whether a small, low cost, practical RLG can be achieved.

3.8.3 Scattering Reduction

A traditional approach to avoid lock-in, is the attempt to reduce coupling by mirror improvement. With current mirror deposition technology, mirrors are extremely low loss and low scattering. Total absorption and scattering loss can be readily held to 0.006% (60 ppm). Using the expression from Table 5.1 as an estimate for lock-in, the expression for the effective scattering rate is

$$(8.1) \qquad \Omega_R = (\tfrac{1}{2}\Omega_L\,\Omega_G)^{1/2}$$

With a measured lock-in threshold of 0.1 deg/sec, a scale factor of 1380 count/deg (20 cm square perimeter), a gain rate (Ω_G) of 10 deg/sec, Eq (8.1) gives a value for the scattering rate as 0.7 deg/sec. This corresponds to a frequency of 970 Hz. Using,

$$(8.2) \qquad \Omega_R = R\,c/(2\pi L)$$

the effective scattering **R** is found as 4×10^{-6}. Power scattering is the square of the above value, resulting in an extremely small value. Reduction in scattering will cause performance improvement, but it is not expected that reduction alone will ever eliminate the need for some form of lock-in avoidance.

Techniques have been demonstrated[61] which make use of the vector properties of the scattering. Certain of the mirrors are moved along the direction of the normal, to allow the beam to strike the surface at different positions. This varies the spacing between the mirrors and allows the vector sum of all the scattering to result in a minimum lock-in. Zero is typically not obtained. The mirror position is actively controlled by means of monitoring the effect of the scattering on the amplitudes of the beams. Although this technique can be implemented, there are problems. In controlling the mirror position, the sense of the discriminant is determined by whether the lock-in is scattering or loss dominated, making production implementation more difficult. It is also difficult to obtain much improvement with gyros that initially have low lock-in, and the added cost and complexity make the technique questionable. The most improvement comes from gyros that have large lock-in, which, with today's technology are usually small sized gyros. Typically, smaller gyros are associated with low-cost applications, where it is more difficult to justify the added cost of the implementation of the lock-in reduction. However, there may be application where higher performance is required from a smaller sized gyro, and the technique may be practical.

3.8.4 Environmental Lock-In Compensation

The most basic approach, which was successfully tested for an anti-ballistic missile application[72], is one where the RLG is used in an extremely severe environment. Then the resultant vibration induced upon the RLG by the vehicle to which it is mounted, provides a mechanical noisy dither that allows the gyro to operate unlocked. This implementation requires a structurally rigid mount for the mirrors, such as a glass frame, into which the holes are drilled for the three independent gyros. One problem is that there are not many applications that provide the required vibration dither. Another is the difficulty of testing, without some other form of lock-in avoidance. A criticism once raised with respect to this type configuration, is the fact that two good gyros are lost, if one is not acceptable. However, with present RLG manufacturing technology, gyro yield is high, and there are cost efficiencies with the usage of a common laser frame. It should be noted that the input axis alignment stability between the three gyros, which is usually a major error driver, is excellent for the common laser frame approach. However, care must be taken in the gyro design to avoid structural change during gyro turn-on due to heat dissipation by the single localized cathode. This structural change will affect the gas flow with a resultant bias transient. Effects can be large.

3.8.5 Common Mechanical Dither Mechanization

An approach which avoids the three-in-one issue, and economizes on the cost of three individual mechanical dither devices, is to mount the three individual gyros on a support that is part of the mechanical dither mechanism[73]. In this mechanization all three gyros can be mechanically dithers with the cost of only one dither mechanism. In that the dither motion is a component of the actual motion along the three input axis of the gyro, there will be reduced performance over an approach in which full dither is applied to each axis. This approach has been taken for small (6 cm path length), low-cost/lower-performance gyros[29].

3.8.6 Common Mechanical Dither With Three-axis Gyro

Another approach that uses a common mechanical dither, also uses the three-in-one design, with a novel modification[74]. The frame, into which the holes are drilled for the three gyros, is of a cubic configuration. For cost efficiencies, the cavity is of a square configuration, with mirrors mounted on the faces of the cube. Thus, each mirror supports two of the gyros. In this way, instead of having twelve mirrors for the three gyros, there are only six. Of the six mirrors, three are path length control mirrors, and each is used to control the path length of two gyros. There is further efficiency in that frame makes use of a common cathode, that supports parallel gas discharges for each of the gyros. This approach allows the usage of a larger sized gyro for a given overall volume, than could be used if three individual gyros were used. In that design margin then exists, other tradeoffs that tend to reduce performance can be made to reduce cost.

3.8.7 DC Bias

As with mechanical dither, lock-in avoidance approaches that have been most useful, are those in which the oscillating frequencies are biased away from the lock-in region. A configuration which uses only a simple dc bias, such as a constant rotation, tends not to be successful. The problem is lack of stability of the dc bias. For typical lock-in values, the bias must be in the range of hundreds of deg/sec to maintain scale factor stability over the dynamic range of hundreds of deg/sec, which is not consistent with bias stability requirements of 0.1 to 0.001 deg/hr. Experiments have used a wide variety of techniques to produce the bias, including physical rotation[75], insertion of passive elements to produce a differential path (spinning wheels, flowing gases and liquids)[76], gas flow of the active gain medium[26,77], and magneto-optic elements[55,70,78].

Major problems must be addressed with the insertion of optical elements in the cavity. One is strain induced birefringence, which leads to large magnetic sensitivity and increased thermal sensitivity. Another is the additional scattering, which tends to cause increased lock-in. These additional elements always cause increased loss, which causes an increase in the fundamental-limit angular random walk. An increase in part count typically leads to an increase in production cost.

To avoid the stringent stability on the dc bias, there have been implementations in which the three orthogonal gyros have been mounted on a housing that provides a common constant input rate for all three gyros[18]. After a fixed number ($\approx$10-20) of rotations, the direction is quickly reversed, which tends to reduce the buildup of bias-instability induced errors. This slow dither also minimizes the number of traversals per unit time, through the lock-in region, resulting in a lower angular random walk. Excellent continuous performance (0.000X deg/hr) for shipboard application, has been achieved with this implementation. The tradeoff is the requirement for a gimbal and a larger than desired mounting configuration.

3.8.8 Faraday Cell Bias

In the early development of the RLG, lock-in avoidance techniques were investigated that avoided the use of mechanical dither. Although the mechanical dither provides a bias-free approach, it has negative features. Its implementation requires a mounting that has higher cost than desired. The presence of mechanical motion imparts a reaction torque on the mount that has an adverse effect on the accelerometers. Because of the added randomization that is required to avoid the dynamic lock-in effects, a random noise is produced on the gyro output. Lastly, the usage of a mechanical dither motion on this highly elegant and simple strapdown gyro goes against the esthetic values of the designers. It was felt that a magneto-optic approach was more compatible with the usage of the Sagnac effect.

To obtain a non-reciprocal effect, the Faraday cell[55] concept was used. The cell consisted of an magnetic field-driven optically active material with the property that it has a different index of refraction for light traveling in opposite directions. This index difference, which results in a phase difference, is equivalent to the phase difference caused by the Sagnac effect. It causes a frequency splitting that can be used to bias the gyro without the need for any inertial rotation. The differential phase is proportional to the magnetic field and the length of the cell. The proportionality constant, known as the Verdet constant, is typically on the order of 0.1 arc-min/gauss/cm. With reasonable magnetic field strengths, the Faraday element can give rise to the required bias of hundeds of deg/sec.

Unfortunately, the implementation of the Faraday element was not simple. The index of refraction effect requires the presence of circularly polarized light. The ring laser cavity tends to support linearly s-polarized light. The choice of the polarization is determined by the higher reflectivity of the mirrors for s-polarization, than p-polarization. Thus it is necessary to place two 1/4 wave plates, with the slow and fast axis crossed, at the ends of the Faraday element. This changes the linearly polarized light to circularly polarized light for transversal through the Faraday element, and back again to linearly polarized light for transversal through the laser cavity. The crossing of the 1/4 wave plates is required to adjust the sense of the polarization such that an index difference occurs. With the improper alignment of the 1/4 wave plates, a differential loss can occur with resultant bias instability. Also magnetic sensitivity tends to be large, requiring a high degree of shielding.

The hundreds of deg/sec bias must be removed to allow measurement of the inertial rates. Because the Verdet constant is temperature dependent, large thermal sensitivities tend to occur. In an attempt to minimize these problems and avoid buildup of bias errors, it was usual to operate the Faraday cell with an alternating magnetic field. This required synchronization of the output to pairs of magnetic cycles. Unlike mechanical dither, where no rectification can occur, the Faraday cell was not necessarily symmetric to the degree required for navigation applications.

Using the FM concept[79] for analysis of the lock-in equation, it was shown that the lock-in equation can be expressed as a series of terms with Bessel function coefficients. Thus by adjusting the amplitude of the Faraday bias, it was thought that it would be possible to eliminate the lowest (or two lowest with a dual frequency modulation) lock-in contributions. Unfortunately, as discussed in the section on dynamic lock-in, other lock-in regions still existed at higher rotation rates. Also, it was not possible to control the stability of the Faraday bias, to account for changes in lock-in threshold. Although the Faraday Cell technique appeared at first sight to offer advantages over the mechanical dither approach, the presence of many problems made it difficult to implement for a production device.

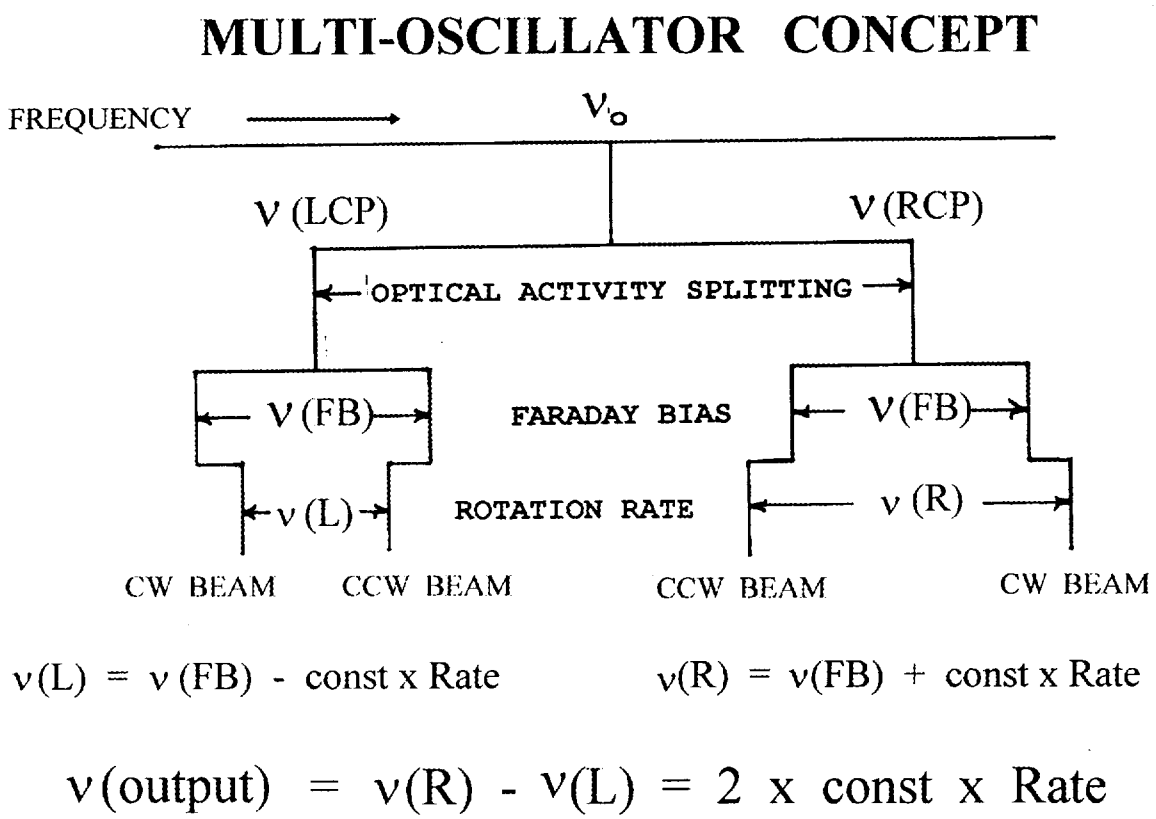


Figure 8.1 This figure illustrates the multi-oscillator concept showing the optical activity splitting of the cavity frequency ν_o, into right an left circular polarized frequencies. The Faraday bias splits the counter clockwise and clockwise components of each. The rotation rate further adds a splitting contribution, the sense of which is shown in the figure. The readout of the different polarized beams are performed independently of each other. These outputs are then subtracted from one another, resulting in a final output that is independent of both the optical activity and Faraday splitting, and only dependent on the rotation splitting.

As problems in producing a Faraday Cell compensation surfaced, a very clever technique to provide automatic compensation of the instabilities was proposed[80-84]. The laser was operated with two separate sets (clockwise and counter clockwise beams) of laser frequencies, each used to measure input rate and both simultaneously unlocked by a common biasing mechanism. Then each set was used to obtain an output consisting of the common bias and the rotation rate. By obtaining the difference between the two outputs, the common bias cancels and the final output consists only of the rotation rate. This approach is described in much more detail in section 4 in this monograph.

The approach is described using Figure 8.1. The cavity frequency is represented by ν_o. By means of an optically active material, such as quartz, a differential phase can be introduced with respect to the left and right circularly polarized (LCP and RCP) beams. The phase introduced by this element is reciprocal with respect to direction. Thus the oscillating frequencies in the cavity are shifted and are expressed as ν_{LCP} and ν_{RCP}. In that no non-reciprocal element has been introduced, each polarization consists of a clockwise and counter clockwise beam, locked to a common frequency. At this stage, a non-reciprocal component, such as a Faraday element, is introduced. Now the degeneracy in direction in each polarization is removed. However, because of the non-reciprocal nature of the Faraday effect with respect to polarization, the RCP and the LCP beams are split in frequency so that the clockwise and counter clockwise beams move in opposite directions, as described in Figure 8.1. Now the beams are all unlocked and hence responsive to the Sagnac effect. For a constant rotation, each beam traveling in the same direction, which is of different polarization, is shifted in the same direction. Hence the Sagnac effect causes an increase and decrease in splitting, respectively, for each polarization. The shifting strategy is shown in Fig 8.1.

To obtain a readout, the outputs are measured for each polarization separately and recorded in the conventional manner. This is done by using polarization discriminating elements, such as 1/4 wave plates with polarizers. Each output contains the non-

reciprocal Faraday splitting, plus the Sagnac effect in addition and subtraction. Thus by subtraction of the two outputs, the Faraday effect can be canceled. The Sagnac effect remains with a factor of two improvement in scale factor.

Note that the outputs are independent of the reciprocal optical activity splitting. It should be noted that in addition to providing the mechanism for circular polarization, the optical activity splitting also frequency separates the beams that have orthogonal polarization and traveling in the same direction. Otherwise coupling would occur with a resultant frequency synchronization into a linear polarized beam.

This technique has been variously described as a four-frequency RLG, multi-oscillator RLG, differential (DILAG) RLG and zero-lock RLG (ZLG). The technique was introduced in the late 1960's and implemented in the early 1970's by companies that were not successful in solving the difficult technical problems. The last nomenclature is used by Litton, which has been successful on overcoming the technical difficulties and is producing gyros using this approach. The analysis of the multi-oscillator RLG in its general form is extremely complex, has previously been reviewed[12-13], and is covered in more detail in section 4 in this monograph.

By operating the multi-oscillator with the Faraday bias, the gyro is well away from the lock-in region for both polarizations. Hence it is not necessary to reverse the Faraday bias and traversal through the lock-in region is avoided. This avoids the angular random walk error introduced by mechanical dither. For applications where the inertial rate approaches the Faraday bias, one of the polarization signals can be used in lieu of the other. In this case, the Faraday bias must be removed electronically. By operating well away from the lock-in region, scale factor non-linearity arising from lock-in is avoided.

Multi-oscillator gyros have been reported as operating at the fundamental limit[85]. However, it should be noted that because of the added cavity loss due to the insertion of optical components, the fundamental limit is larger than it would be for a clear path gyro. With this tradeoff, angular random walk has been reported as being comparable with the two types of gyros that are of comparable size. In order to reduce the number of optical elements in the cavity, an out-of-plane geometry was implemented as a replacement for the quartz reciprocal optical activity element[12,86]. By this it is meant that the beams on reflection from the mirrors (even number required) do not travel in a plane. The frequency splitting occurs because the LCP and RCP polarization vectors can accumulate a different phase shift per pass, based on the individual phase shifts of the s and p vector upon reflection. It has been shown that a unique scale factor and input axis exists for this type of configuration[86]. The scale factor is determined by the projected area formed by the cavity perimeter on a surface normal to the input axis.

Other means, such as Zeeman splitting[87] and multi-mode coupling[12], to obtain the non-reciprocal splitting without the need for the insertion of a Faraday cell in the cavity have been proposed. No practical device using these techniques has been demonstrated due to the large error mechanisms associated with them.

Over the years of its development, there have been difficulties in implementing the multi-oscillator. Its concept is the operation of two independent gyros in the same cavity, which are separated in frequency by the reciprocal splitting. The concept of having two independent gyros means that because of the extremely high sensitivity of the RLG (small magnitude of the Sagnac effect), not all effects will be of common mode and hence cancel. For example, a gyro with the two sets of frequencies separated due to their operating with different polarizations, will, when the difference is obtained from the two outputs, have the Faraday effect cancel, but only to first order. Due to the dispersion of the Verdet constant of the Faraday rotation material over the symmetric frequency splitting, the Faraday splitting for the RCP and LCP modes differs. For typical designs, the difference, which can be equivalent to a value on the order of a deg/hr, can vary with temperature and must be compensated.

Another effect is due to the dispersion effects of the active gain medium. Because of line shape asymmetries, the Faraday splittings are unequal and dependent on magnitude of splitting, gas composition, laser intensity, external magnetic fields and temperature. Also, because of the bias sensitivity, more care must be given to the path length control of the multi-oscillator, than for the two-frequency RLG, with respect to tightness of control.

In designing the multi-oscillator with a Faraday bias element, insertion is a concern. Materials with large Verdet constants tend to have large coefficients of expansion, with resultant thermal sensitivity. The element (glass with permanent magnets) must be attached to the laser frame in a manner consistent with no thermal stress. This is required to avoid high thermal and magnetic sensitivity. In that the elements, including mounting material (typically indium) are inserted into the cavity, cleanliness and out-gassing must be considered.

A further issue is the back-scattering, which although is compensated by the Faraday element, can lead to bias sensitivity[88]. The mechanism is thought to be different scattering onto the modes of the two separate polarizations. This results in scale factor differences, which translate to bias offset that do not cancel in the common mode cancellation of the readout. This bias mechanism has a large thermal sensitivity. To avoid this issue, it is possible to make use of the fact that scattering causes a

modulation on the intensities, as well as the frequencies. Then calibration techniques can be used to compensate the beat frequency output with intensity measurements. Long term stability of the coefficients becomes the issue.

At this time it is still an open issue as to the superiority of the multi-oscillator vs. the dithered two-frequency RLG. This issue may never be resolved, but if is, it will be resolved based on a combination of cost, performance and reliability.

3.8.9 Magnetic Mirrors

In spite of the fact that no work is known to be ongoing at this time, a final lock-in compensation approach will be discussed because of its elegance. The concept is to obtain the differential phase shift needed to bias the locked beams, by using the reflectivity of the mirrors[89]. This approach avoids the insertion of optical elements in the cavity and the need for mechanical dither. In its first implementation, use was made of the transverse Kerr magneto-optic effect and it was known as a magnetic mirror. For the transverse Kerr effect, a ferromagnetic material, such as iron, iron garnets, permalloy, etc., is deposited on the surface of the mirror. In the presence of a magnetic field along the surface of the mirror and perpendicular to the plane formed by the laser beams, analysis shows that by matching the boundary conditions at the surface, a phase shift occurs at reflection. In that the phase shift is asymmetric with respect to the direction of the beams, an effective inertial bias is introduced by means of the differential phase shift. To avoid the previously discussed problems of a dc bias, alternating the direction of the magnetic field provides an ac bias.

As can be expected, there are numerous difficulties associated with the use of a magnetic mirror, although none are insurmountable. The first is that the transverse Kerr magneto-optic effect requires the beams to be of p-polarization. In that the reflectivity is higher for s-polarization with non-normal incidence of the beams, some method is required to force the p-polarization to oscillate. In the original implementation, modular gain tubes were used, which required Brewster angle windows to minimize reflectivity-type loss. Hence the windows could be oriented for p-polarization. However, this approach negated the advantages of a clear path RLG, and was discontinued. Another approach, which has been successfully implemented, made use of the phenomena that there is a phase shift difference between s and p-polarization light upon reflection from the mirrors at non-normal incidence. Then for cavities having an odd number of mirrors, the set of longitudinal modes of the s and p-polarizations will be spaced one half of a wavelength apart. Thus midway between each s-polarization cavity mode, there will be a p-polarization mode. Then with a polarizer on the output, the path length control can be designed such that the gyro operates with p-polarization. This is more readily achieved with smaller gyros, that have a larger frequency separation between longitudinal modes. Because of the lower reflectivity for the p-polarization, cavity losses will be higher, with a corresponding increase of error terms. In that the magnetic mirror approach is most suitable to low-cost lower performance applications, this performance degradation could be acceptable. It should be noted that although the reflectivity is lower for the p-polarization than for the s-polarization, the transmission is higher and sufficient power is available to drive the output electronics.

Another consideration is that the ferromagnetics are exceedingly lossy in the visible and this requires operation using the He-Ne infrared 1.15 μm transition. Losses are still high but it is possible to reduce them by placing the magnetic layer inside the stack[70]. There is then a tradeoff between loss and differential phase shift, which is reduced as the magnetic layer is buried. The usage of the infrared transition results in reduced performance in that error terms tend to be higher. For example the dispersion terms scale with wavelength and are higher, resulting in higher gas flow errors. They also scale with gain and higher losses require a larger operating gain for the usual design of a constant gain/loss ratio. Because of the reduced Doppler width, a higher gain/loss ratio would be required to maintain adequate output power. Also, detectors are less sensitive, requiring additional power.

Because of the performance disadvantages of operating with the IR transition, an attempt was made to operate a magnetic mirror RLG using the visible transition[90]. A permalloy layer was deeply buried in the mirror stack, thus allowing low enough loss such that the laser was able to oscillate. Although the bias was sufficient to unlock the gyro, its magnitude was not sufficient to obtain a practical instrument.

Early designs of the magnetic mirror, had large linear magnetizations that required magnetic field current reversals having slow switching speeds and high power consumption. In addition, rectification and bias instability was an issue. A later design using an iron-permalloy magnetic mirror with a low magnetic field saturation level, and large hysteresis, required only pulsing to drive between saturation levels[91]. Thus no power was required to maintain the saturated field and with a switching time only on the order of a microsecond, switching frequencies in the tens of KHz were achievable. With this high frequency dither, angular white noise was dominated by the quantum noise. For a 10 cm perimeter, noise was determined as being in the 0.1 deg/hr $^{1/2}$ range.

A problem area, that causes performance degradation, is that it is not possible to obtain a differential phase shift without a differential loss occurring. At best, an optimal design can be made that trades the magnitude of the two. With these

performance issues, it appears that although the magnetic mirror can be implemented, especially using a monolithic housing with three orthogonal gyros, production issues and competitive technologies have raised enough questions so that the magnetic mirror RLG is no longer being pursued.

LIST OF REFERENCES

1 C. V. Heer, Bulletin of the American Physical Society 6, 58, (1961)
2 A. Rosenthal, Journal of the Optical Society America 52, 1143, (1962)
3 W. M. Macek and D. T. M. Davis, Jr., Applied Physics Letters 2, 67, (1963)
4 E. McCartney, Journal of the Institute of Navigation 13, 260, (1966)
5 J. Killpatrick, IEEE Spectrum 4, 44, (1967)
6 G. Sagnac, C. R. Academy of Science 157, 708, (1913)
7 P. Savage, AGARD Lecture Series #95 "Strap-Down Inertial Systems" June 1978, NTIS, Springfield VA
8 P. J. Klass, "Honeywell Breaks Into Inertial Market", Av. Week & Space Techn (11/19/79), pp. 78-85
9 C. V. Heer, "Physics of Optics of Ring Gyros" SPIE Vol 487, (1/7-10/84) pp. 2-12
10 D. Mackenzie, Technology and Culture, 34, 475 (1993)
11 F. Aronowitz, The Laser Gyro in Laser Applications Vol I, edited by M. Ross, Academic Press (1971)
12 W W. Chow, J. Gea-Banacloche, L. M Pedrotti, V. E. Sanders, W. Schleich and M. O. Scully, Rev. Mod. Phys, 57, 61 (1985)
13 J. R. Wilkinson, Progress in Quantum Electronics, Vol 11 Pergamon Press 1987
14 R. W. Dithchburn, "Light", p. 339, Wiley, New York, 1959
15 M. P. Langerin, C.R. Academy of Science, 173, 183, (1921)
16 E. J. Post, Reviews of Modern Physics, 39, 475 (1967)
17 P. Savage, IEEE PLANS, San Diego, (11/1-3/76)
18 J. Fulfrost and W. J. Checco, IEEE PLANS, pp. 214-9, (1984)
19 R. G. Majure and T. A. Robinson, IEEE PLANS pp. 96-100, (1986)
20 T. Callaghan, S. Callaghan, J. Hanse, C. Tettemer and F. Aronowitz, "Symposium Gyro Technology 1983" Stuttgart,Germany
21 T. J. Podgorski U.S. Patent # 3,390,606, July 2, 1968
22 "Laser Gyro Comes In Quartz", Electronics, 39, 183 (1966)
23 F. Aronowitz, Physical Review, A139, 635 (1965)
24 W. L. Lim, J. P. Hauck and J. W. Raquet, U.S. Patent # 4,705,398 (11/10/87)
25 S. Chao, W. Lim, J. Hammond, "Physics of Optical Ring Lasers", SPIE Vol 487 1984, pp. 50-57
26 T. Podgorski and F. Aronowitz, IEEE Journal of Quantum Electronics 4,11 (1968)
27 B. Doyle, unpublished analysis at Honeywell, 1966
28 Schott Glass Technologies, Inc., Duryea Pa. 18642
29 J. Killpatrick and M. Weber, Eighteenth Joint Services Data Exchange - Inertial Systems, (10/28-30/86)
30 Unpublished work at Raytheon, 1983
31 B. H. G. Ljung, J. G. Keper, U.S. Patent # 4,159,059, June 26,1979
32 T. Dorschner, unpublished experiments at Ratheon
33 J. H. Simpson, Proceedings of IEEE NAECON 1980, pp 80
34 T. Dorschner, H. Haus, M. Holz, I. Smith and H. Statz, IEEE Journal of Quantum Electronics, QE16, 1376 (1980)
35 M. Tehrani, unpublished analysis at Honeywell, 1979
36 M. Sargent III, M. O. Scully, W. E. Lamb Jr., "Laser Physics", Adison-Wesley (1974)
37 F. Aronowitz, Ph.D. Thesis, New York University (1969)
38 W. E. Lamb, Physical Review, A134, 1429 (1964)
39 Refer to references 11-13 for numerous publications.
40 F. Aronowitz and R. J. Collins, Applied Physics Letters, 9, 55 (1966)
41 C. V. Heer and R. D. Graft, Physical Review 140, A1088 (1965)
42 C. V. Heer, Physical Review, 134, 799 (1964)
43 F. Aronowitz, Applied Optics, 11, 2146 (1972)
44 P. W. Smith, International Quantum Electronics Conference, Montreal, May 1972
45 F. Aronowitz, J. E. Killpatrick and S. P. Callaghan, IEEE Journal of Quantum Electronics, 10, 201, (1974)
46 Y .L. Klimontovich, P. S. Landa, and E. G. Lariontsev, Soviet Phys, JETP 25, 1076 (1967)
47 S. Stenholm and W. E. Lamb, Jr., Physical Review 181, 618 (1969)
48 R.H. Cordover and P.A. Bonczyk, Physical Review, 188, 696 (1969)
49 D. Smith and D. Shernoff, Applied Optics, 24, 1722 (1985)
50 S. Roden, D. Andrews, C. Knipe and T. King, SPIE, 1694, 208 (1992)
51 F. Aronowitz and W. L. Lim, SPIE Laser Inertial Rotation Sensors, 157, 7 (1978)
52 I. Langmuir, Journal of the Franklin Institute, 196, 751 (1923)
53 Unpublished experiments at Rockwell using a square 38 cm cavity.
54 R. Adler, Proceeding of the IRE, 34, 351 (1946)

55 W. Macek, D.Davis, R.Olthvis, J. Schneider and G. White, "Optical Masers", pp.199, Polytech. Instit. of Brooklyn Press, 1963
56 F. Aronowitz and R. J. Collins, Journal of Applied Physics, 41, 130 (1970)
57 F. Aronowitz and W. Lim, IEEE Journal of Quantum Electronics, 13, 338 (1977)
58 C. C. Wang, Proc. Symposium on Modern Optics, Polytechnic Institute of Brooklyn, 1967
59 F. Aronowitz, Journal of Applied Physics, 41, 2453 (1970)
60 H. Haus, H. Statz and I. Smith, IEEE Journal of Quantum Electronics QE21, 78 (1985)
61 T. J. Podgorski, U.S. Patent 4,152,071 May 1, 1979
62 T.J. Hutchings, U.S. Patent # 4,281,930 (1981)
63 R. Patterson, B. Ljung and D. Smith, "Physics of the Optics of Ring Lasers", SPIE Vol 487, 1984 pp. 78-84
64 F. Aronowitz, unpublished analysis at Honeywell, 1977
65 W. Egli, unpublished analysis at Honeywell, 1977
66 J. E. Killpatrick, unpublished analysis at Honeywell, 1966
67 J. E. Killpatrick, SPIE "Physics of the Optics of Ring Lasers", Vol 487 1984 pp 85
68 T. Hutchings and D. Stjern, IEEE National Aerospace and Electonics Conference, 1978 pp. 549-555
69 J. Killpatrick, U.S. Patent #3,467,472 Sept. 16, 1969
70 J. Krebs, W. Maisch, G. Prinz and D. Forester, IEEE Transactions On Magnetics, 16, 1179 (1980)
71 N. Buholz and M. Chodorow, IEEE Journal of Quantum Electronics 3, 454 (1967)
72 Laser Gyro Comes In Quartz, Electronics, 39, 183 (1966)
73 J. Hanse and J. Killpatrick, U.S. Patent #4,751,718 June 14, 1988
74 D. Weber, Journal of the Institute of Navigation, 35, 15 (1988)
75 A. Thomson and P. King, Electronics Letters, 2, 382 (1966)
76 W. Macek, J. Schneider and R. Salamon, Journal of Applied Physics, 35, 2556 (1964)
77 B. Dessus, J. Catherin and J. Migne, C.R. Academy of Science, B262, 1691 (1966)
78 T. Hutchings, J. Winocur, R. Durrett, E. Jacobs and W. Zingery, Physical Review, 152, 467 (1967)
79 T. Hutchings and D. Stjern, Proc. IEEE Nat. Aerospace and Electronics Conference pp.549-555, 1978
80 H. de Lang, Phillips Res. Repts, 19, 429-440 (1964), PhD Dissertation, University of Utrecht, 1966
81 K. Andringa, U.S. Patent #3,741,657 (1973), Electr. Prog. 17, 20 (1974)
82 R. Warner, U.S. Patent #3,890,047 (1975)
83 G. Yntema, D. Grant, and R. Warner, U.S. Patent #3,862,803 (1975)
84 W. Chow, J. Hambenne, T. Hutchings, V. Sanders, M. Sargent and M. Scully, IEEE J. Quant. Electr., QE16, 918 (1980)
85 T.A. Dorschner, H.A.Haus, M. Holz, I.W. Smith, H. Statz, IEEE Journal Quantum Electronics QE-16, 1376 (1980)
86 H. Statz, T. Dorschner, M. Holz, I.W. Smith, "The Multi-oscillator Ring Laser Gyro", Raytheon Research Report ER83-4131, January 1983
87 V. Sanders, S. Madan, W. Chow, M. Scully, Optics Letters 5,99 (1980)
88 Litton Guidance and Control Systems, "Nondithered Ring Laser Gyro Study", Avionic Laboratory Report AFWAL-TR-119, November 1986
89 R.F. Morrison, E. Levinson, B.L. Bryant, Proc. 1977 IEEE National Aerospace Electronics Conference, p.1045
90 Unpublished work at Honeywell
91 R. Eichner, R. Hansen, R. Ouellette, Institute of Navigation., January 22-24, 1991

4. MULTIOSCILLATOR RING LASER GYROSCOPES AND THEIR APPLICATIONS

C. H. Volk, S. C. Gillespie, J. G. Mark, D. A. Tazartes

Litton Guidance and Control Systems
5500 Canoga Ave
Woodland Hills, CA 91367-6698

INTRODUCTION

The ring laser gyroscope has been both a technological and commercial success, having displaced the mechanical gyroscope as the angular sensor of choice for a variety of applications including navigation and motion control. The ring laser gyroscope provides excellent performance at lower cost and with higher reliability than is achievable with the mechanical gyroscopes, which are inherently more complicated and prone to wear-out.

Although the ring laser gyroscope provides a very elegant solution to the measurement of angular rate, it does have one significant problem; at low rates frequency locking between the two counter-propagating beams occurs which results in a null output. This effect is referred to as lock-in. Lock-in occurs because scatter from the ring laser's reflecting surfaces couple the two counter-propagating laser beams. As with any coupled oscillator system, frequency pulling will result and with sufficiently strong coupling, synchronization of the oscillating frequencies will ultimately occur. The earliest investigations of the ring laser gyroscope concerned the development of an optical solution to the lock-in problem; however, significant performance problems were encountered and mechanical dithering of the laser gyroscope was finally adopted as a practical solution to the lock-in problem. Although the introduction of dither allowed the ring laser gyroscope to achieve preeminent stature for navigation applications, it did result in some limitations. These include mechanical angle noise, acoustic noise, dither cross-coupling, as well as other detrimental system level effects. Moreover, the use of dither limits RLG performance to less than that theoretically possible. The multioscillator ring laser gyroscope, as embodied in Litton's Zero-Lock Laser Gyro® (ZLG®), resolves the early problem encountered with applying optical bias and overcomes the disadvantages present in dithered ring laser gyroscopes.

HISTORY

H. de Lang [1] was evidently the first to realize that the frequency locking problem associated with the use of a two traveling wave ring laser gyroscope at low input rates could be eliminated by a four traveling wave ring laser device. He suggested that a ring laser with Faraday rotation and phase anisotropy chosen such that the total phase anisotropy was roughly circular would yield a four frequency ring laser gyroscope which would not suffer the locking problem. He also showed how to extract the rate dependent

term from the four individual frequencies. De Lang used the Jones matrix [2] approach to calculate the eigenfrequencies for the ring interferometer with Faraday rotation and phase anisotropy.

United Aircraft Corporation was apparently the first group to become interested in this concept; their US patent [3] filed in 1968 discloses the necessary embodiments for this device referred to as the Differential Laser Gyro (DILAG). The DILAG multioscillator concept used a polarization isotropic ring cavity in conjunction with two types of polarization anisotropic biasing elements in the laser cavity (Figure 1 is from the patent which pictorially shows how four frequency operation is obtained). In the DILAG patent, reciprocal (non-propagation direction dependent) polarization anisotropy (RCP is right circular polarization, LCP is left circular polarization) is achieved via the use of a quartz crystal, while the directional anisotropy is achieved via a Faraday cell.
The DILAG patent was issued in January 1975.

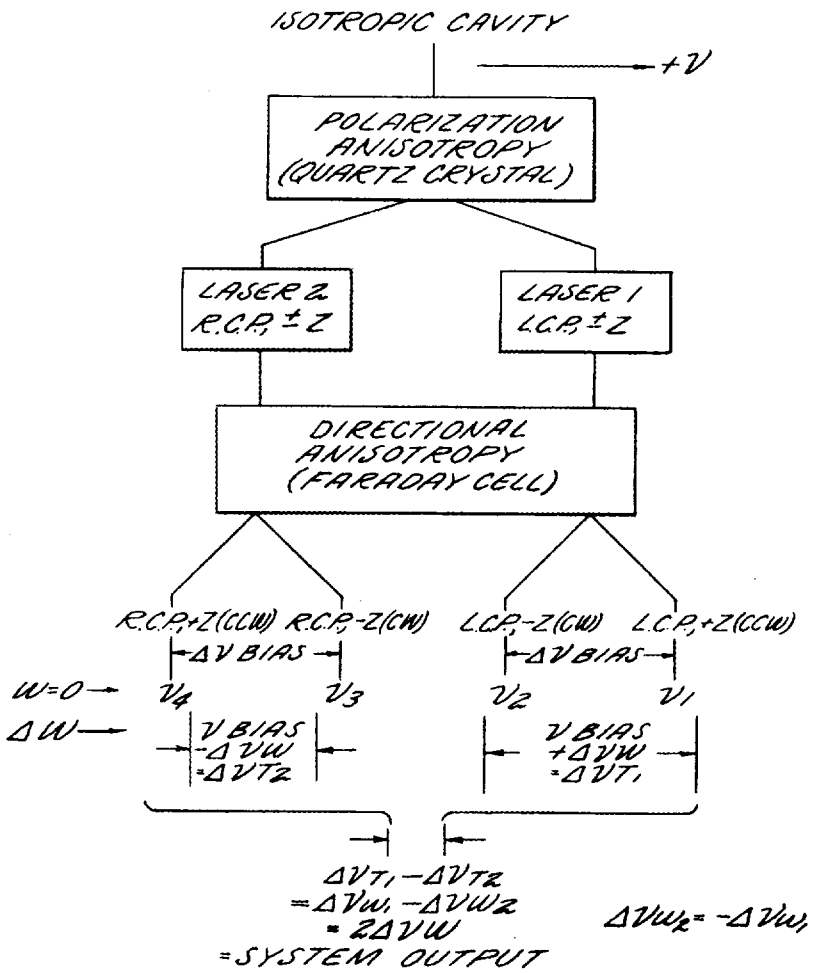


Figure 1 - DILAG Principle

Figure 2 shows Hamilton Standard's Division of United Aircraft Corporation's DILAG configuration circa 1977 [5].

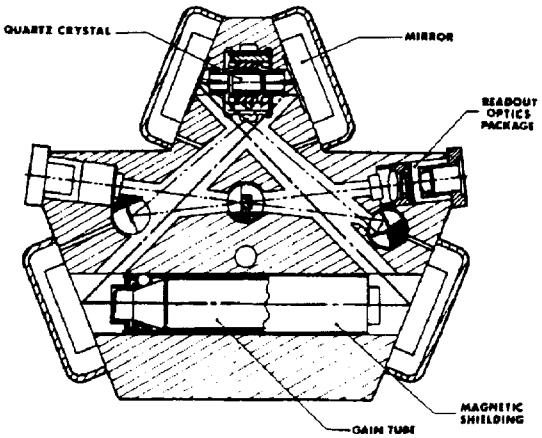


Figure 2 - United Aircraft's DILAG Configuration (circa 1977)

Andringa [4] (Raytheon Corporation) independently arrived at the same conclusions—Andringa was issued a patent involving the same concept in 1973 after filing in 1971.

A team from Litton Industries Incorporated and the Optical Sciences Center from the University of Arizona also embarked on a multioscillator ring laser gyroscope development activity in the mid 1970's. This group investigated a variety of different four frequency ring laser gyroscope approaches [6,7,8].

All these early approaches to "optical bias" resulted in performance limitations. Thus, implementation of body dither [9] provided the first practical solution to the lock-in problem and this technique was employed in the first generation of production ring laser gyros.

The critical step in devising a practical multioscillator ring laser gyroscope was the realization that the reciprocal polarization anisotropy could be achieved via the use of a non-planar ring laser optical cavity by Dorschner and Smith [10,11]. Elimination of the quartz crystal is important for minimizing the temperature sensitivity of the multioscillator ring laser gyroscope; additionally, reduction in the cavity loss leads to lower angle random walk instruments. Elimination of the quartz crystal for the reciprocal polarization element led to the possibility of a clear path multioscillator ring laser gyroscope via the use of alternate methods of non-reciprocal biasing. Besides use of a Faraday element for non-reciprocal splitting, the use of biasing via Zeeman splitting of the gain media and the use of a magnetic mirror based on the polar Kerr effect were considered [12, 13]. Figure 3 shows an early (circa 1979) Litton out-of-plane ring laser gyro.

Figure 3 - Litton non-planar ring laser gyro (circa 1979)

The goal of a clear path multioscillator ring laser gyroscope has been elusive. Magnetic mirror biasing has not proven to be practical due to high loss from the magnetic mirrors. Zeeman non-reciprocal biasing introduces gyro sensitivity to pathlength and magnetic field variations [12].

In 1985 Litton Industries acquired the Raytheon multioscillator ring laser gyroscope technology. Development of a production multioscillator ring laser gyroscope (the Zero-Lock Laser Gyro® (ZLG®)) was initiated [14]. The ZLG is a four frequency ring laser gyroscope which uses a Faraday element for the non-reciprocal biasing and a non-planar ring resonator configuration to achieve the reciprocal polarization anisotropy. Key to the success of this approach has been the development of extremely low loss anti-reflective coatings (<100 ppm) for the Faraday rotator glass surfaces. The ZLG has been in production since 1991 (over 1500 commercial systems utilizing this instrument have been manufactured).

BASIC PERFORMANCE

Principles of Operation

For earlier reviews on multioscillator ring laser gyroscopes see [16], [17] and [18].

Multioscillator Ring Laser Gyroscope Concept

For an electromagnetic oscillator in an inertial frame of reference, the solutions to Maxwell's equations are four-fold degenerate. Two states of energy momentum coexist for each direction of propagation. In non-inertial systems, Maxwell's equations predict the four-fold degeneracy will be split into components depending on polarization state and propagation direction. De Lang [1] realized that a ring laser with Faraday rotation and

phase anisotropy would operate at four frequencies and could be used as a differential gyro to detect rotation rate via the Sagnac [15] effect.

Figure 4 shows the elements of a multioscillator ring laser gyroscope.

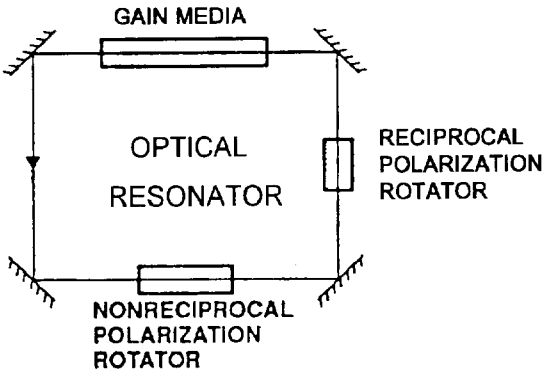


Figure 4 - Elements of a Multioscillator Ring Laser Gyroscope

An optical ring resonator with gain provides the basis for the ring laser gyroscope. In order to achieve a multioscillator ring laser gyroscope configuration, reciprocal and non-reciprocal polarization rotators are required. The reciprocal (i. e. non-direction dependent) frequency splitting is typically achieved via the use of a non-planar optical resonator which removes the frequency degeneracy between the polarization eigenmodes (left circularly polarized like (LCP) and right circularly polarized like (RCP)). The reciprocal splitting is determined by the cavity image rotation which is dependent on the cavity out-of-planeness. Since each polarization eigenmode can oscillate as a anti-clockwise (aw) or clockwise (cw) traveling wave, the possibility of four frequency operation is open. Non-reciprocal (propagation direction dependent) frequency splitting is achieved via the use of a Faraday rotator or use of a magnetic mirror or Zeeman bias. Figure 5 shows the resulting four frequencies relative to the Doppler broadened gain curve.

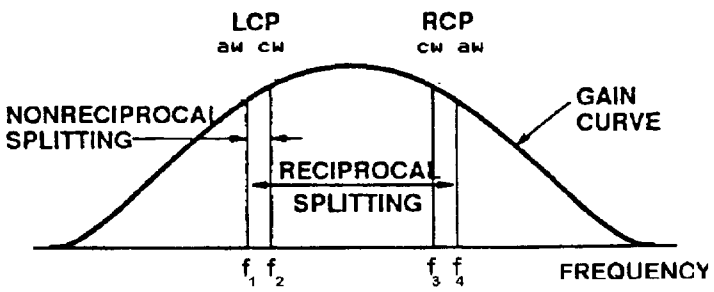


Figure 5 - Modes of a Multioscillator Ring Laser Gyroscope

Note that all four frequencies must oscillate simultaneously to achieve a multioscillator ring laser gyroscope. Reciprocal splitting is set by the degree of cavity out-of-planeness (of the order of 800 MHz for the ZLG) while the Faraday splitting is of the order of 1 MHz.

Differential Laser Gyroscope

Removing the direction dependent frequency degeneracy via the Faraday effect creates two oppositely biased two-frequency (nominally right and left circularly polarized) laser gyroscopes in one cavity. Since each gyro is optically biased by the Faraday effect, lock-in which occurs in conventional RLGs is eliminated. Application of rate increases the frequency splitting of one gyro while decreasing the splitting of the other via the Sagnac effect. A differential gyro is achieved by taking the difference between the two frequency gyros

$$f = (f_4 - f_3) - (f_2 - f_1)$$

With differential output, variations in the non-reciprocal bias are common mode rejected while twice the rate sensitivity relative to a two-frequency RLG of the same geometry is achieved. The frequency locking problem observed in the two frequency RLG is avoided under all operating conditions as long as the Faraday bias is larger than the Sagnac induced frequency splitting at the highest rates required for gyro performance. This is usually the case since typical Faraday biases in the range of one MHz are readily achievable.

Non-planar ring laser gyroscopes

Sagnac effect for non-planar ring laser gyroscopes

Statz etal [16] have shown a non-planar ring laser gyroscope experiences the Sagnac effect calculated to first order in velocity divided by the speed of light. They find that for a four frequency non-planar ring laser gyroscope with co-moving optical media the output frequency is given by

$$f = \frac{8A}{n\lambda_0 L}\Omega \cdot \mathbf{a}$$

where Ω is the input rotation rate, n is the refractive index of the co-moving media, λ_0 is the free space wavelength, L is the pathlength , A is the magnitude of the enclosed area and **a** is the unit area vector.

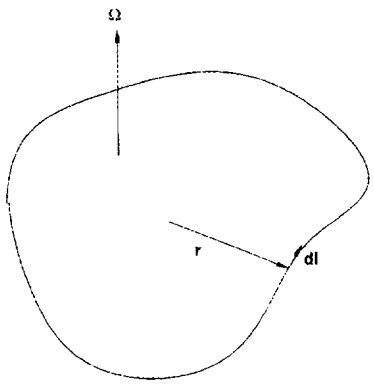


Figure 6 - Line Elements of Ring Cavity

The vectorial area is defined as (see figure 6)

$$\mathbf{A} = A\mathbf{a} = \frac{1}{2}\oint n^2(1-\alpha)\mathbf{r} \times \mathbf{dl}$$

where **dl** is an incremental vector line element pointed along the ring contour at position **r**, n is the refractive index of the co-moving medium and α is the Fresnel-Fizeau drag coefficient. The sensitive axis of the ring is uniquely determined by the above equation.

Image Rotation in Non-planar Ring Resonators

Non-planar cavities are used to provide reciprocal polarization rotation via image rotation. Upon each traversal of the non-planar resonator, an image is rotated about the direction of propagation. The amount of rotation depends on the geometry of the ring. The entire transverse field distribution is rotated and with it the polarization states. Statz etal [16] (and references listed therein) have discussed the image rotation properties of non-planar optical resonators. In order to calculate the image rotation for a non-planar ring, we remember that for an ideal mirror there can be no net transverse electric field components in the mirror plane and the normal component of the electric field must be conserved. For the kth ideal mirror of a non-planar resonator, the reflection condition can be written in matrix form as [16]

$$\mathbf{E}_{rk} = F_k\mathbf{E}_{ik}$$

where $\mathbf{E}_{rk}$ is the reflected electric field from the kth mirror, $\mathbf{E}_{ik}$ is the electric field incident on the kth mirror and F_k is the orthonormal reflection matrix

$$F_k := \begin{bmatrix} 2\cdot a^2 - 1 & 2\cdot a\cdot b & 2\cdot a\cdot c \\ 2\cdot a\cdot b & 2\cdot b^2 - 1 & 2\cdot b\cdot c \\ 2\cdot a\cdot c & 2\cdot b\cdot c & 2\cdot c^2 - 1 \end{bmatrix}$$

where a, b and c are the direction cosines for the kth mirror. Considering the mirrors to be the only direction-changing elements in the non-planar resonator, the total change in the electric field after successive reflection at M mirrors is given by

$$\mathbf{E}_r = \mathbf{F}_M\mathbf{F}_{M-1}\ldots\ldots\mathbf{F}_2\mathbf{F}_1\ \mathbf{E}_i$$

The production of two successive reflections is equivalent to a rotation, represented by an orthonormal matrix. Thus, for cavities with an even number of mirrors M=2k, the rotation can be written as [16]

$$\mathbf{R}=\mathbf{R}_k\ldots\mathbf{R}_1$$

where $\mathbf{R}_k = \mathbf{F}_M\mathbf{F}_{M-1}$. The rotation axis corresponding to R gives the direction of the optic axis of the ring and the equivalent rotation about that axis is the image rotation i for the resonator. The rotation matrix can be written [16] as

$$R(i) := I\cdot\cos(i) + (1 - \cos(i))\cdot\begin{bmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{bmatrix} + \sin(i)\cdot\begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix}$$

where I is the identity matrix. The cosine of the image rotation is given by [16]

$$\cos(i) := \frac{1}{2}\cdot(\mathrm{Trace}(R(i)) - 1)$$

The square root of the sum of the squares of the three elements of the skew symmetric part of R gives the sine of the image rotation. The skew symmetric elements divided by the sine of the image rotation are the direction cosines for the axis of rotation.

For a planar ring resonator of even number of mirrors, the effect of the planar reflections is to provide no image rotation.

Skew Rhombus Rings

In order to get an idea of how the reciprocal splitting is controlled by the properties of the out-of-planeness of the resonator, consider a skew rhombus ring which is formed by taking a planar rhombus and folding it out of plane about one of its diagonals. The image rotation is determined by the fold angle (γ) and twice the angle of incidence of the mirrors

(θ_1 and θ_2). Figure 7 shows the geometry of the skew rhombus ring resonator configuration—the planes of incidence of each of the four mirrors forming the non-planar resonator are formed from each pair of adjacent line segments.

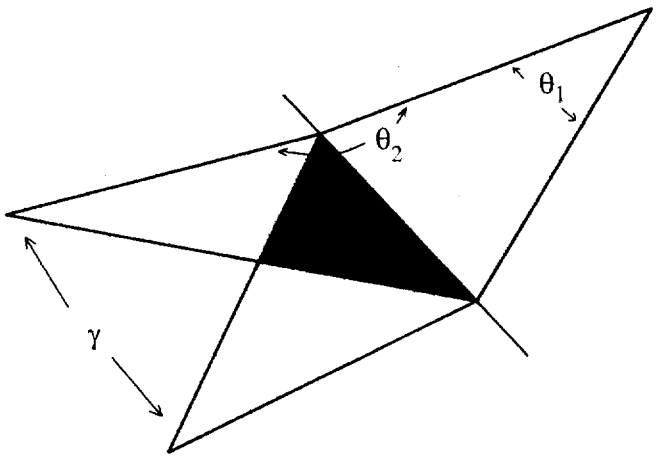


Figure 7 - Skew Rhombus Out-of-Plane Geometry

Figure 8 shows a plot of image rotation versus angle of incidence for the special case where all mirrors have the same angle of incidence.

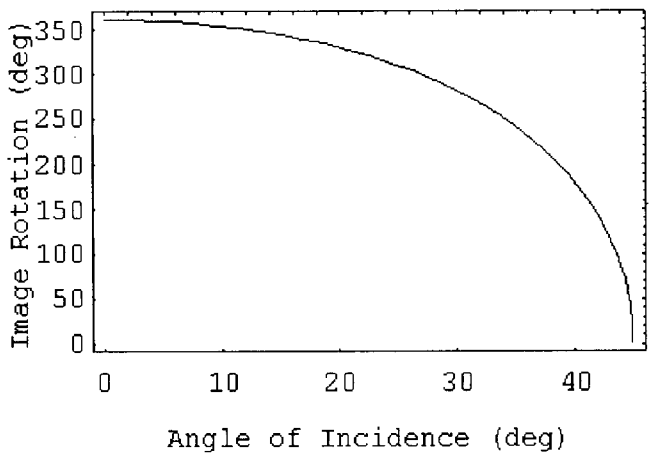


Figure 8 - Image Rotation (degrees) vs. Mirror Angle of Incidence for Skew Rhombus with All Mirror Angles of Incidence Equal ($\theta_1=\theta_2=\theta$)

The reciprocal frequency splitting is given by

$$\Delta f_{reciprocal} := \frac{c}{L} \cdot \frac{i}{\pi}$$

where i is the image rotation.

Bias

Although the multioscillator ZLG is a differential device, there are nevertheless sources of null offset. Statz etal [16] have considered a number of sources of null offset. Some of the more important error sources are discussed in more detail below.

Bias from the dispersion of the Verdet coefficient

The two dual-frequency gyros associated with the nominally LCP and RCP cavity polarizations are oppositely non-reciprocally biased via the Faraday effect. The Faraday bias is the product of the Verdet coefficient of the thin slab of amorphous glass times the longitudinal component of the applied magnetic field times the slab's thickness.

$$\Delta f_{non-reciprocal} = \frac{c}{\pi L} \cdot V(\lambda) \cdot B \cdot$$

where $V(\lambda)$ is the Verdet coefficient as a function of wavelength, B is the longitudinal magnetic field and t is the thickness of the Faraday rotator glass. The Faraday biases for the LCP and RCP modes are slightly different due to dispersion of the Verdet coefficient over the reciprocal splitting of the instrument achieved via the cavity out-of-planeness. For typical splittings, biases of the order of 5 Hz result from this effect.

Bias related to dispersion of the gain media

In the presence of the excited He-Ne gas mixture, the empty cavity frequencies are shifted. In the simplest approximation each of the four empty cavity frequencies are pulled towards the center of the laser transition in proportion to the product of their loss and distance from line center and in inverse proportion to the laser transition linewidth. We can write this shift as

$$\Delta f_{dis,\,linear} = -\text{loss} \cdot \frac{c}{2 \cdot \pi \cdot L} \cdot \frac{(f - f_0)}{f_D}$$

Where c is the speed of light, L is the cavity pathlength, loss is the fractional loss per pass of the mode , f_D is the Doppler width of the emission line, f is the empty cavity frequency of the mode and f_0 is the center frequency of the laser transition. Applying the values for the four frequencies and four losses yields the following result:

$$\Delta f_{disp_bias} = -4 \cdot \frac{c}{2 \cdot \pi \cdot L \cdot f_D} \cdot \left[MCDL \cdot \left(\frac{f_R}{2}\right) + DPL \cdot \left(\frac{f_F}{2} + f_B\right) + DDL \cdot f_d \right]$$

where f_R is the reciprocal splitting, MCDL is the magnetic circular dichroic loss defined as $(l_{Ra} - l_{Rc} + l_{Lc} - l_{La})/4$ (l_{Ra} is the loss of the nominally right circularly polarized anti-clockwise mode for example), DPL is the differential polarization loss defined as $(l_{Ra} + l_{Rc}$ -

l_{Lc} - l_{La})/4, f_F is the non-reciprocal (Faraday) bias, f_B is the frequency offset of the gain center line for a given helicity mode due to the value of any longitudinal magnetic field applied to the plasma, DDL is the differential direction loss defined as (l_{Ra} - l_{Rc} - l_{Lc} + l_{La})/4 and f_d is the detuning of the four frequencies relative to cavity center.

In the next higher order approximation to determine the mode pulling as a function of frequency, a term could be added to the model which is linear in loss but cubic in frequency difference

$$\Delta f_{dis,\,cubic} := loss \cdot \frac{c}{2 \cdot \pi \cdot L} \cdot k_3 \cdot \left(\frac{f - f_0}{f_D} \right)$$

where k_3 is a constant of the order of one.

In the model of the dispersive effects of the gain media given by the sum of the first and third order terms, it is also observed that bias effects are present which depend on the value of all types of differential loss as well as the average cavity loss defined as (l_{Ra} + l_{Rc} + l_{Lc} + l_{La})/4, the value of any longitudinal magnetic fields applied to the plasma (a longitudinal magnetic field on the gain media will shift the gain curve for the LCP_{cw} and RCP_{aw} modes (+ helicity) oppositely to the gain curve for the LCP_{aw} and RCP_{cw} modes (- helicity) relative to the zero magnetic field gain curve center) and also on the detuning of the four frequencies relative to cavity center. Bias instability arises in the multioscillator ZLG from inability to maintain a constant longitudinal field on the plasma and by the inability to hold the detuning of the cavity constant. However, with the composite dispersion model it is possible to show that a longitudinal field value exists which will make the derivative of the dispersion bias with respect to detuning zero and also a detuning exists which will make the derivative of the dispersion bias with respect to longitudinal field on the plasma zero.

In the particular embodiments of the multioscillator ring laser gyroscope which do not use Zeeman biasing for the non-reciprocal splitting a technique can be employed, called dispersion equalization, in which the proper value of longitudinal magnetic field is applied to the gain media and the proper pathlength control offset are selected to desensitize the gyro to magnetic or pathlength control variations. Pathlength detuning insensitivity in these instruments can be obtained via magnetic field generating coils which are placed in the vicinity of the plasma so that a longitudinal field of the proper value can be applied to the gain media.

Bias related to cavity backscatter

Cavity scatter sources which give rise to scale factor (SF) non-linearities in a two-frequency RLG will give rise to a multioscillator bias because the scatter for the LCP-like two frequency gyro is not generally the same as for the RCP-like two frequency gyro. The output frequency (f) for a two frequency RLG in the presence of backscatter is given by Reference [19] as :

$$f := f_i - \frac{1}{2 \cdot f_i} \cdot \left[(f_L)^2 - (f_2)^2 \cdot \frac{(f_i)^2}{\left[(\Omega_g)^2 + (f_i)^2 \right]} \right]$$

Or the SF error can be written as

$$\Delta SF := \frac{f - f_i}{f_i}$$

$$\Delta SF := -\frac{1}{2} \cdot \left(\frac{f_L}{f_i} \right)^2 + \frac{1}{2} \cdot \left[\frac{(f_2)^2}{(\Omega_g)^2 + (f_i)^2} \right]$$

The first term is the usual negative SF error; the second term has the opposite sign and is called the positive SF error. The terms are defined as follows:

$$f_L := \frac{c}{L} \cdot \left[(s_{cw})^2 + (s_{aw})^2 + 2\, s_{cw} \cdot s_{aw} \cdot \cos(\varepsilon_{cw} + \varepsilon_{aw}) \right]^{\frac{1}{2}}$$

$$f_2 := \frac{c}{L} \cdot \left[(s_{cw})^2 + (s_{aw})^2 - 2 \cdot s_{cw} \cdot s_{aw} \cdot \cos(\varepsilon_{cw} + \varepsilon_{aw}) \right]^{\frac{1}{2}}$$

$$\Omega_g := \frac{c}{L \cdot F} \cdot \left(\frac{g}{loss} - 1 \right) \cdot \left(\frac{1 - h}{1 + h} \right)$$

f_i = apparent rate

$s_{cw,aw}$ = cavity backscatter coefficient for the cw, aw directions

$\varepsilon_{cw,aw}$ = cavity backscatter phase for cw, aw directions

g = gain

h = hole overlap factor

F= cavity finesse

For scatter, consider the multioscillator as two independent gyros whose frequencies are "Adler pulled" from what they would be in the absence of scatter. Each of the two gyros in the multioscillator looks effectively as a two frequency gyro operating with a rate of the Faraday bias. At zero input rate Ω_g is much less than the apparent rate (which is the Faraday bias), so both terms give SF errors approximately proportional to the inverse square of the Faraday bias. Backscatter coupling is only considered between the Faraday split frequencies, since the reciprocal splitting is large enough to make interpair coupling

negligible compared to intrapair coupling. The cavity scatter for the LCP-like and RCP-like pairs will generally be different since the light is polarized differently and the frequency of the light is different. By subtracting the frequency pulling terms for two different pairs, we will obtain a bias which is dependent on the difference in scatter between the LCP-like and RCP-like pair and is approximately dependent on the reciprocal of the Faraday bias. Notice that the scatter terms in the SF error depend on the backscatter amplitudes and phases of the cavity. These backscatter terms are the vectorial sum of the scatter vectors for each of the optical elements which scatter light. For example

$$s_{cw} \cdot \exp\left(i\varepsilon_{cw}\right) := \sum_j s_{cw,j} \cdot \exp\left(2 \cdot i \cdot k \cdot z_j\right)$$

where k is the wavenumber of the mode, and $s_{cw,j}$ is the scatter from the jth scatter at position z_j along the beam path.

Scale factor

The scale factor of the multioscillator ZLG can be written as the product of a geometrical term, a gas dispersion dependent term, and a scatter dependent term. The geometric scale factor (SF_G) is $8A/\lambda L$ where A is the appropriate enclosed area, L is the perimeter, and λ is the operating wavelength. The geometrical scale factor is modified by small gain and loss dependent terms and cavity scatter effects. The scale factor can be written as

$$SF := SF_G \cdot D \cdot S$$

where D and S are the terms relating to dispersion and scatter respectively. The S term is different from one by terms that are proportional to the inverse of the Faraday bias squared. The scatter correction is <<1 ppm for typical Faraday bias and cavity scatter parameters.

Quantum Noise

The uncertainty principle of quantum mechanics imposes a limit on the measurement capability of any system. Dorschner etal [20] considered the operation of a four-frequency ring laser gyroscope at the quantum limit. They showed that a lower limit resultant from phase perturbations due to spontaneously emitted photons could be placed on the rotation-rate equivalent noise. This "quantum noise" results in angle random walk on the integrated rate output of the ring laser gyro. Quantum noise limits the accuracy of the measurement of angular rate as the reciprocal of the square-root of the output data sampling period. For a four-frequency ring laser gyroscope the noise is described by

$$\sigma_\Omega = S_\Omega \, T^{-1/2}$$

where σ_Ω is the rotation rate equivalent noise, T is the data sampling period and S_Ω is the angle random walk coefficient given by

$$S_{\Omega} := \left(1 - \frac{pop_1 \cdot g_2}{pop_2 \cdot g_1}\right)^{\frac{-1}{2}} \cdot \frac{\lambda \cdot c}{8 \cdot \pi \cdot A \cdot \sqrt{2}} \cdot loss \cdot \sqrt{\frac{h \cdot c}{\lambda \cdot P}}$$

where $pop_{1,2}$ and $g_{1,2}$ are the population levels and g factors for the laser levels (1 is lower, 2 is higher), h is Planck's constant and P is the power per mode. The angle random walk coefficient for zero-lock laser gyros operating at 632.8 nm has been experimentally determined to close to the theoretically predicted limit. The 18 cm pathlength ZLG angle random walk coefficient is typically in the vicinity of 0.0008 deg/sqrt hr.

Zero-Lock Laser Gyroscope®

The Zero-Lock Laser Gyroscope [14] is a Litton Industries embodiment of the multioscillator ring laser gyroscope which uses a non-planar cavity resonator in conjunction with a Faraday element to achieve four frequency operation. Associated functions are gain media excitation and control, gain media dispersion equalization and gyro readout. As in other ring laser gyroscopes, temperature sensors are provided for thermal compensation of the output data.

The ZLG utilizes a split dc He-Ne plasma to provide laser gain. A high voltage power supply is used along with plasma current regulation to provide constant current operation in each of the two discharge legs. A split discharge is utilized to mitigate bias caused by gain media motion relative to the gyro body.

A diaphragm mirror in conjunction with a piezoelectric transducer maintains pathlength over changing environmental conditions. The intensity from either of the beam directions can be used as the error source for a pathlength control system when pathlength modulation is applied.

Gain media dispersion control is used to mitigate the effects of mode pushing and pulling. As discussed previously, a longitudinal field is applied to the gain media and the cavity is slightly detuned. The magnetic field and the amount of detuning are determined during calibration. The desired values for these parameters are those that will lead to minimum sensitivity to deviations in magnetic field or further changes in cavity detuning.

The gyro readout method employs a combining prism in conjunction with a dual spot photodiode and electronic signal processing. The photodiode outputs are double sideband suppressed carrier signals which can be electronically processed to obtain the two dual-frequency gyro outputs. The two output frequencies are counted—the multioscillator output is the difference between the two frequencies.

ADVANTAGES

The pursuit of a practical optically biased ring laser gyroscope was motivated by more than a desire to have an elegant solution to the lock-in problem. Although it was been clearly demonstrated that body dithered ring laser gyroscopes perform well in many applications, there are still shortfalls using this approach and further there are many applications in which the body dithered approach is not applicable. In this section we describe particular advantages of the multioscillator ring laser gyroscope as embodied in the ZLG approach.

Angle Random Walk

The application of body dither is effective in allowing the ring laser gyroscope to measure very low rotation rates; rates below which the ring laser would normally lock-in. Lock-in effects still occur during each of the dither reversals. At the dither turnarounds there will be some loss of phase information [21]. Since the output angle of a ring laser gyroscope is derived from the frequency difference of the two laser beams, an error will accumulate due to the lock-in effect at the dither reversals. Even though the periods of time when the phase loss occurs are usually very short and thus the resulting output error would be very small, the resulting error is cumulative and in time can grow to become significant for high precision applications.

The magnitude of this effect is a function of the precise lock-in rate for the particular instrument and the amplitude of the dither motion. The lock-in rate is a function of the scatter in the optical cavity. The net scatter in the laser's optical cavity is a phasor addition of the individual scatter points from each of the reflecting optical surfaces. The phasor addition, in turn, is a result of the precise phase relationship of each of the scatter sites. This phase relationship will be a function of temperature because even though the pathlength of the optical cavity is controlled as the block material changes dimension over temperature, the distance between the optical surfaces will change as the moveable mirror adjusts in response to block dimensional changes. The change in relative distance between the scatter sites changes the phase relationship and thus the net magnitude of the scatter effect. This results in a temperature dependence to the dithered induced noise.

The optical bias used in the ZLG eliminates the problem of dither-induced noise. The noise for the ZLG is solely a function of the quality of the optics used for the laser cavity and the internal circulating power which is mainly a function of the current used to drive the laser plasma. There have been a number of schemes developed to circumvent the dithered-induced noise in the ring laser gyro [22 and 23]; however, these schemes add complexity to the instrument and are not completely effective. It then follows that the ZLG approach will always exhibit a lower random walk coefficient compared to a dithered ring laser gyroscope of comparable size. In addition the angle random walk coefficient of the ZLG is nearly independent of temperature.

Scale Factor

As has been discussed previously, the fact that the ZLG is equivalent to two ring lasers in the same optical cavity results in a doubling of the scale factor compared to a two-mode ring laser gyroscope of comparable size. More importantly, however, conventional RLG scale factor linearity is affected by lock-in that occurs during dither reversals in a manner similar to that discussed for phase information loss. This results in the ZLG exhibiting a much improved scale factor linearity performance. Scale factor linearity of a ZLG is at least an order of magnitude better than that of a dithered ring laser gyroscope of the same geometry.

Laser Gyro Readout

The readout signal from a laser gyroscope is found by combining the two counter-propagating beams with appropriate optics such as a prism. The mixing of the beams leads to an interference pattern consisting of a series of light and dark bands that move across the surface of the detector. In a conventional (two-frequency) RLG, the motion of the interference fringes corresponds to the inertial rotation of the gyro. A shift from light to dark or dark to light results from an angular motion whose arc of length is one half the wavelength of the laser beam. For navigation grade laser gyroscopes, this is on the order of 1.5 arc seconds. In conventional RLGs, two photodetectors are required in order to determine the direction of rotation. The motion of the fringes across the two detectors determines this direction.

In a ZLG, the optically biased gyroscopes are operating simultaneously. The optical bias due to the Faraday effect causes approximately 10^6 light to dark and 10^6 dark to light transitions per second. The number of transitions is modified by the angle of rotation. Detectors and detector electronics are configured to measure the intensity transitions for each gyro (LCP, and RCP) independently. At rest and in the absence of gyro bias, both streams of fringe transitions occur continuously in equal amounts due to the Faraday bias. In the presence of angular motion, the number of transitions in one stream will decrease while the number in the other stream will increase by the same amount. The difference in the number of fringe transitions represents the output of the differential gyro discussed earlier. For a given angle of rotation, the differential ZLG output is twice that of an RLG of the same size and is proportional to the angle of rotation since the effect of the optical (Faraday) bias has canceled out.

Resolution Enhancement

The normal method of detecting angle changes in laser gyroscopes (i.e. counting of fringe light to dark and/or dark to light transitions) is very advantageous as it provides a direct digital output of rotation information. In a navigation accuracy ZLG, each light to dark or dark to light transition corresponds to approximately 1.5 arc sec of angle change. With the differential ZLG output, an additional factor of two is gained, leading to typically a 0.75 arc sec of angle change. This, nonetheless, results in a finite output angle resolution

even though the interference pattern is tracking a continuous measure of angular displacement. No information is lost in the long term, however, an angle quantization error is incurred on an instantaneous basis.

The quantization error resulting from the readout detection mechanization leads to potentially noisy angular rate estimates in wide bandwidth systems. For example, a 0.75 arcsec angle error over a 1000 Hz interval corresponds to a 750 deg/hr angular rate error. For the ZLG, a very effective, purely digital method of reducing the quantization error to negligible values has been developed. The method takes advantage of the optical biasing present in the ZLG. The use of resolution enhancement gives the ZLG a clear advantage over a conventional dithered RLG in terms of noise characteristics. The ZLG output does not suffer from dither induced contamination. Further, resolution enhancement is far more difficult to achieve on a dithered gyros which are biased with a high frequency sinusoidal motion.

Figure 9a schematically shows the continuous and quantized angular displacement or fringe phases from a laser gyroscope as a function of time. The difference between the two curves is the quantization error which is shown in Figure 9b. For a constant angular rate, the quantization error is a periodic sawtooth function. Each of the two fringe phases is quantized independently leading to a difference phase, corresponding to rotation angle, whose quantization error behaves as displayed in Figure 9c. The resulting patterns preserve correct angle information in the long term but clearly limit angle resolution in the short term.

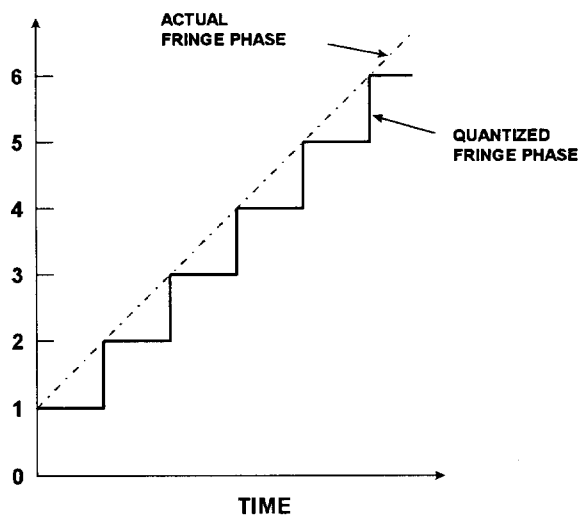


Figure 9a - Gyro Fringe Phase

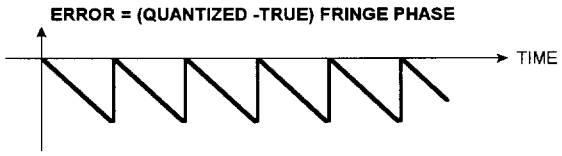


Figure 9b - Gyro Quantization Error

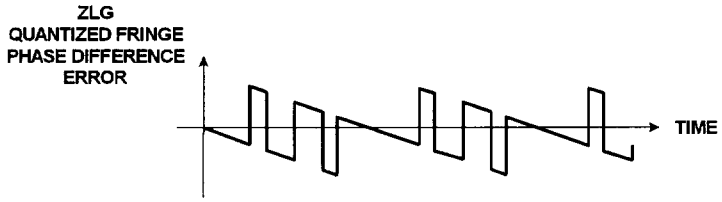


Figure 9c - ZLG Quantized Fringe Phase Differential Error

Digital Resolution Enhancement Technique

In the ZLG, resolution enhancement is accomplished by reducing the quantization error through the use of high speed digital signal processing; for a detailed discussion of this technique see references 24 and 25. Since the technique described is purely digital, it has excellent noise immunity and can be implemented at very low cost in conventional digital electronics. In a ZLG, the two fringe phases and quantization processes are independent and the angle output is a linear function of the two. Therefore, the processing of each fringe stream can be analyzed separately. Figure 10a shows how laser gyroscope data are normally processed to form an incremental angle. In Figure 10b, a moving average filter has been added to reduce the quantization noise. In the ZLG, the two gyro outputs are processed separately and combined after filtering. Because the moving average is a linear filter, the quantized signals can be combined and then filtered as shown in Figure 10c.

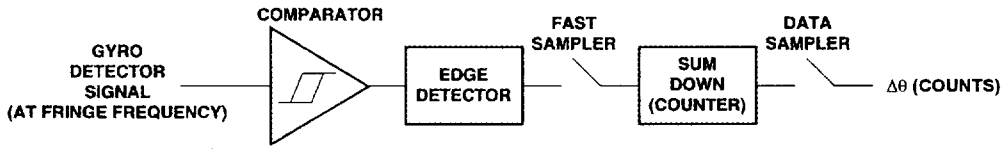


Figure 10a - Laser Gyro Conventional Data Processing

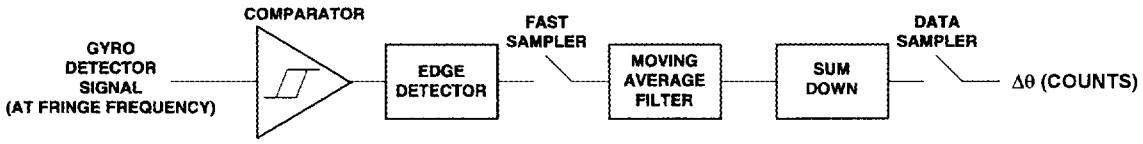


Figure 10b - Laser Gyro Data Processing with Filter

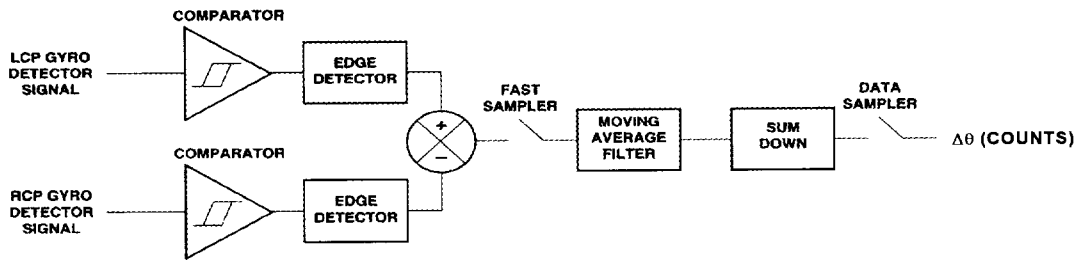


Figure 10c - ZLG Data Processing with Filter

The key to accomplishing a significant enhancement in angle readout resolution is the use of an optical bias that creates a frequency difference between the two counterpropagating laser beams that far exceeds the required angle tracking bandwidth of the system. A low pass filter can then greatly attenuate the quantization error without degrading the useful gyro information. Practical implementations of this technique have demonstrated three orders of magnitude improvement in angle readout resolution for the ZLG; thus a nominal quantization on the order of an arcsecond can be reduced to the order of 0.001 arcseconds.

APPLICATIONS

The ZLG is intended to fulfill the total range of applications possible for laser gyros including among others aircraft, missile, helicopter, ship, and land vehicle navigation, launch system guidance, flight control, motion compensation of other sensor systems, multi-function inertial systems, space related pointing and low noise applications. As opposed to the dithered RLG, the ZLG is a true strapdown angle sensor with no moving parts and no mechanical biasing. The absence of dither, the extremely low quantization noise (with resolution ehnancement), and the low angle random walk makes the ZLG ideally suited to many additional requirements which cannot be satisfied with conventional RLGs. For example, the ZLG may be used for spacecraft, submarines, and other platforms where low mechanical and acoustic noise is critical. Further, the ZLG can be hard-mounted to permit accurate, wide-bandwidth tracking of attitude. The ZLG's scale factor linearity and stability also permits its use in applications with large, persistent angular rates.

Navigation

ZLG-based strapdown inertial navigation systems are standard on many commercial and military platforms. The strapdown inertial system provides self-contained navigation and attitude information. It does not rely on external sources (such as radio waves) to formulate its solution. Therefore, the inertial system is immune to external interference and can operate in a wide range of conditions.

Inertial navigation remains the lowest noise and most accurate method of tracking high frequency motion. However, due to instrument errors and drifts, the inertial navigation and attitude degrade with time. Thus, to ensure long term performance, the sensors

(gyroscope and accelerometers) must conform to very stringent performance requirements. Strapdown inertial navigation systems use accelerometers to measure vehicle acceleration in three dimensions. However, in order to relate the orientation of the vehicle (and hence the accelerometers) with respect to the navigation reference frame, it is also necessary to track the vehicle's attitude. Gyroscopes measure rotation of the vehicle and their outputs are appropriately integrated to provide the attitude of the vehicle in the form of a direction cosine matrix. The accelerometer outputs are transformed through the attitude matrix to the reference frame where they are integrated into velocity and position. Thus, the inertial navigation system can provide attitude information as well as velocity and position information.

The key parameters for a strapdown navigation gyroscope are its dynamic range, angle random walk, bias stability, and scale factor accuracy and linearity. ZLG-based strapdown navigation systems are standard equipment on many aircraft and other vehicles. Approximately half of the ZLGs produced are used in commercial airliner navigation systems. These systems maintain position error drift rates typically better than one nautical mile per hour, and will better 2.0 nautical miles per hour error rates at least 95% of the time. Typical military aircraft navigation systems exhibit a position error drift rate better than 0.8 nautical miles per hour.

In many instances, the Global Positioning Satellite (GPS) System is used to bound the long-term performance of the inertial navigation system. An integrated system of this type provides significant advantages. Low noise and short term accuracy, as well as immunity to external disturbances are provided by the ZLG-based inertial navigation system. Long-term accuracy is ensured by the GPS system which provides drift-free updates.

Specifications

Table I gives a specification of a typical ZLG-based inertial navigation system and Figure 11 shows typical navigation performance for such a system.

Resolution	0.75 arcsec/pulse
Bias Repeatability	0.003 °/hr
Scale Factor Stability	0.2 ppm
Random Walk	<0.0015 °/√hr

Table 1 - ZLG Based Inertial Navigation System Gyro Performance Requirements

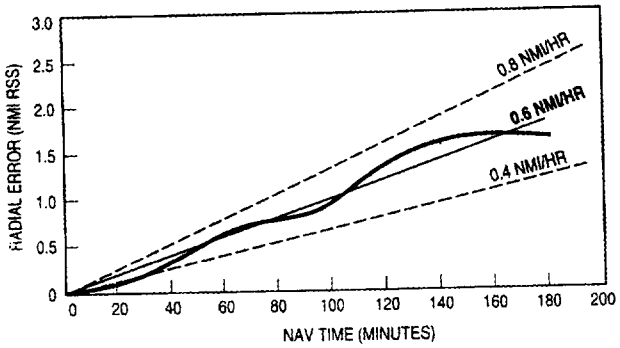


Figure-11 Typical Navigation Performance for ZLG Based Inertial System

The gyro parameters are key to achieving proper system performance as discussed below:

Dynamic Range

In a strapdown system, the sensors experience the full dynamic range of the vehicle motion. Thus, they must be able to operate at high angular rates (400 °/sec typical in fighter aircraft). At the same time, they must be able to accurately measure extremely low rates (0.005°/hr or less) in order to maintain accuracy of the navigation solution. The dynamic range required of a navigation grade strapdown gyroscope spans approximately nine orders of magnitude. The ZLG can easily support this requirement since the Faraday bias is chosen to be much larger than the expected maximum angular rate-induced frequency shift and since the ZLG has no locking effects. The dynamic range must also be maintained over a bandwidth wide enough to track rapid attitude changes. The ZLG, like all laser gyroscopes, is an optical sensor based on the relativistic Sagnac effect. The sensor has extremely wide bandwidth, essentially limited only by the sampling rate used to extract the rotation information.

Angle Random Walk

In the initial alignment of an inertial navigation system, the gyros are used to locate the azimuth (or heading) of the system relative to the Earth rate vector. The accuracy of the alignment (also known as gyrocompass alignment) depends on the accuracy with which angular rate can be measured in a given period of time. The angle random walk coefficient determines the uncertainty in the estimation of angular rate:

$$\sigma_{\Omega} = \frac{RWC}{\sqrt{T}}$$

Where σ_{Ω} is the uncertainty in the measurement expressed as a standard deviation, RWC is the random walk coefficient and T is the averaging time (alignment time).

This measurement uncertainty can be expressed as a heading uncertainty by dividing by the horizontal component of Earth rate.

$$\sigma_\psi = \frac{\sigma_\Omega}{\Omega_E \cos\Phi} = \frac{RWC}{\Omega_E \cos\Phi\sqrt{T}}$$

Where Ω_E is the magnitude of Earth rate and Φ is the latitude.

Gyro random walk also causes system errors in long flights as it causes attitude random walk.

As discussed in a previous section, the body dithered ring laser gyroscope experiences additional sources of noise due to light scatter and lock-in effects at dither reversal. The multioscillator ring laser gyroscope provides noise improvement and hence improved navigational accuracy with short ground alignment times.

Bias

Gyro bias errors lead to alignment errors and to navigation errors. Gyro bias is a key parameter in position error drift of an inertial navigation system. For a 0.8 nautical mile/hour system, gyro bias error must be below 0.005 °/hr. Since the ZLG is operated as a differential gyro, the optical bias (due to Faraday effect) cancels out. However, as discussed previously, the ZLG does exhibit other bias mechanisms related to dispersion and scatter. In order to minimize dispersion bias effects, highly effective cavity length and magnetic field control methods [26] have been developed. Scatter effects are minimized through the use of state-of-the-art multi-layer dielectric coatings. Residual bias effects are calibrated and modeled as a function of temperature. While the temperature bias models for a ZLG tend to be more complex than those for conventional RLGs, the bias residuals (i.e. unmodeled errors) are comparable for instruments of the same size. Navigation accuracy ZLGs reliably meet the required bias error specification.

Scale Factor

The gyro scale factor accuracy determines the ability of the navigation system to faithfully track vehicle rotational dynamics. For high performance aircraft, gyro scale factor accuracy of better than 5 ppm is required. Lower performance vehicles such as commercial airlines can tolerate somewhat larger errors. The ZLG, as previously discussed, offers excellent scale factor accuracy and linearity, substantially outperforming the conventional dithered RLG which suffers from locking effects.

Flight Control

In many applications, inertial sensors are needed to measure dynamic parameters used as observable in control systems. For example, angular rate and angular acceleration measurements are often required for flight control systems. Key requirements in these applications include bandwidth, latency (phase lag), noise, and drift rate. Resolution enhanced ZLGs are very well suited to these applications because of their very low noise (low quantization and no dither), their very high bandwidth (optical gyroscope with no

mechanical dither interface), and their low drift rates. ZLGs can also be hard-mounted to provide high-fidelity tracking of the structure dynamics. This is usually not possible with dithered RLGs.

Sensor Referencing/Motion Compensation

Synthetic aperture radar is a current application that requires both precise navigation information and compensation of the motion of the host vehicle [27,28]. A synthetic aperture radar image is formed by combining all the range and Doppler information from the returned radar pulses while keeping proper account of the relative position and velocity of the vehicle during the scene time.

An important aspect of the quality of the final image is related to how accurately the vehicle's motion is compensated. Poor motion compensation results in a distorted response function, which results in an uncertainty in the location of the image. The amount of disturbance of the image is related to the root mean square fit of the position error over the scene time relative to a constant velocity error. It can be shown that this disturbance is driven by the random walk of the gyros providing the motion compensation. The optically biased ring laser gyro, with its inherently low noise is a superior choice for this application.

Dither also produces sinusoidal motion of the sensor block. Aliasing of these frequencies sensed by the accelerometer to frequencies with periods near the scene time can lead to additional types of motion compensation errors. This type of error will cause high side lobes or false targets on either side of the actual target. The optically biased ring laser gyro, such as the ZLG, avoids the possibility of aliasing of sensor block dither reaction motions as well as avoiding induced instruments errors due to rectification effects, often referred to as sculling and coning errors.

Multi-Function Systems

Due to the extreme versatility of the ZLG, this instrument is ideally suited for use in multi-function inertial systems. These systems can provide not only position information but also precise velocity, attitude, angular rate, and acceleration information, thus satisfying many requirements with a single unit. Often, multi-function navigation systems will also include GPS receivers to aid the inertial system and bound the long term errors.

Space and Underwater Applications

In addition to the above mentioned performance improvements using the multioscillator ring laser gyroscope, the lack of dither motion makes this technology very attractive to space and underwater applications in which small motions imparted to the vehicle must be avoided. In addition, there is no acoustic noise generated by mechanical dither.

FUTURE DIRECTIONS

In addition to the applications discussed above, changing world politics and the advent of the Global Positioning System have resulted in a growing need for very compact guidance and control instruments for precision pointing and tracking, pinpoint strikes and accurate target discrimination. Requirements for these systems are not compatible with conventional dithered ring laser gyro technology and are too stringent in terms of random walk versus size for any projected fiber optics systems. The multioscillator technology as embodied in the ZLG is well suited to fulfill this need. Eliminating the dither motion from the laser gyro has demonstrated superior performance advantages. The elimination of the dither has also now made the three instruments in one monolithic block structure concept truly viable. Using the ZLG based technology it is now possible to fabricate three orthogonal ring laser gyros in a volume of glass about the size of a golf ball. This configuration would provide the lowest random walk performance per unit volume for any developed technology today, which includes both the fiber optic and the silicon micromachined. With this size device, precise platform stabilization can be accomplished for a large variety of applications such as infrared search and track, synthetic aperture radar and laser communications. This type of device could also be used to provide the "flywheel" for a GPS navigation system during satellite outages due to jamming effects or other natural blockages.

In summary, it is expected that the ZLG will continue to fulfil a role as an ideal sensor for multiple applications. Increasingly, the ZLG will be combined with GPS to permit the use of smaller, lighter systems capable of providing accurate wide bandwidth information from the inertial sensors while maintaining long term accuracy using the GPS.

References

[1] H. de Lang, "Eigenstates of Polarization in Lasers," Phillips Res. Repts 19, pp. 429 - 440, 1964.

[2] R. C. Jones, "A new calculus for the treatment of optical systems," J. Opt. Soc. Amer., vol 31, pp. 488-493, 1941 and vol 32 pp. 486-493, 1942.

[3] G. B. Yntema, D. C. Grant, Jr., and R. T. Warner, "Differential Laser Gyro System", U. S. Patent 3,862,803, January 1975.

[4] K. Andringa, "Laser Gyroscope", U. S. Patent 3,741,657, June 1973.

[5] J. Breman, H. Cook and D. Lysobey, "Differential Laser Gyro Development," Navigation: Journal of the Institute of Navigation, vol 24, No 2, pp. 153-159, 1977.

[6] D. Grant, J. Hambenne, T. Hutchings, V. Sanders and M. Scully, " A Multioscillator (Four-Mode) Ring Laser Gyro Report," Proceedings of the National Aerospace and Electronics Conference (NAECON), pp. 1029-1032, 1977.

[7] D. Grant, S. Madan, V. Sanders, D. Anderson, W. Chow and M. Scully, "Litton Ring Laser Gyro Update", AIAA Strategic/Tactical Missile Systems and Space Sciences Meeting pp. 73-79, 1980.

[8] V. Sanders, S. Madan, W. Chow and M. Scully, "Properties of a Zeeman Multioscillator Ring Laser Gyro," Proceedings of the IEEE National Aerospace and Electronics Conference, vol 1 pp. 2-6, 1979.

[9] J. Kilpatrick, "Laser Angular Rate Sensor," U. S. Patent 3,373,650, 1968

[10] I. Smith and T. Dorschner, "Electromagnetic Wave Ring Resonator," U. S. Patent 4,110,045, October 1980.

[11] T. Dorschner and I. Smith, "Clear-Path Four-Frequency Resonators for Ring Laser Gyros," J. Opt. Soc. Amer. Vol 68, pp. 1381, 1978.

[12] I. Smith and T. Dorschner, "Biasing the Raytheon Four-Frequency Ring Laser Gyroscope," Laser Inertial Rotation Sensors, SPIE vol 157, pp 21-29, 1978.

[13] T. Dorschner, I. Smith, H. Statz, "A Manganese Bismuth Magnetic Mirror for the Raytheon Laser Gyroscope," Proc. IEEE 1978 National Aerospace and Electronics Conference pp. 569-573, 1978.

[14] C. H. Volk, I. Longstaff, J. M. Canfield and S. C. Gillespie, "Litton's Second Generation Ring Laser Gyroscope," Proceedings of the 15th Annual Biennial Guidance Test Symposium, pp 493-502, 1991.

[15] E. Post, "Sagnac Effect," Reviews of Modern Physics, vol 39 pp. 475-493, 1967.

[16] H. Statz, T. A. Dorschner, M. Holtz, and I. W. Smith, "The Multioscillator Ring Laser Gyroscope," Laser Handbook, vol 4, North Holland Press, pp 229-332, 1985.

[17] W. Chow, J. Hambenne, T. Hutchings, V. Sanders, M. Sargent III, and M. Scully, "Multioscillator Laser Gyros," IEEE Journal of Quantum Electronics, vol QE 16, no 9 pp. 918-935, 1980.

[18] W. Chow, J. Gea-Banacloche, L. Pedrotti, V. Sanders, W. Schleich, and M. Scully, "The ring laser gyro," Reviews of Modern Physics, vol 57, no. 1, pp 61-104, January 1985.

[19] V. A. Markelov and A. A. Turkin, "Investigation of back reflections in a ring laser," Soviet Journal of Quantum Electronics, Vol 4, No 10, April 1975, pp. 1282-1283.

[20] A. Dorschner, H. A. Haus, M. Holtz, I. W. Smith and H. Statz, "Laser Gyro at Quantum Limit," IEEE Journal of Quantum Electronics, Vol QE-16, No. 12, December 1980 pp 1376-1378.

[21] Mechanically Dithered RLG At the Quantum Limit, S. W. Hammons and V. J. Ashby, NAECON Proceedings, 1981.

[22] D. A. Tazartes and J. G. Mark, "Apparatus and Method for Determining Ring Laser Gyroscope Phase Error at Turnaround," U. S. Patent 4,875,774, October 1989

[23] H. A. Gustafson, W. H. Lim and F. H. Zeman, "Ring Laser Lock-in Correction Apparatus," U. S. Patent 4,641,970, February 1987

[24] M. Fernandez, R. Ebner, N. Dahlen, "Zero-Lock Laser Gyro", Proceedings: 12th Annual AAS Guidance and Control Conference, Feb 4-8, 1989, Keystone, Colorado

[25] J. G. Mark and D. A. Tazartes, "A Resolution Enhancement Technique for Laser Gyros," Proceedings of 4th International Conference on Integrated Navigation Systems, May 26-28, 1997.

[26] "Multi-oscillator ring laser gyroscope adaptive digitally controlled cavity length control system," U.S. Patent 5,208,653, May 4, 1993.

[27] J. Lipman, "Millimeter Navigation Performance for High Quality SAR Imaging and Target Location," Proceedings of the Precision Strike Technology Symposium, October 9-10, 1996.

[28] J. Lipman and R. Cox, "Tradeoffs and Technical Issues in the Integration of Modern Navigation Systems and Synthetic Aperture Radars," Proceedings of the Precision Strike Technology Symposium, September 27-28, 1995.

5. ZEEMAN LASER GYROSCOPES

V.V. Azarova, Yu.D. Golyaev, V.G. Dmitriev, M.S. Drozdov, A.A. Kazakov, A.V. Melnikov, M.M. Nazarenko, V.N. Svirin, T.I. Soloviova, N.V. Tikhmenev

5.1 INTRODUCTION

5.2. ZEEMAN EFFECT AND BASIC PRINCIPLES OF THE LASER GYROS ON THIS EFFECT

5.3. OUT-OF-PLANE CAVITY. POLAZITAION CHARACTERISTICS. CAVITY MODES

5.4. ENERGY COUPLING AND LOCK-IN EFFECT FOR COUNTER-PROPAGATING WAVES IN THE RING LASER WITH THE OUT-OF-PLANE CAVITY

5.5. ERROR MODEL OF ZEEMAN LASER GYRO AND METHODS OF ERROR DECREASING

5.6. ZEEMAN LASER GYRO DESIGN

5.7. ZEEMAN LASER GYRO BEHAVIOR UNDER SEVERE ENVIRONMENT

5.8. APPLICATION OF ZEEMAN LASER GYRO

5. ZEEMAN LASER GYROSCOPES

V.V. Azarova, Yu.D. Golyaev, V.G. Dmitriev, M.S. Drozdov, A.A. Kazakov, A.V. Melnikov, M.M. Nazarenko, V.N. Svirin, T.I. Soloviova, N.V. Tikhmenev

Research & Development Institute "Polyus"
3 Vvedensky str, 117342 Moscow, Russia

• 5.1. INTRODUCTION

The possibility of rotation parameters measuring by means of a ring laser was shown in several works [1, 2, et al].

Zeeman laser gyro with nonplanar resonator (ZLGN) is one of the many different types of the laser gyros. ZLGN is based on ring laser with out-of-plane resonator and magnetooptical biasing on Zeeman effect. Out-of-plane resonator provides the stable existence for the waves with right and left circular polarizations and reciprocal (not dependent on propagating direction) splitting of frequencies for orthogonally polarized waves. Magnetic field applied directly to active medium causes nonreciprocal splitting of the of the counter-propagating waves frequencies owing to Zeeman effect. That is why there are not only two frequencies modifications, but also four frequencies modification of ZLGN.

Sometimes such devices are called "motionless laser gyros with clear-path resonator". The main advantages of such ZLGN in compare with other laser gyros are the following:
1) absence of any moving parts (such as dither mechanism);
2) absence of any intracavity components (such as quartz) for creation of the circular polarized waves;
3) absence of any intracavity components (such as Faraday cell) for creation of biasing;
4) increased stability of out-of-plane resonator to the manufacturing errors and disadjustment in compare with a planar resonator;
5) easy control of a biasing shape by the magnetic field.

Because of advantages (1) - (4) ZLGN has an increased stability to mechanical influences, short ready time, stable operation in the wide temperature range.

Owing to feature (5) the output characteristic of ZLGN can be effectively linearized by the choosing of optimal shape of the current in the coil exiting a magnetic field for biasing (rectangular, triangular, "swallow tail").

The concept of ZLGN with out-of-plane resonator was suggested firstly by Russian scientists of R&DI "Polyus" in 1967 [3]. Later this concept was researched by American scientists [4, 5, 6].

About 30 years the scientists and the engineers of R&DI "Polyus" were developing theory, design and technology of ZLGN sensors and created family of so-called ZLK devices: ZLK-28, ZLK-20, ZLK-16, ZLK-12 (Fig. 5.1). Some members of this family are produced at plant, some of them are

under experimental production. The results of tests in laboratory and in flights proved the advantages of ZLK. Also some problems were found [7]. One problem caused by the higher losses at the mirrors for the circular polarization in compare with linear polarization. It required higher gain of active medium that led to bigger errors connected with interaction of laser irradiation with active medium.

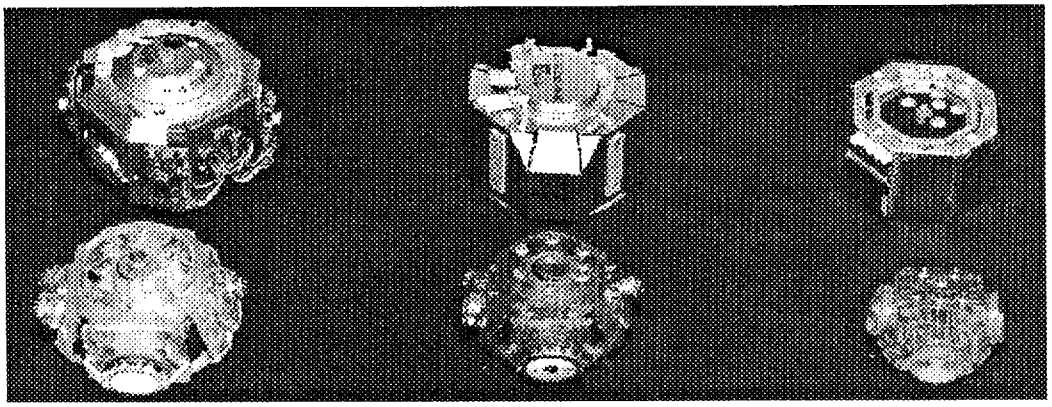


Fig. 5.1. Zeeman laser sensors ZLK-20, ZLK-16, ZLK-12.

Other problem was inherent high sensitivity to magnetic field due to magnetooptical type of the biasing. It decreased the accuracy in conditions of high external magnetic fields.

Specialists of R&DI "Polyus" were investigating these problems. To decide it and to reach the required accuracy the improved mirrors technology and the optimal shielding were developed. Specific algorithms with using of Zeeman laser intrinsic information about internal temperature (algorithms "correction on biasing") were developed to get the effective compensation of the errors.

As a result R&DI "Polyus" from 1980 regularly produces several modifications of ZLGN at the plant and applies it in the systems of stabilization, orientation and navigation.

Now the perspective ZLGN with smaller size and better performance are being designed and manufactured. Besides the unique out-of-plane solid-state YAG:Nd ring lasers (Fig. 5.2) with magnetooptical biasing are being investigated and they are expected to become a new generation of the miniature gyros [8-12, 13].

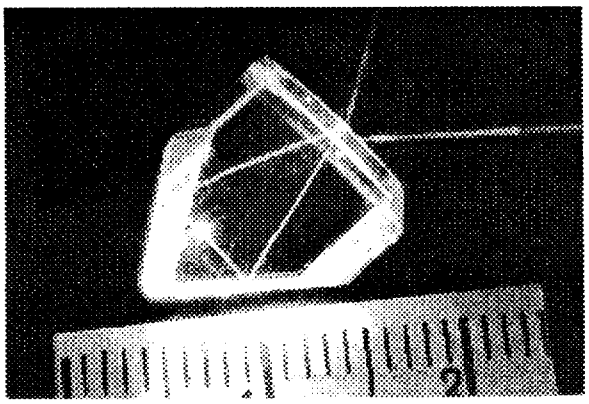


Fig 5.2. Nonplanar solid-state YAG:Nd ring laser.

• 5.2. ZEEMAN EFFECT AND BASIC PRINCIPLES OF THE LASER GYROS ON THIS EFFECT

Gain curve Zeeman splitting

Zeeman effect was discovered in 1896 by Holland physicist Zeeman. There are two types of this effect: longitudinal and transversal ones. Both types represent energy levels splitting of some optical gas medium being affected by magnetic field [14, 15]. Due to levels splitting initial optical emission lines also become splitted. For longitudinal effect the splitting is double. Each subline corresponds possible optical transitions between the magnetic sublevels. For the Ne line of the He-Ne laser 0.6329 μm the level diagram is presented in Fig. 5.3. Finally additional selection rule $\Delta m=1$ determines the allowed optical

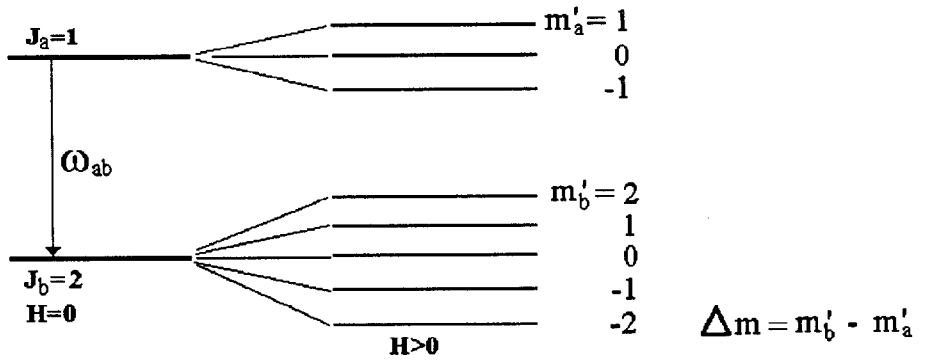


Fig. 5.3. Energy levels of diagram of 0.63 μm lines of the He-Ne laser in the absence and presence of the magnetic field. J_a, J_b - angular momentum of levels, m_{ab} - angular momentum magnetic field components.

transition, where Δm is the changing of angular momentum. For the transversal effect the selection rule is $\Delta m=0; \pm 1$ and the initial emission line is splitted on three lines (Fig. 5.4). The value of splitting (frequency difference between nonsplitted and neighboring splitted lines) is determined by expression:

$$\delta\nu_D = \frac{e \cdot g}{4 \cdot \pi \cdot \mu \cdot c} \cdot H \qquad (5.1)$$

where $e=1.6 \cdot 10^{-19}$C - electron charge, $\mu= 0.91 \cdot 10^{-30}$ kg - electron mass, $c=3 \cdot 10^{10}$cm/s - light velocity, H - magnetic field strength (Oe), g=1.3 - Lande factor. By using these quantitative physical constants we have the value of $\delta\nu_Z = 1.8 \cdot 10^6 \cdot H$

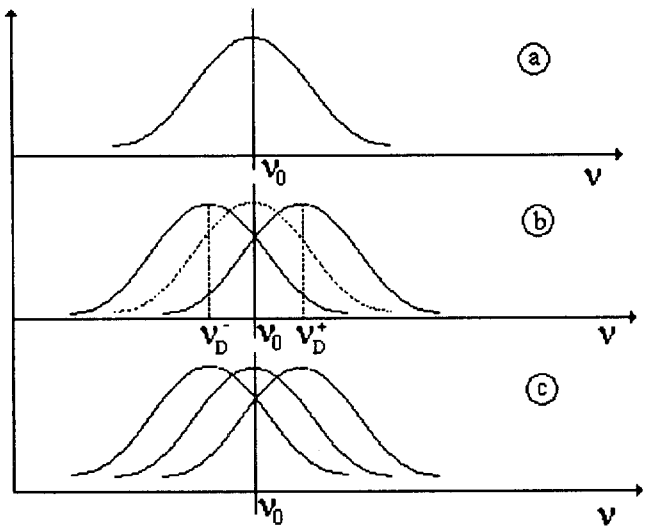


Fig. 5.4. Diagram of Ne emission line splitting in He-Ne laser by applying of magnetic field: a - nonsplitted line without magnetic field, b - splitting under longitudinal magnetic field, c - splitting under transversal magnetic field; ν_0 - center of nonsplitted Doppler line.

Because of longitudinal effect we have two gain curves (Fig. 5.4), separated by frequency distance (Zeeman splitting), given by expression:

$$\left(\nu_D^+ - \nu_D^-\right) = \Delta\,\nu_Z = 2\cdot\delta\cdot\nu_D = 3.6\cdot 10^6 H \qquad (5.2)$$

Each curve is capable of spontaneous or stimulated emission only one of the two possible circular polarization light, namely: left curve (low frequency) for the left circular polarization (LCP) looking along the magnetic field **H** and right curve (high frequency) for right circular polarization RCP.

Zeeman biasing

To achieve biasing effect in the laser with Zeeman method it is necessary to provide inside active medium the waves with opposite circular polarization LCP and RCP with respect to the magnetic field direction. That means in laser ring cavity it should be generated two counter-propagaiting waves with the same circular polarization states LCP or RCP (or both) looking along their wave travel vectors. For this purpose the out-of-plane optical resonator is used which creats the circular polarization waves and simultaneously causes the reciprocal frequency splitting.

Firstly we will consider two frequency ring laser. When the magnetic field is absent we have nonsplitted gain curve (Fig. 5.5). The two counter-propagating waves have the same cavity frequencies $\nu_{CW}=\nu_{CCW}$ usually tuned to the gain curve center ν_o and opposite (LCP and RCP) polarizations, looking along vector **H**.

This initial gain curve is splitted while the magnetic field is on and we have two gain curves low frequency (ν_D^-) and high frequency $\left(\nu_D^+\right)$.

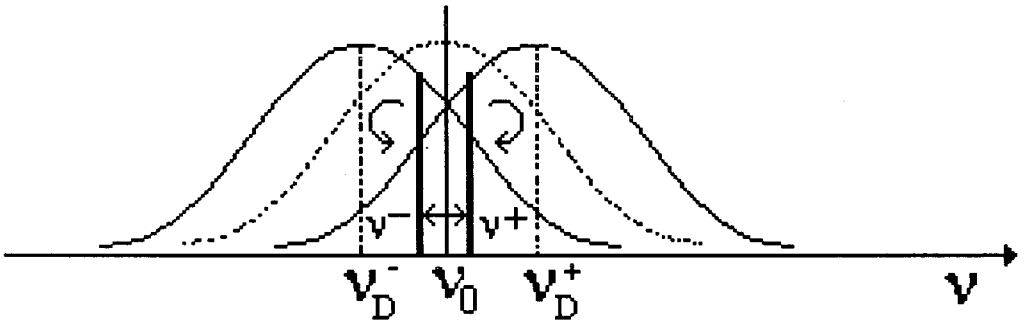


Fig. 5.5. Zeeman counter-propagaiting waves splitting: ν_o - modes frequency without magnetic field, ν^-, ν^+ - splitted counter-propagaiting modes frequencies while magnetic field is on.

Each curve is effective only for one of the two states of circular polarizations LCP or RCP (relatevely magnetic field direction). Because the counter-propagating waves have opposite circular polarizations respectively vector H they have acquired separate own gain curves ν_D^- or ν_D^-. But the centers of these curves due to Zeeman splitting already do not coincide with the initial frequency ν_o of the waves. Due to well known phenomenon "mode frequency pulling" the real generating mode frequencies shift toward the corresponding gain curve center [16].

The new frequencies of the waves ν_{CW}, ν_{CCW} are equal:

$$\nu^{\pm} = \nu_C + 0.94 \cdot \frac{\Delta\nu_C}{\Delta\nu_D} \cdot \frac{G_0}{\sigma} \cdot \left(\nu_D^{\pm} - \nu_C\right), \tag{5.3}$$

where $\nu^{\pm}$ - possible values of counter-propagating waves frequencies ν_{CW}, ν_{CCW} depend on magnetic field direction; ν_C -initial mode frequency, usually $\nu_C=\nu_O$, $\Delta\nu_C$, $\Delta\nu_D$ - bandwidths of resonator mode and laser emission line, G_o, σ - gain and optical losses coefficients of the cavity. The difference between frequencies (ν_{CW} - ν_{CCW}) actually is a value of biasing of the laser gyro and by using 5.2, 5.3 it is equal:

$$\nu_{CW} - \nu_{CCW} = \nu_B = 0.94 \cdot \frac{\Delta\nu_C}{\Delta\nu_D} \cdot \frac{G_0}{\sigma} \cdot \left(\nu_D^+ - \nu_D^-\right) \tag{5.4}$$

By using (5.2) the formula (5.4) is transferred to the expression:

$$\nu_B = 2.63 \cdot 10^6 \cdot \frac{\Delta\nu_C}{\Delta\nu_D} \cdot \frac{G_0}{f} \cdot H \tag{5.5}$$

For typical ZLGN the values of Zeeman splitting and biasing are equal ($\Delta\nu_C$=1.5 MHz, $\Delta\nu_D$=1.5 GHz, G_O/σ=2, H=20 Oe):

$$\Delta\nu_Z = \left(\nu_D^+ - \nu_D^-\right) = 7.2 \cdot 10^7 Hz, \qquad \nu_B = 0.94 \cdot \frac{1.5 \cdot 10^6}{1.5 \cdot 10^9} \cdot 2 \cdot 5.6 \cdot 10^7 = 135\text{kHz}.$$

For the four frequency ring laser the biasing method is similar to the two frequency ring laser. For four frequency case the out-of-plane cavity must have such fold angle (as we can see later in the part 5.3.) that two cavity modes with right and left circular polarization should be as close as possible to each other but without interrection between them through the active medium. In the longitudinal magnetic field they are splitted due to Zeeman effect in two frequenses for clockwise and counter clockwise pathes. For increasing the range of existence of four frequency regime a special mixture of three gasesis used: Ne^{20}, Ne^{22}, He^{3}. For this gas mixture the gain curve bandwidth is equal $\Delta\nu_D$ =2.3 GHz. The scheme of frequency biasing for the four frequency ring laser is shown on Fig. 5.6.

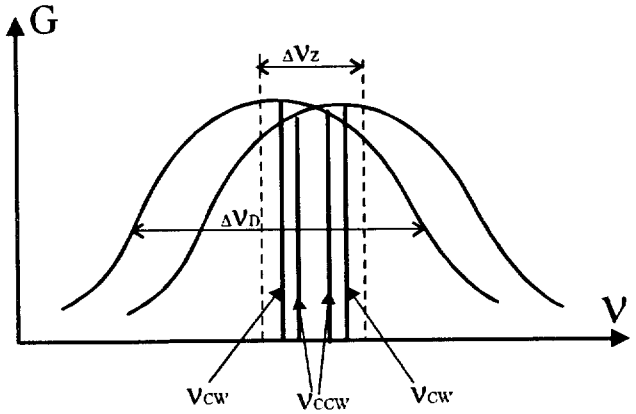


Fig.5.6. The scheme of Zeeman biasing for four frequency ring laser sensor ZLK-28.

More detailed description of ZLGN amplitude-frequency characteristics with meaning of the real working regimes (high gain, self absorption and other) was written in [17-24].

- **5.3. OUT-OF-PLANE CAVITY. POLAZITAION CHARACTERISTICS. CAVITY MODES**

The main principle of ZLGN operation is the same as for any laser gyro based on a general scheme of four-mirrors ring laser (Fig.5.7). At the same time ring cavity in ZLGN is out-of-plane, that is why it has the specifical mode spectrum. Frequency degeneracy for own oscillation modes with the opposite directions of rotation of the polarization plane is absent in such cavities. The calculation of the spectrum of own frequencies of out-of-plane cavity based on the matrix equation is given in [25]:

$$M \cdot \vec{E} = \gamma \cdot \vec{E}, \qquad (5.6)$$

where $M=\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$ is a matrix describing the effect of all elements of the cavity on a light wave circulating in the cavity; $\vec{E}=\begin{pmatrix} E_x \\ E_y \end{pmatrix}$ is an electric vector of the light wave that includes an amplitude and a phase; γ is a factor of the matrix M.

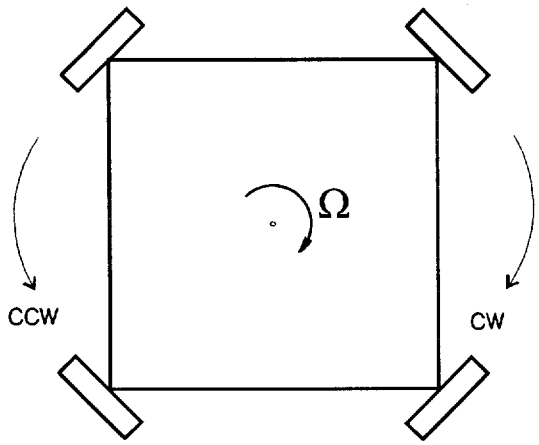


Fig. 5.7 The four-mirrors ring laser gyroscopic sensor rotating with an angular rate Ω.

The meaning of γ is simple: this is a factor describing the change of the amplitude and the phase shift of the light wave during its circulating in the cavity. For natural frequencies of the cavity the phase shift should be multipled by 2π. The matrix equation (5.6) is a system of two equations for projections of the vector $\vec{E}(E_X, E_Y)$ on the X, Y axes, while the Z axis is coincident with the direction of circulating in the cavity. This system of equations has two solutions that correspond to two modes with orthogonal polarizations of the electric field:

$$\gamma_{1,2} = (1/2)\mathrm{SpM} \pm \sqrt{4\mathrm{SpM} - \mathrm{DetM}}, \tag{5.7}$$

$$|E_x/E_y|_{1,2} = M_{12}/(\gamma_{1,2} - M_{11}). \tag{5.8}$$

For out-of-plane cavities the expression under the root is always negative ($4SpM\text{-}DetM < 0$), due to difference in phase shifts and correspondingly in frequencies between modes with different polarization states (because of the signs "+" and "-" before the root).
The matrixes of full round of out-of-plane cavity with ideal mirrors for clockwise, M^{cw}, and counterclockwise directions, M^{ccw}, are equal and equivalent to the matrix of rotation of axes through a fold angle ρ_Σ:

$$M^{ccw} = M^{cw} = \begin{bmatrix} \cos(\rho_\Sigma) & -\sin(\rho_\Sigma) \\ \sin(\rho_\Sigma) & \cos(\rho_\Sigma) \end{bmatrix} = S(\rho_\Sigma), \tag{5.9}$$

where ρ_Σ is a full angle of rotation of the image while full round of the cavity. By substituting (5.9) in (5.7 and 5.8) we obtain $\gamma_{1,2} = \exp(\pm i\rho_\Sigma)$, and $|E_x/E_y|_{1,2} = \exp(\pm i\pi/2) = \pm i$. Hence for out-of-plane cavities natural oscillation modes has RCP and LCP independently on a fold angle of the

cavity, and the phase shift between them is equal to the double angle of rotation of axes in full round of the cavity, $2\rho_\Sigma$.

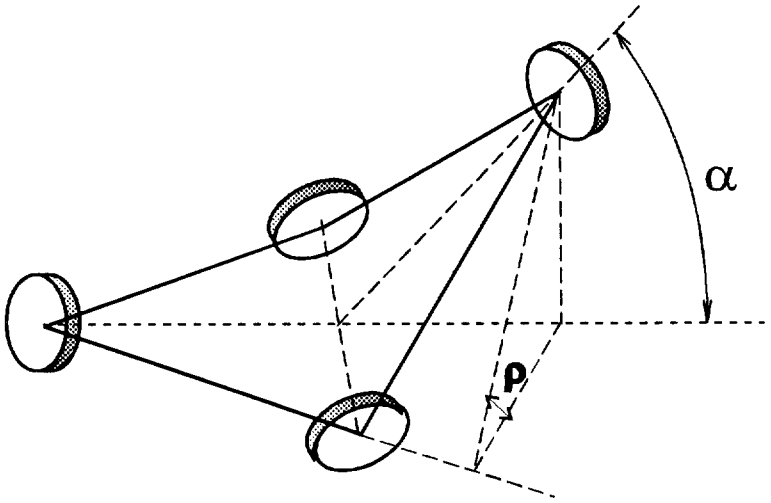


Fig.5.8. The scheme of the out-of-plane cavity with a fold angle.

Depending on a fold angle of out-of-planarty α (Fig.5.8) and correspondingly on a value ρ_Σ the resonance frequency shifts for RCP and LCP change, that is, the cavity spectrum also changes. By changing the fold angle a cavity with a required frequency spectrum can be designed. Forming of the frequency spectrum in cavities for one of the directions is illustrated in Fig.5.9.

In case of two-frequency Zeeman lasers, the major problem is to achieve a maximum possible frequency separation for waves with orthogonal polarizations for operation only at one of them. This condition is provided by the equidistant spectrum of all cavity modes. In turn, such mode spectrum is provided by a special cavity design, where the rotation of the image while round of the cavity is $\rho_\Sigma = 90°$. For the two-frequency sensors ZLK-12, 16, 20 the image rotation $\rho_\Sigma = 90°$ is provided by the fold angle $\alpha = 32°$ and the angle between two neighbor incidence planes $\rho = 22.5°$ (Fig. 5.8).

With designing four-frequency laser gyros, the problem arises consisting in optimization of frequency separation between modes with RCP and LCP to get the four-frequency mode. In Fig. 5.9 the spectrum of natural frequencies of the ring cavity with the fold angle $\alpha \approx 57°$ and the total angle of rotation $\rho_\Sigma \approx 189°$ is shown.

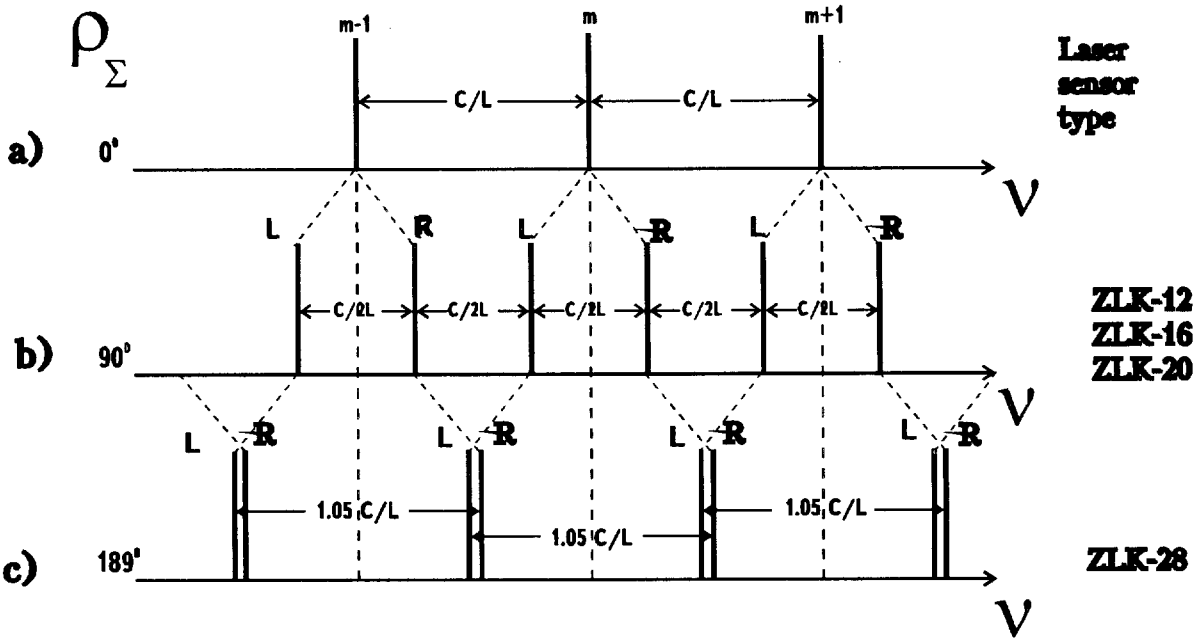


Fig.5.9. The spectrum of natural frequencies of out-of-plane cavities as a function of a total angle of rotation ρ_Σ and types of sensors with out-of-plane cavities: a) a cavity with an equivalent plane contour, b),c) out-of-plane laser cavities: b) two-frequency ring lasers, c) four-frequency ring lasers.

- **5.4. ENERGY COUPLING AND LOCK-IN EFFECT FOR COUNTER-PROPAGATING WAVES IN THE RING LASER WITH THE OUT-OF-PLANE CAVITY**

The physical phenomenon of energy coupling and lock-in effect in the ring laser gyro is well known. The ring laser with the out-of-plane cavity has some specific features. These features are connected with the circular polarization of the counter-propagating waves and we consider its here.

Generally, energy coupling between two counter-propagating waves is based on the backscattering of these waves. For each cavity direction we have two waves: the main wave and the wave backscattered from the opposite wave. If the polarization states of the main and backscattered waves are the same these waves interfere and through this mechanism both main waves have energy coupling [26]. Due to this coupling the counter- propagating waves frequencies are shifted toward each other. For small rotation rate the phenomenon "lock-in" occurs, when the frequencies are sticked together. If the polarization states of the main and backscattered waves are orthogonal or oppositly circular no interference take place and consequently no energy coupling and no lock-in take place.

For two-frequencies out-of-plane cavity two counter-propagating laser waves are generating. These waves have the same LCP (or RCP) polarization looking along wave travel vectors [25]. For ideal mirrors surfaces backscattering waves would have opposite circular polarization comparing main waves. At that case the laser gyro have no lock-in effect. But in real situation the cavity mirrors are not ideal and backscattered waves consist of mixture of both polarization states: LCP and RCP. So, possible polarization discoupling mechanism is working only partially. In any case the counter-propagating waves coupling in out-of-plane ring laser exists and lock-in effect takes place. Due to this effect output frequency is equal to zero even though laser rotates within relatively wide rate region (0...1000 deg/h), that is in this region laser gyro can not measure the rotation. To avoid this

negative phenomenon the introduction of externally controlled null shift or "biasing" is using. But the value of biasing also can drifts. To avoid biasing drift instead constant biasing the alternating biasing is used. In this scheme the biasing is operated in both the positive and negative rotation rate direction. Over each of these cycles the net biasing averages to zero, so that the output of the gyro, being intrinsically an integrated magnitude should only reflect the net rotation rate [26, 27].

There are two alternating biasing techniques: the mechanical one and magnitooptical one. The first technique is based on real mechanical vibration of laser body - "dither" (or on reverse rotation of laser body - "rate biasing"). The second one is based an Faraday, Kerr or Zeeman effect.

Both techniques have their own advantages and disadvantages. The main disadvantage of the laser gyros with the magnetooptical biasing is enough high sensitivity to external magnetic field. To reduce the sensitivity the several methods are used: at first magnetic shielding, then active compensation systems and finally four frequency oscillation regime of laser. The essential advantage of the magnetooptical biasing is a possibility to create practically any timeshape of alternating biasing, and so manner to improve the accuracy of the laser gyro.

• 5.5. ERROR MODEL OF ZEEMAN LASER GYRO AND METHODS OF ERROR DECREASING

ZLGN errors, caused by coupling through backscattering

The counterpropagaiting waves coupling phenomena caused by backscattering is well known. The methods of its reduction are being investigated and advanced continuously [26-29, 31-33]. The basic effects of counter-propagaiting wave coupling in ZLGN are the same as in usual laser gyro with linear polarization and with the dither biasing. The same parameters and the same analytic formulas describe the input-output characteristic of ZLGN [34, 35]. In range of rotation rates nearest to magnitude of alternative biasing $|\Omega - 2\pi f_{\Pi}\Omega| >> \Omega_L$ the nonlinearity and the nonstability of scale factor are defined by coefficients of the backscattering and the cavity dynamic bandwidth:

$$\Delta K_M \sim \frac{R_e(\sigma_1\sigma_2)}{(\Delta\nu_d)^2}, \qquad (5.10)$$

where σ_1, σ_2 - are the complex coupling coefficients; $\Delta\nu_d$ - is the cavity dynamic bandwidth for the contrary waves intensities difference.

In range of rotation rates far away from alternating biasing magnitude ($|\Omega - 2\pi f_{\Pi}\Omega| \sim \Omega_L$) the nonlinearity and the nonstability of scale factor are defined by relation of lock-in to biasing magnitude:

$$\Delta K_M \sim \frac{\Omega_L}{2\pi\nu_B}, \qquad (5.11)$$

where Ω_L - lock-in zone, ν_B - magnitude of alternating biasing.

Advantage of the laser gyro with magnitooptical biasing is the possibility to create any shape of biasing. Special biasing form can reduce significantly the scale factor errors [38].

It is well known [35- 37] that alternating biasing laser gyro input-output characteristic has the dynamic synchronization zones (see Fig. 5.10).

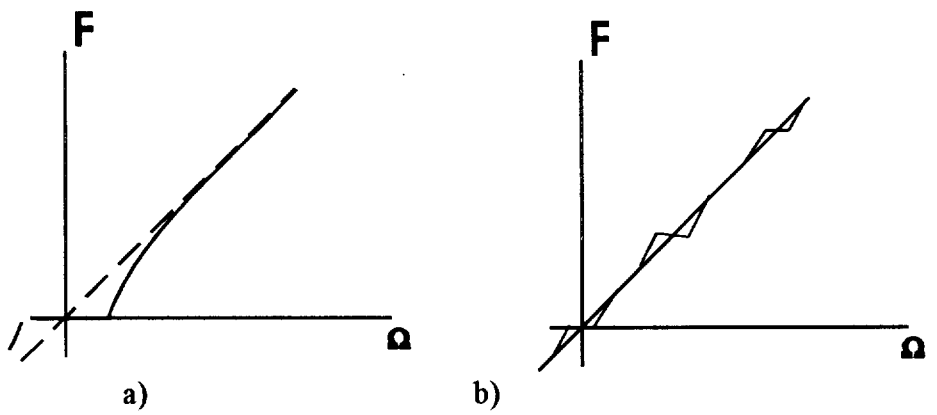


Fig. 5.10. Input-output characteristics of the ring laser gyro. As a dashed line, the ideal characteristic is shown. As a continuous line: a) with coupling between counterpropagating waves, b) with alternating biasing.

The dynamic zones values depend strongly on alternating biasing shape. For the laser with linear polarization of cavity radiation and with dither biasing the value of maximum dynamic zone is defined by formula:

$$\Omega_{dL} = \frac{2\Omega_L}{\sqrt{2\pi\nu_B \cdot T_K}}, \tag{5.12}$$

where T_K - is the period of biasing.

For ZLGN with rectangular reversive biasing the dynamic zone value is:

$$\Omega_{dZ} = \frac{\Omega_L}{\pi\nu_B \cdot T_K} \tag{5.13}$$

The dynamic synchronization zone errors can be reduced by additional random or periodic small frequency biasing so-called desinchronization. In any case it will be as effectual as less initial value of dynamic zone. Even if the desinchronization is effectual so-called random drift remains, when the rms phases difference increases as a square root of the time $\langle\varphi\rangle = D\sqrt{t}$.

The random drift depends on the derivative of current value of biasing rate (or angular acceleration in the point of the rate sign changing):

$$D = \frac{\Omega_L}{\sqrt{T_K \cdot 2\pi \frac{d\nu_B}{dt}(\nu_B = 0)}} \tag{5.14}$$

where $2\pi \frac{d\nu_B}{dt}(\nu_B = 0)$ - the current value of biasing rate in the point of the sign changing.

It is clear from formulas (5.12 - 5.14) that it is important to provide the short front of biasing sign changing. The magnitooptic biasing permits to reduce he dynamic zones and random drift (due to short edge) and at the same time to have the complete shape of biasing for reducing the scale factor nonlinearity [38].

ZLGN errors, caused by current drift (Langmuir drift)
One of the main source of zero shift in ZLGN is so called current drift. The current drift is caused by moving of gas mixture atoms along the discharge channels of ring laser. There are a lot of methods used for compensation of this influence [31, 39, 40]. However, these methods are not ideal and in any case some current drift influence remains. The value of this influence depends on laser body temperature. The further reducing of the current drift value can be achieved by means of algorithmic correction based on temperature. This idea requires a system of temperature sensors located in different parts of the ring laser. The algorithmic correction will use the information from those sensors for zero shift compensation.

In ZLGN there is a possibility to use a simplier way to gather the necessary information for correction instead of using temperature sensors system. This way is to use the biasing frequency value ν_B measured just right after ring laser turn-on as a data for correction of a component of zero shift caused by current drift.

This possibility based on correlation between the current drift and the biasing frequency value of ZLGN. The reason for this is the identical physical mechanism of current drift and Zeeman biasing creation. As it was mentioned before the current drift is caused by a direct movement of exited atoms of the gas mixture. This direct movement shifts the Zeeman gain curves for counterpropagating waves to the one direction. This shift gives an additional linear pulling of coutureprepagating waves frequencies. So, the current drift appears. The same mechanism of the gain curves shift works while Zeeman biasing creation. That's why the dependencies of the current drift vs temperature and biasing frequency vs temperature have a good correlation.

On practice there is a little deviation of this correlation from theoretical, because of mechanical deformations of ring laser contour while temperature changings. This deformation can cause the laser beam displacement in the channels which changes the value of gas atoms effective speed. The possible inflation of ring laser contour deformation can be taken into account as a nonlinear part in the formula (5.15).

Summing up the all mentioned facts, the value of current drift can be found from Zeeman biasing value by formula:

$$\nu_C = \nu_{C0} + c_C \nu_B + \nu_{CN}(\nu_B), \qquad (5.15)$$

where ν_{C0} - constant part of current drift, $c_C \nu_B$ - linear part of current drift, where c_C - the coefficient of proportionality of current drift and ν_B biasing frequency value, ν_{CN} - nonlinear part of current drift.

The experimental results showed that the nonlinear part always presents, but its value significantly lower than linear part. The experimental results of using the biasing frequency value for algorithmic correction of ZLGN are shown in part 5.7.

• 5.6. ZEEMAN LASER GYRO DESIGN

Two-frequency sensors

The family of two-frequencies Zeeman laser sensors (Fig. 5.1, 5.11, 5.14) are produced at the base of the same design principles and on the common technology.

The ring resonator is formed by the four mirrors, three of them are flat and fourth one is spherical. Two flat mirrors placed at the same diagonal have a special shape, and they are connected with piezodrivers for the optical path control.

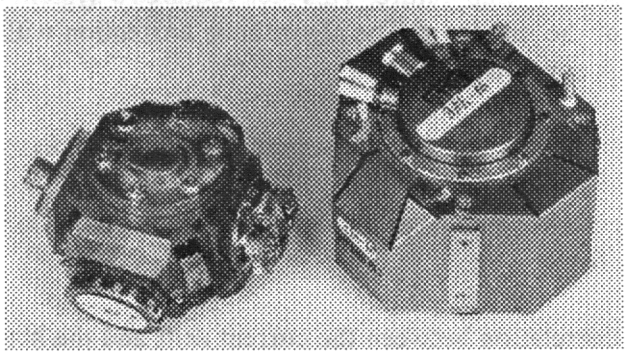


Fig. 5.11. Two frequencies Zeeman laser sensor

The resonators have a symmetrical out-of-plane configuration. They have perimeter 20 cm (ZLK-20), 16 cm (ZLK-16), 12 cm (ZLK-12). The resonator body is made from citall (analogue of Cer-Vit) with extra low thermal expansion coefficient. The resonator body is the base of the design. In the body the channels for the active medium are drilled with diaphragm in one of the channels for the axial modes selection. The base surfaces for the mirrors are oriented so that the angle between two neighbor incident planes is equal 22.5°. To apply the optical contact for the mirrors placing the surfaces are processed by extrapolishing. The cold cathode is mounted at the center of the resonator body. Two anodes, two burning electrodes and exhaust tube are sealed in citall disk which is placed at the resonator body by the optical contact. Internal volume of the resonator is filled by the gas mixture He^3 and Ne^{20}. Gas discharge is exited in four legs of resonator, and the current direction in two legs is reversed in compare with neighbor ones. Such design has a high symmetry. The counter-propagating waves are combined at the output of the resonator with the combining prism. The special two-plates photoreceiver transfers the optical signals to two electrical sinusoidal signals with information on the rate and direction of the rotation. Another photoreceiver registers the intensity of one of the counterpropagating beam and operates in the system of the optical path control. The biasing is formed by the applying of the sign-reversed magnetic field to the active medium with help of four identical coils. The resonator is mounted at the frame by the flat spring. The μ-material shield protects the sensor from the magnetic fields.

In 3-axis ZLGN three sensors are firmly mounted at the same frame (Fig. 5.12) so that their rate sensitivity axes are mutually orthogonal. The service electronics of sensors are placed at the same

frame, such as the power supply, the discharge current and burning current stabilizers, the biasing control system, the optical path control system, the synchronizer of all electronics, the processor of output data, ROM for memorizing of the correction coefficients for each sensor.

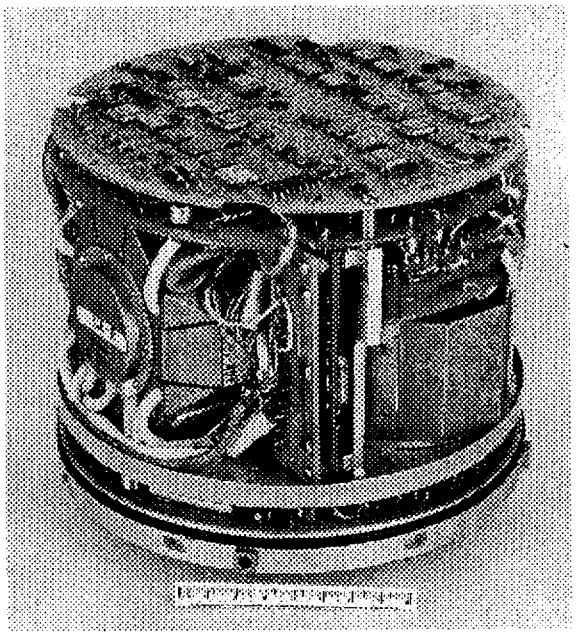


Fig. 5.12. Three-axis ZLGN

All assembly is surrounded by the hermetic case which is simultaneously two-layers shield. Internal volume is filled with the dry nitrogen

.

The common source of voltage about 800 V and six stabilizers are used for creation of dc discharge in active medium.

To reduce the discharge initiating time the additional electrodes are placed between cathode and anode. The pulse of high voltage is applied to the additional electrodes when gyro is turned on.

The alternating biasing is created synchronously in all sensors by the passing of the current through the coils connected in series. The shape of the alternating symmetrical current pulses is close to rectangular. High frequency of the switching of the current sign (1 kHz) determines the short time of delay (less 1 ms) of output information about the rate changes. The small inductance of the coils permits to switch the biasing current very fast (the pulse edge duration is less 40 ms), and it permits to reach very small dynamic synchronization zones.

The principle of the optical path control system operation can be seen from Fig. 5.13.

While the magnetic field is applied to the active medium the gain line is shifted on frequency and the generation frequency does not coincide with the center of the gain line. Correspondingly the generation power is decreased. When the magnetic field sign is changed the gain line jumps to other side and the resonator mode passes to another slope of the gain line. If the resonator frequency coincides with the center of the nonshifted gain line (without the magnetic field) and gain line is symmetrical, the switch of the magnetic field sign will not lead to change the generation power. But

if the resonator frequency is shifted (detuned) relatively the gain line center, the switch of the magnetic field sign will cause to the changes by jump of the generation power in one beam. If the magnetic field is varying in time on rectangular law with the frequency ν, the intensity modulation depth in one beam is proportional to the resonator detuning from the center of the nonshifted gain line, and the modulation phase shift relatively the magnetic field phase (0 or π) determines the detuning sign. This signal of detuning permits to build the optical path control system on the base of a classical scheme with a synchronous detector, an integrator and an output amplifier. The input signal of the synchronous detector is the photoreceiver signal registering the intensity in one beam, and the reference signal is the rectangular logical signal which is synchronous with the magnetic field at the active medium. The output amplifier is connected with two piezodrivers joined in parallel.

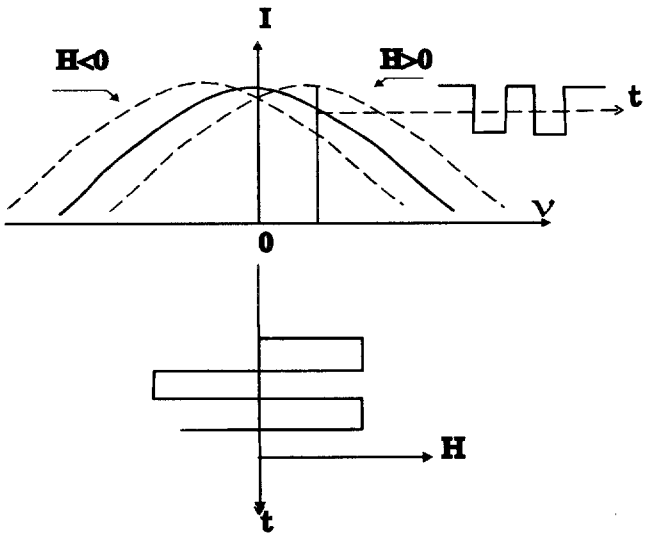


Fig. 5.13. The principle of the optical path control system operation:
1 - the generation power in one beam; ν - the generation frequency detuning from the nonshifted gain line center; t - time; H - magnetic field at the active medium (biasing magnetic field)

If necessary the biasing system is supplied with a desynchronizer which adds the noise signal with the small amplitude and quasi-white spectrum. The desynchronizer destructs the dynamic synchronization zones on the input-output characteristic, but with simultaneous small increasing of the noise component in the output signal.

The processor of output data transfers the sensors sinusoidal signals to the pulse trains with the duration equal to the biasing pulse duration and frequency corresponding the rate.

ROM is intended for the memorizing of the individual sensors data: scale factor, rate sensitivity axes misalignment, constants of functions describing the zero shift dependence on output temperature, on time and on internal sensor parameters, such as the biasing value, voltage at the piezocorrector etc.

Four-frequency sensors

As well as two-frequencies resonator four-frequencies one is formed by the four mirrors, three of them are flat and fourth one is spherical. Four-frequencies resonator has the fold angles of the out-of-plane optical path which permit the existence of RCP and LCP waves in both directions, totally four frequencies (see [25, 30, 44] and Fig. 5.14).

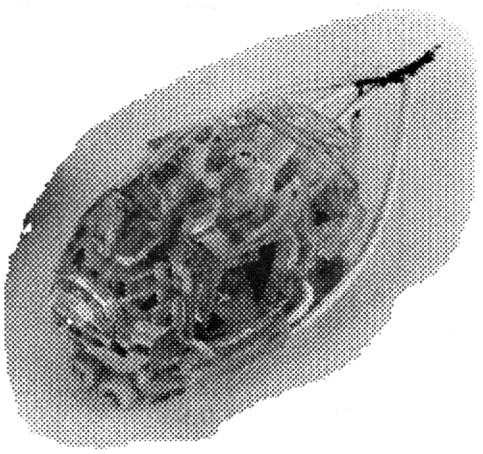


Fig. 5.14. Four frequencies Zeeman laser sensor

The active medium consists of the mixture He^3, Ne^{20}, Ne^{22}. The isotopes of Ne^{20} and Ne^{22} are in ratio about 1:1 and they are used for the extension of the gain line and increasing the area of the existence of the four-frequencies mode. For the separation of the orthogonally polarized beams in the output signal the depolarization of the output mirror is decreased up to minimum by reduction of the incidence angle up to 15°.

The optical path control system stabilizes the length of resonator so that the biasing frequencies to be equal. The design of four-frequencies sensor is showed at the Fig. 5.14.

The main advantages of four-frequencies sensors in compare with two-frequencies ones are the low sensitivity to the external magnetic fields and the receiving the additional information while four signals are processed. The usage of this information for the forming of the correction signals permits to reduce the error due to magnetic sensitivity up to 0.03 degree/h·Oe, to stabilize the resonator perimeter up to $(10^{-3} - 10^{-4})\lambda$, and apply the correction of the active medium degradation during the operation and storage time.

The Zeeman laser sensors have relatevely high sensitivity to external magnetic field. This disadvantage is particularly compensated by the possibility of the compensation of the drifts which are connected with the active medium parameters and losses instability. This is because of the dependence between the magnetic sensitivity and the active medium parameters and losses in the resonator is the same as the connection between the main components of the drift and the active medium parameters and losses.

The main parameters of the produced Zeeman laser sensors are presented at the table 5.1.

Table 5.1

	Two-frequences ZLK			Four- frequences ZLK
	ZLK-12	ZLK-16	ZLK-20	ZLK-28
Perimeter length, cm	12.8	16	20	28.5
Scale factor, arcsec/pulse	4.0	3.3	2.7	3.0
STATIONARY REGIME				
Zero drift stability, deg / h	0.1	0.05	0.05	0.03
Zero drift repeatability, deg / h	0.3	0.15	0.15	0.1
Random drift, deg / √ h	0.02	0.01	0.01	0.005
TRANSIENT REGIME				
Zero drift repeatability, deg / h	5-10	1-3	1-2	0.5
Scale factor repeatability, ppm	200	100	100	50
CONDITIONS				
Shocks, g	500	500	500	300
Vibration, g	40	35	25	5
Temperature, °C	-55 .. +60	-55 .. +60	-55 .. +60	-10 ..+ 60
Magnetic sensitivity, deg /h·Oe	2	0.2	1	0.03

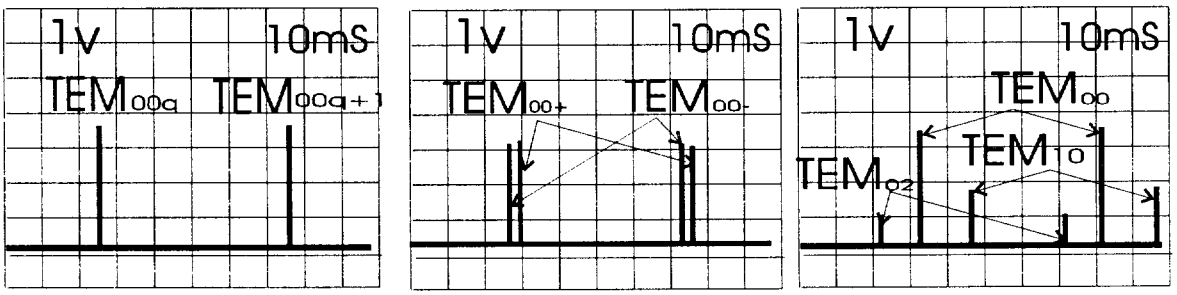
The ways of the Zeeman ring lasers improvements

As it was shown in [27, 31] the main sources of the drift of the laser gyro are increased with the increasing of the resonator losses. The improvement of the ZLGN accuracy is available firstly with help of the improvement of the mirrors quality: decreasing the losses and the depolarization of the reflected and backscattered waves. Currently developed mirrors which are produced by the method of ion-beam sputtering deposition technology permits to reduce the losses in the resonator more than 5 times and to decrease the "lock-in" zone. The scattering indicatrix graphs from the different mirrors which were plotted with help of the differential scattering measurement equipment show that backscattered waves depolarization is decreased at the improved mirrors more than on order.

The additional way of the improvement of the Zeeman ring laser accuracy is the increasing of the manufacturing and adjustment accuracy of the resonators. To decide this problem the equipment for the adjustment and test of resonator was developed. The principle of this equipment functioning is

based at the analyses of the transmission spectrum of the tested resonators [52]. The transmission pulses durations at the frequencies of the resonator own modes and the mode distances are measured by scanning the sounding beam frequency.

The typical oscillograms of the transmission spectrum of the out-of-plane resonators are presented in Fig. 5.15. Measured intracavity losses are determined by the formula $\sigma=2\pi\Delta t/T$, where Δt is the transmission pulse duration, T - time interval between transmission pulses.



a) b) c)

Fig. 5.15. Typical oscillogramms of transmission spectrums of out-of-plane resonators: a) - two neighbor axial modes with RCP; b) - modes with RCP and LCP; c) - circular polarized modes with different lateral indexes.

After measurement Δt and T the total intracavity losses are calculated, and they are checked during assembly and technological operationof the resonators.

The resonators are adjusted on the minimum of the diffraction losses of the high order modes, and it helps to improve accuracy of sensor manufacturing.

• 5.7. ZEEMAN LASER GYRO BEHAVIOR UNDER SEVERE ENVIRONMENT

Operation of ZLGN under high magnetic fields

As it was mentioned before, ZLGN has enough high sensitivity to external magnetic fields because of magnetooptical type of the biasing. This problem is common for any laser gyros with the biasing on magnetooptical effects (including Faraday and Kerr effects) [41].

The character of the magnetic sensitivity of ZLGN can be understood from Fig. 5.16. Here an out-of-plane optical path configuration 1-2-3-4 is shown. Also there are the projections of the magnetic sensitivity vectors of the resonator legs on three orthogonal planes α, β, γ.

It is easily seen that on the plane α the sum vector of the magnetic sensitivity $S_{\alpha\Sigma}==S_{\alpha 12}+S_{\alpha 14}+S_{\alpha 34}+S_{\alpha 32}=0$. On the planes β and γ the sum vectors of the magnetic sensitivities are equal correspondingly $S_{\beta\Sigma}=S_{\beta 32}+S_{\beta 12}+S_{\beta 34}+S_{\beta 14}\neq 0$ and $S_{\gamma\Sigma}=S_{\gamma 34}+S_{\gamma 32}+S_{\gamma 14}+S_{\gamma 12}\neq 0$. Therefore $S_{\Sigma}=S_{\beta\Sigma}+S_{\gamma\Sigma}$ is not equal to zero and directs perpendecular to the plane α.

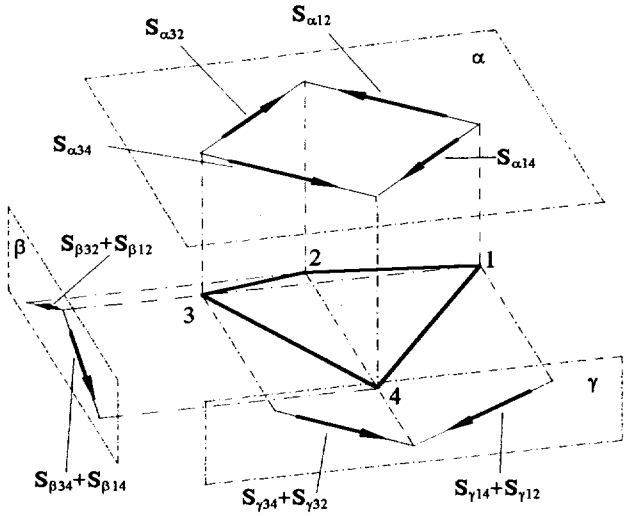


Fig 5.16. Projections of out-of-plane optical path 1-2-3-4 in resonator of ZLGN on the orthogonal planes α, β, γ.

It can be noticed that the magnetic sensitivity vector of the ZLGN coincides in direction with the ring laser rate sensitivity axis.

Therefore it is necessary to protect such laser gyros in case of operation under the strong or changing in time magnetic fields, which can influence on ZLGN accuracy.

Two concepts of the protection were investigated in R&DI"Polyus": by the active methods of the magnetic field compensation (on magnetic sensor data or inherent laser gyro data) and by the passive methods (with magnetic shielding) [42].

For currently produced ZLGN the magnetic shields are used. R&DI"Polyus" designers were first to suggest and realize in 1984 a new technical concept of 3-axis laser gyro unit which provides simultaneously the effective protection from the external magnetic fields and the reduction of weight and size. Before that a powerful (three-layers) own shield was used for each ring laser of 3-axis unit. But this way led to a very heavy 3-axis laser gyro unit not to meet to the state-of-the-art requirements for the weight and size. In accordance to a new technical concept a thin (single-layer) shield is used for each ring laser and a high protective (two-layers) common shield is used for the compact 3-axis unit, integrating three ring lasers and electronics.

This concept of integrated design permitted to reduce significantly weight and size (in 2.5 and 4 times correspondingly) of 3-axis unit. First 3-axis ZLGN based on this concept was model ZLK-20-1 with weight 5.5 kg and volume 3.6 dm^3.

As for other countries Honeywell in USA also began to apply from 1993 integrated type of 3-axis unit design in their advanced model GG-1320DGL [43].

At the next steps of development of ZLK family the shields design was improved and they were made from the μ-material with high relative magnetic permeability μ=50 000-70 000. Now

ZLGN have magnetic sensitivity 0.1-0.2 deg/h·Oe for two frequencies modifications and 0.03 deg/h ·Oe for four frequency modifications (with autocompensation of magnetic component of error).

For further reduction of the ZLGNs magnetic sensitivity not only a shield design will be optimized, but also the active methods of the magnetic field compensation are planned to be applied.

Operation of ZLGN under shocks and vibrations
Owing to absence of dither and any intracavity optical components ZLGN are inherently very stable to vibrations and shocks. Also, the resonator with out-of-plane optical path is about 20 times more stable to deformations than the resonator with plane optical path.

A single element of design which can sense g is piezomirror driving the optical path. That is why R&DI "Polyus" designers put a special attention on this element to prevent the errors due to optical path changing.

It was found [44] the design and width of piezomirror to make the resonance frequency much higher than the typical vibration range. During detail experiments ZLGN were proved to have very low sensitivity to vibrations. Only additional random noise in the output signal appeared at the relatively big amplitudes of random vibrations as it can be seen from Fig. 5.17.

In case of very big shocks (more 100 g) a special signal is formed at the output of ZLGN, if the path changing is more >0,2λ, then INS computer filters the output laser gyro information considered wrong and exchanges it with previous correct data. Duration of shocks is very short (up to 10 ms) and even though optical path changed no error appears owing to filtering.

ZLGN were proved to be stable to vibrations up to 35g in 50 - 2 000 Hz range and shocks up to 500 g [45 - 47].

Operation of ZLGN in wide temperature range
When temperature is changed the zero drift of laser gyro is changed also and it will cause the error. So usually the temperature of laser gyro body is measured by the special temperature sensors and then the temperature error is compensated in accordance with the algorithm of temperature correction.

For ZLGN it was developed the original method of temperature errors compensation. The value of Zeeman biasing is connected with the internal temperature of gas plasma, and this value can be easily measured at the output of ZLGN.

Connection of the Zeeman biasing with the internal temperature is caused by physical processes in Zeeman ring lasers. As it was showed above, the counter-propagating waves splitting (biasing) ν_B is described by the expression (5.5).
Doppler line width is determined by the formula [16]:

$$\Delta \nu_D = 2\,\nu_o \sqrt{\frac{2kT_K \cdot \ln 2}{Mc^2}}\,, \qquad (5.16)$$

where k - Boltzmann constant; M - atom mass; T_K - active medium temperature in Kelvin.

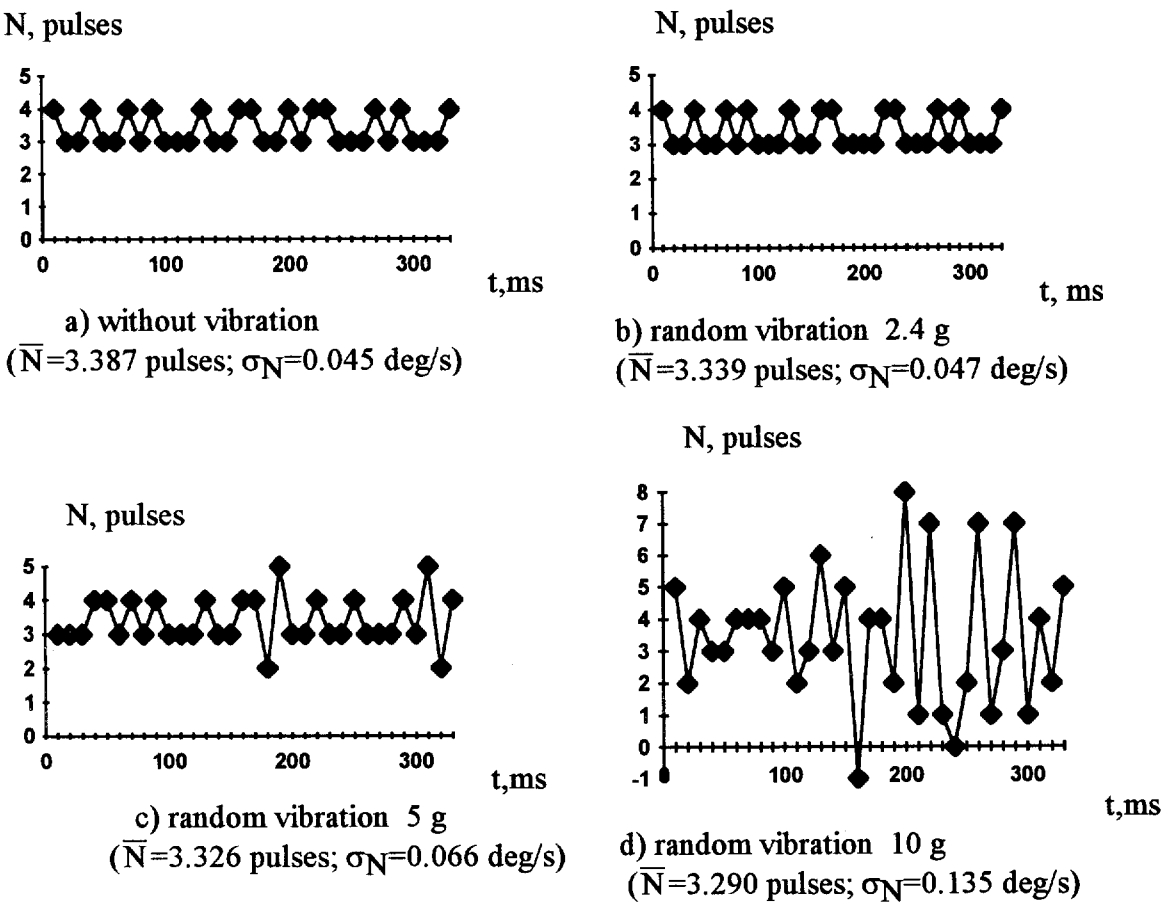


Fig. 5.17. Output signal of ZLGN under random vibration in the range 20 - 2000 Hz (2.4 g; 5 g; 10 g). Time of operation 20 s; cycle time 10 ms; 1 pulse corresponds 3.3" (Mean value $\overline{N}$ changes from (a) to (d) due to zero temperature drift; rms σ_N changes due to rise of vibration).

Substitution of formula (5.16) into formula (5.5) gives the following relations for the ν_B (while gain and losses are equal):

$$\nu_B \approx 1.3 \cdot 10^6 H \cdot \frac{\Delta\nu_c \sqrt{Mc^2}}{\nu_o \sqrt{2kT_K \cdot \ln 2}} = 1.3 \cdot 10^6 H \cdot \frac{\Delta\nu_c \sqrt{Mc^2}}{\nu_o \sqrt{2k \cdot \ln 2}} \cdot \frac{1}{\sqrt{T_K}}. \qquad (5.17)$$

Thus, according to formula (5.17), if the strength of the magnetic field is constant the biasing value is inversely propotional to the root of the gas plasma temperature $\sqrt{T_K}$ (see Fig. 5.18). The range of the operational temperature of the laser gyro corresponds to the practically linear part of the dependence $\Delta\nu_B = f(\sqrt{T_K})$. Therefore for the real applications the biasing value can be considered to depend linearly on the internal temperature in the operational temperature range.

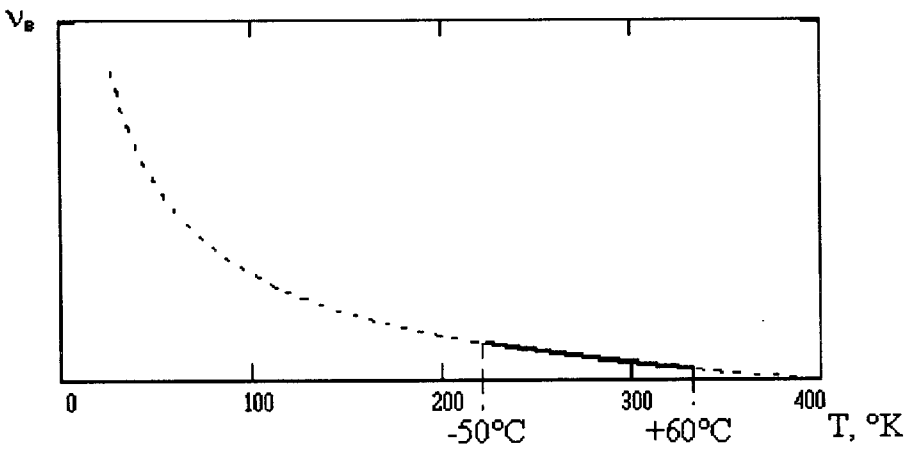


Fig. 5.18. Theoretical dependance ν_B versus absolute temperature.

It can be also shown by the transformation of the formula (5.17). Let's designate all constants for our consideration in formula (5.17) as ν_{BC}

$$\nu_{BC} = 1.3 \cdot 10^6 H \cdot \frac{\Delta \nu_c \sqrt{Mc^2}}{\nu_o \sqrt{2k \cdot \ln 2}} . \tag{5.18}$$

After turning to Celsius degrees and simplifying the formula (5.17) we will get:

$$\nu_B \approx \nu_{BO}\left(1 - \frac{T_C}{546}\right), \tag{5.19}$$

where ν_{BO}- biasing value while temperature T_C=0°C, $\nu_{BO}=\nu_{BC}/\sqrt{273}$.

Experimental dependences of the biasing value versus temperature are presented in Fig. 5.19. As it is seen from Fig. 5.19 the biasing value is connected with temperature by practically linear dependence in operational temperature range.

In ZLGN the dependences of the zero drift versus biasing value are determined experimentally in the whole temperature range for the each ring laser during manufacturing. Then while ZLGN operation the biasing value is measured at the output of laser gyro and with application of previously determined dependences the temperature changing of zero drift is compensated.

In compare with the usual compensation method with placing of the temperature sensor at the resonator body the described method "correction on biasing" is more effective because the biasing value is directly connected with the internal temperature of gas plasma and better correlated with the zero drift.

Thus ZLGN has the simple mechanism of the connection between internal temperature and output signals which provides the continuous and reliable temperature measurement. This feature of ZLGN is used for the algorithmic correction of zero drift temperature changing.

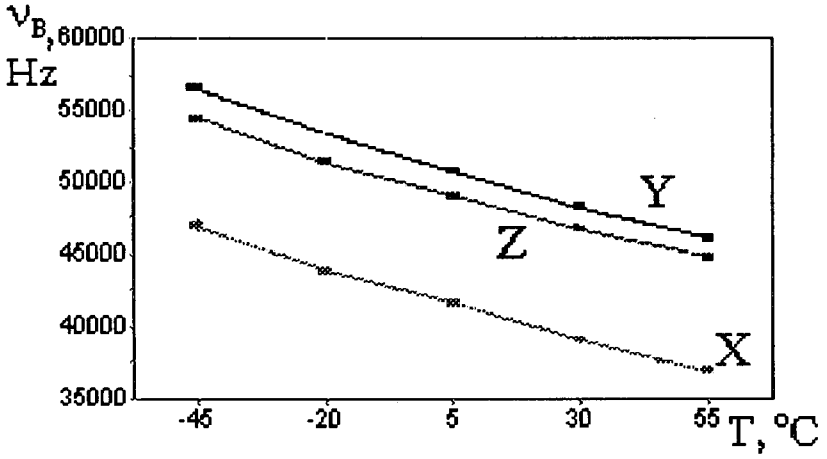


Fig. 5.19. Experimental dependences of the biasing value ν_B versus temperature

Remind that formula (5.19) was derived for the case of equality of gain and losses. Really the ring laser works with gain higher losses, and in this meaning the formula (5.19) is approximate.

Let's consider formula (5.19) for the common case when gain is higher losses. After transformation of expression from [48] to convenient for our consideration form and turning to our designations we get the formula:

$$\nu_B \approx \nu_{BO}\left(1 - \frac{T_C}{546}\right)\cdot\left(C_G G + C_L \sigma\right), \qquad (5.20)$$

where G - gain; σ - losses coefficient; C_G, C_L - constants for our consideration.
Thus, in this more common formula the dependence $\Delta\nu_B$ versus T is the same as in formula (5.19).

Physical meaning of the term in brackets ($C_G G + C_L \sigma$) is the expression of the dependence of the biasing value versus gain and losses coefficient. For example, if parameters of the active medium are constant, but the losses are increasing due to mirrors degradation, the biasing value will increase. If mirrors parameters are stable, but the gain decreases due to pollutions of the gas, the biasing value will decline.

The mirrors degradation and gas composition changing are very slow processes therefore the biasing changing due to these processes are easily distinguished from the fast and more significant changes connected with temperature variation during self-warming or under the external temperature influences.

That is why the biasing dependence on temperature is reasonable to use for algorithmic correction of zero drift changes with temperature during the operational time. Analyses of the slow changing of the biasing value will show the dynamic of the physical processes, first of all, changing of gain and losses during storage.

It is recommended to measure not only the biasing value, but also output optical power and discharge voltage. It gives the unique possibility to recognize the reason of the parameters changing and to distinguish the degradation of mirrors from degradation of the active medium.

Also a feedback available with control of the optical power to keep gain to losses ratio constant during long time, and it will increase the accuracy of the temperature "correction on biasing". Further the algorithmic "correction on biasing" is considered.

Algorithmic correction

Typical zero drift changing in time after turn-on is presented in Fig. 5.20. For describing of this process the following combining function was proved to be the best (further everywhere temperature in Celsius degrees) [46]:

$$\Omega(t, \nu_B) = \Omega_0(\nu_B) + \Omega_1(\nu_B)\, e^{-t/\tau_1} + \Omega_2(\nu_B)\, e^{-t/\tau_2}, \qquad (5.21)$$

where $\Omega_0(\nu)$ - constant component of the zero shift, depending on temperature; $\Omega_1(\nu)$ and $\Omega_2(\nu)$ - coefficients at exponents; τ_1 - time constant about 1 hour, corresponding the resonator self-warming time; τ_2 - time constant about 1 min, determining the time of stabilization of convection flows in active medium and the time of stabilization of temperuture gradients along resonator body.

Investgations showed the coefficients to be well described by polynomials of the second order:

$$\begin{aligned}
\Omega_0(\nu_B) &= A_1 - A_2\nu_B + A_3\nu_B^2;\\
\Omega_1(\nu_B) &= A_1 - B_2\nu_B + B_3\nu_B^2;\\
\Omega_2(\nu_B) &= C_1 - C_2\nu_B + C_3\nu_B^2;\\
\tau_1(\nu_B) &= D_1 - D_2\nu_B;\\
\tau_2(\nu_B) &= E_1 - E_2\nu_B.
\end{aligned} \qquad (5.22)$$

The coefficients $A_{1,2,3}$, $B_{1,2,3}$, $C_{1,2,3}$, $D_{1,2}$, $E_{1,2}$ are determined during manufacturing and testing of ZLGN, then are memorized in ROM of every device.

While application the value of biasing is measured at the ouput of ZLGN and with considering the ROM coefficient the algorithmic correction is provided in INS computer according to formula (5.21).

Summerising the rezults of the numerous tests we can prove that accuracy was increased in 10-20 times after application of the algorithmic correction on biasing value (see Fig. 5.20) . ZLGN can keep accuracy in wide range of temperature from -55 to 60°C.

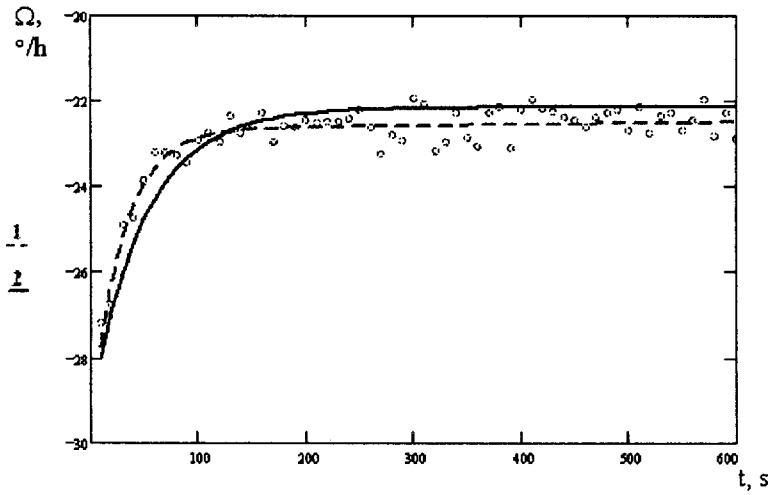


Fig. 5.20. Zeeman laser gyro typical dependence of zero drift Ω vs time after turn-on (line 1 - average line of output data; line 2 - line of algorithmic correction).

• 5.8. APPLICATION OF ZEEMAN LASER GYRO

From 1980 most of the new inertial systems are being developed on laser gyros. ZLGN is one of the best candidate for the modern inertial systems due to following reasons [53]:

1) being completely optical device without any moving parts (no dither mechanism) ZLGN in compare with dithered laser gyros doesn't generate any acoustic noises and excludes mechanical cross-coupling between gyros which are connected with the manner of mounting of dithered laser gyros and cause their coning motions appearing as a navigation error;

2) applying of magnetooptical (Zeeman) biasing ZLGN has wide possibilities to control easy the amplitude of biasing and even makes available to change this amplitude during operation inside the system, that permits to optimize the inertial system alignment procedure thereby getting higher accuracy and shorter ready time;

3) having significantly better performance in complex than fiber optic gyros (lower drift and noises, higher scale factor stability, higher resistance to shocks and vibrations) ZLGN better fits for severe environmental applications with wide dynamic range of angular rates, large temperature range, hard mechanical influences;

4) reaching the mass-production stage in contrast with other new gyros (fused quartz hemispherical resonator gyros, micromechanical gyros etc.) ZLGN can be adopted and supplied to many currently developed modern inertial systems.

Results of tests in conditions of real exploitation (from -55 to +60 °C, up to 500 g shocks and 35 g, 50 - 2,000 Hz vibrations) proved the capability of ZLGN [8, 45-47].

One of the concrete applications of ZLGN is system BINSU (Strapdown Inertial Navigation and Control System). BINSU is inexpensive system combines navigation and control functions owing to

ZLGN usage. BINSU provides autonomous flight and pointing for maneuverable vehicles, has mean values of mass and size. Another application is maneuverable vehicles guidance into precalculated position according to information of the inertial navigation system, control and stabilization on every stage of trajectory, pointing at the final stage. Main specifications of BINSU are as follows: accuracy of determing of navigation parameters (σ) regardless error of initial alignment for 120 s of movement along axes is no more 200 m, along angles is no more 0.5 deg; continuous operation time 1,200 s, time of readiness 3 s, measured angular range ±600 deg/s.

Integration with GPS permits to apply ZLGN also in precise INS [49].

In summery the described and other available areas of applications of ZLGN are followings:
- low-cost middle-accuracy inertial or GPS-aided precise systems for aircraft, helicopters, spacecrafts, ships, land maneuverable vehicles;
- inertial survey and geodetic systems;
- railroad checking;
- well bores control.

REFERENCE

1. Khromykh A.M., JETF, 1966, V.50, p.1.
2. Macek W.M., Davis D.T., Appl. Phys. Lett, 1963, V.2, N67.
3. Melnikov A.V., Prosvetov V.K., Ribakov B.V., Skulachenko S.S., Khromykh A.M., Yudin I.I., Patent USSR N745242, 16.05.1967.
4. Sanders V.E., Anderson D.Z., US Patent 4.247.832, 1981.
5. Dorschner T.A., Smith I.W., J.Opt. Soc. Amer., 1978, V.68, p.1381.
6. Smith I.W., Dorschner T.A., US Patent 4.110.045, Aug 1978.
7. Podgorskiy T.J., Aronowitz F., IEEE J.Quantum Electron., 1923, QE,1, N1, p. 11.
8. Golyaev Yu.D., Dmitriev V.G., Kolbas Y.Y., Livintsev A.L., Melnikov A.V., Mescheryakov B.T., Soloviova T.I., Telegin G.I., Tikhmenev N.V., Proceedings of 2nd Soviet – Chinese Simposium on Inertial Technicks, Leningrad: CNII "Electropribor", 1991, p. 143 - 145.
9. Golyaev Yu.D., Livintsev A.L., Soloviova T.I., Tikhmenev N.V., Proceedings of Conference "Development and application of perspective devices for measurement of rate and acceleration, Kovrov, NTC "Informtechknika", 1991", p. 12 - 14.
10. Bogatov A.G., Golyaev Yu.D., Dedish V.V., Dmitriev V.G., Kolodnyi G.Ya, Kravtzov N.V., Livintsev A.L., Nanii O.E., Nadtocheev V.E., Soloviova T.I., Firson V.V., Izvestiya Akademii Nauk, ser. Fizicheskaya, 1992, V.56, N9, p. 170 - 174.
11. Golyaev Yu.D., Dedish V.V., Dmitriev V.G., Kravtzov N.V., Lariontzev E.G., Livintsev A.L., Nadtocheev V.E., Nanii O.E., Soloviova T.I., Firson V.V. Veselovskaya T.V., Izvestiya Akademii Nauk, ser. Fizicheskaya, 1992, V.56, N9, p. 163 - 169.
12. Golyaev Yu.D., Dedish V.V., Dmitriev V.G., Kravtzov N.V., Livintsev A.L., Nadtocheev V.E., Nanii O.E., ., Soloviova T.I., Firson V.V., Lazernaya Technika i optoelectronika, 1993, N1,2(68,69), p.51 - 54.
13. Boyko D.L., Goliaev Y.D., Dmitriev V.G., Kravtzov N.V., Lazernye novosty, 1997, V.2, p. 11.
14. Born M., Atomic physics, 1965.
15. Chow W.W. et al, Reviews of Modern Physics, 1985, V.57, N1.
16. Bennet V., Gas masers, Mir, 1964.
17. Beterov I.M., Matushin Y.A., Chebotaev V.P., Optika i spectroscopiya, V.28, N2, p. 357.

18. Ribakov B.V., Skulachenko S.S., Khromykh A.M.,Yudin I.I., JETF, 1973, V.64, p. 1146.
19. Khromykh A.M., Yakushev A.I., Kvantovaya electronika, 1977, V.4, N1, p. 27.
20. Saveliev I.I., Khromykh A.M., Yakushev A.I., Kvantovaya electronika, 1979, V.6, N6, p. 1155.
21. Saveliev I.I., Timonin P.V., Yakushev A.I., Kvantovaya electronika, 1979, V.6, N7, p. 1549.
22. Saveliev I.I., Kvantovaya electronika, 1979, V.6, N3, p. 632.
23. Mogylnaya T.Y., Saveliev I.I., Timonin P.V., Yakushev A.I., Kvantovaya electronika, 1983, V.10, N10, p.2032.
24. Nazarenko M.M., Saveliev I.I., Skulachenko S.S., Khromykh A.M., Yudin I.I., Kvantovaya electronika, 1977, V.4, N8, p.1738.
25. Saveliev I.I., Khromykh A.M., Kvantovaya electronika, 1976, V.3, N7, p.1517.
26. Bershtain I.L., DAN USSR, 1965, V.163, p.60.
27. Klimontovich Y.L., Kuriatov V.N., Landa P.S., JETF 1966, V.51, p. 3.
28. Kuriatov V.N., Landa P.S., Larionov E.G., Radiophisika, 1968, V.11, p. 1839.
29. Ribakov B.V., Demidenkov Y.V., Skrotskiy S.G., JETF, 1969, V.57, S.4(10), p. 1184.
30. Azarova V., Golyaev Yu., Dmitriev V., Melnikov A., Nazarenko M., Svirin V., Tikhmenev N., Proceeding of 52nd Annual Meeting Navigational Technology for the 3rd Millenium, Cambridge, Massuchusetts, 1996, p.697 - 703.
31. Aronowitz F., The Laser Gyro in Laser Application, New York, Academic Press, 1971, p. 133.
32. Aronowitz F., Collins R., J Appl.Phys.Lett., 1966, V.9 p.55.
33. Klimontovich Y.L., "Nauka", 1974.
34. Bambini A., Stenholm S., Optics Communication, 1984, V.9, N4, p. 269.
35. Khromykh A.M., Electronnaya tehknika, 1990, V.11, S.2(54), p.44.
36. Khromykh A.M., Lasernaya technika i optoelectronika, 1990, V.11, S.3(55), p.49.
37. Golyaev Yu., Tikhmenev N., Yaremenko S., Lasernaya tehknika i optoelectronika, 1991, V.11, S.2(58), p. 59.
38. Naida O.N., Rudenko V.V., Lasernaya tehknika i optoelectronika, 1991, V.11, S.1(57), p. 75.
39. Langmuir I., Franklin Inst., 1923, V.196, p. 751.
40. Golyaev Yu.D., Dmitriev V.G., Kazakov A.A., Melnikov A.V., Soloviova T.I., Son J.Y., Proc. SPIE, 1993, V.2778, Part 11, p. 1084 - 1085.
41. Saveliev A.M., Soloviova T.I., Zarubezhnaya radioelectronika, 1981, N8, p. 77 - 92.
42. Astakhov K.V., Golyaev Yu.D., Drozdov M.S., Dmitriev V.G., Kolbas Y.Y., Soloviova T.I., Titov A.N., Chemiris A.I., Proceedings of 3nd St. Peterburg International Conference on Integrated Navigation Systems, St.Petersburg: CNII "Electropribor", 1996, p. 151 - 157.
43. Mullarkey J.E., Proceedings of 2nd St.Petersburg Internationaal Conference on Gyroscopic Technology and Navigation, St.Petersburg: CNNI "Electropribor", 1995, p. 20 - 27.
44. Azarova V.V., Golyaev Yu.D., Dmitriev V.G., Melnikov A.V.,.Zapotilko N.R. Nazarenko M.M., Svirin V.N., Tikhmenev N.V., Proceedings of 2nd St.Petersburg International Conference on Gyroscopic Technology and Navigation, St.Petersburg: CNII "Electropribor", 1995, p. 49 - 50.
45. Golyaev Yu.D., Dmitriev V.G., Melnikov A.V.,. Soloviova T.I., Proceedings of International Conference "Imitation of flight'92", 1992, Moscow, ZAGI, p. 131 - 132.
46. Golyaev Yu.D., Dmitriev V.G., Melnikov A.V.,. Soloviova T.I., Proceedings of 1st St.Petersburg International Conference on Gyroscopic Technique, St.Petersburg, CNII "Electropribor", 1994, p. 178 - 190.
47. Golyaev Yu.D., Soloviova T.I., Scientific-Technical Proceedings of "Eurika'95" Brussel International Exibition, Moscow, NTC "Informtechnika", 1996, N5, p. 10 - 26.
48. Khromykh A.M., Yakushev A.I., Kvantovaya electronika, 1977, V.4, N1, p. 27 - 34.
49. Fomichev A.A., Dmitriev V.G., Kolchev A.B., Skrozkiy S.G., Chemiris A.I., Opticheskaya technika, 1995, N1(5), p. 28 - 31.

50. Azarova V., Golyaev Yu., Dmitriev V., Melnikov A., Nazarenko M., Svirin V., Tikhmenev N., Proceedings of 1st St.Petersburg International Conference on Gyroscopic Technique, St.Peterburg, CNII "Electropribor", 1994, p.141 - 152.
51. Azarova V., Golyaev Yu., Dmitriev V., Melnikov A., Nazarenko M., Svirin V., Tikhmenev N., Proceedings of 2nd St.Petersburg International Conference on Gyroscopic Technology and Navigation, St.Petersburg, CNII "Electropribor", 1995, p.49 - 50.
52. Azarova V.,.Giruts E., Kopilov S., Nazarenko M., Izvestiya Academii Nauk, Seriya physicheskaya, 1994, V.58, N2, p.4 - 10.
53. Azarova V., Golyaev Yu., Dmitriev V., Drozdov M., Kazakov A., Melnikov A., Soloviova T., Tikhmenev N., Titov A., Gavrilin B., Slaschinin M., Tikhmenev V., Selivanova L., Proceedings of Symposium Gyro Technology 1998, Stuttgart (Germany).

6. LASER GYROS WITH TOTAL REFLECTION PRISMS.

Y.V.Bakin, G.N.Ziouzev, M.B.Lioudomirski

Bauman Moscow State Technical University
No.5 2nd Baumanskaja St.
Moscow, 107005 RUSSIA

• 6.1. HISTORY

From the first steps of laser gyroscopy development it was obvious, that a ring cavity with high Q-factor and low back-scattering level is a basis of precise laser gyro (LG).

To provide minimal losses inside the cavity reflective elements with maximal reflectivity are to be used. Essentially greatest reflectivity (equal to unit) can be provided if the total internal reflection effect is used. Because of this matter the ring cavities techniques development initially went in two directions - investigations of mirror cavities and cavities with total reflection prisms (TRP).

The first publications on TRP ring cavities concern to the beginning of 1970th [6.1, 6.2].

Various opinions of LG experts about expediency of TRP cavities application are known. For example, F. Aronowitz [6.3] referred them to as unpromising because of their temperature sensitivity and magnetic sensitivity.

Experience of laser gyros with prism cavities development in Russia demonstrates the ability to achieve high accuracy specifications.

On the Fig. 6.1 there are shown the main types of total reflection prisms for application in the TRP cavities.

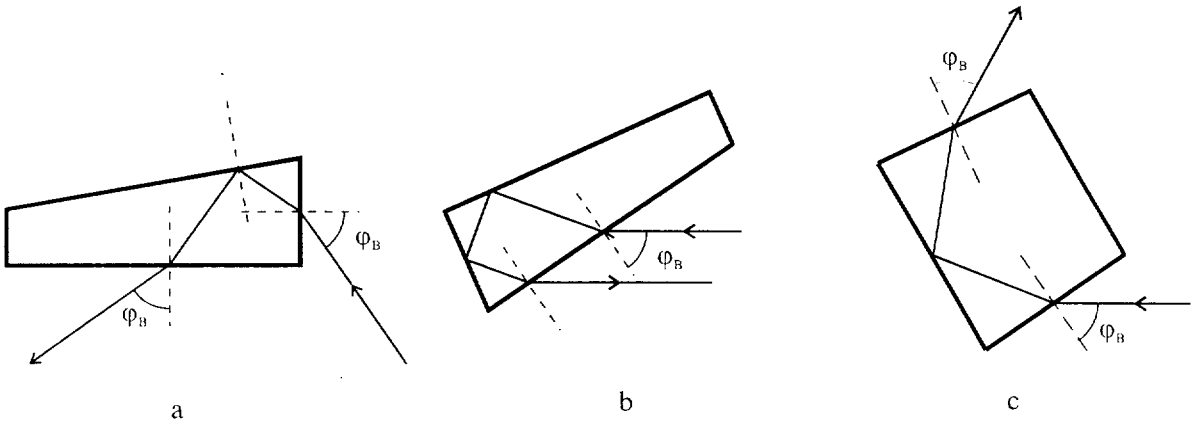


Fig. 6.1.

On the Fig.6.1a 90° beam rotating TRP for rectangular four-prism resonator is shown. The 180° beam rotating TRP for two-prism rectangular cavity is shown on the Fig.6.1b. TRP for the triangular cavity that provides 60° beam rotation is shown on the Fig.6.1c.

Authors have the experience in application of TRP ring cavities of below shown (Fig.6.2a...6.2e) configurations.

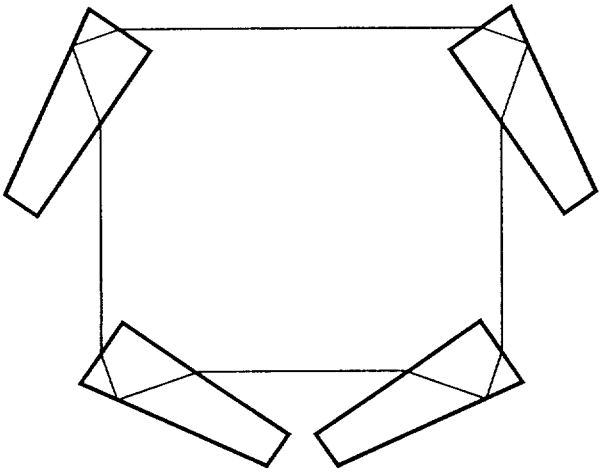


Fig. 6.2a.

Four-prism Cavity with Reversed Second Prisms.

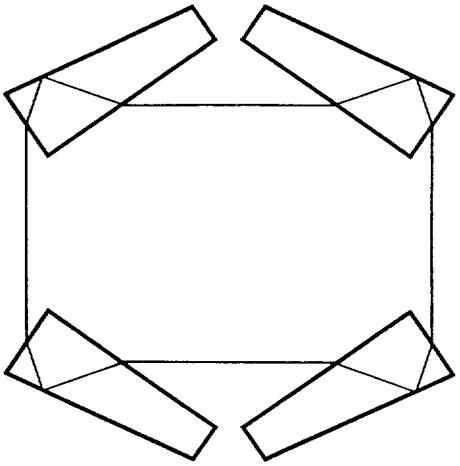


Fig. 6.2b

Four-prism Cavity

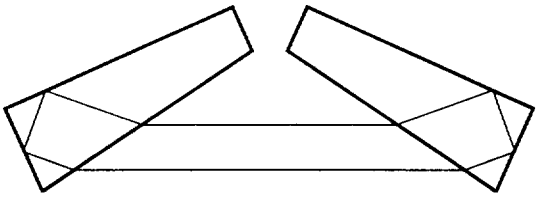


Fig. 6.2c

Two-prism Cavity No.1

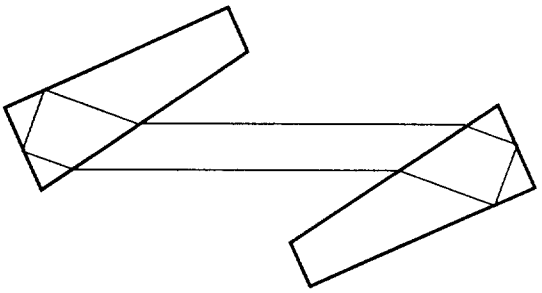


Fig. 6.2d

Two-prism Cavity No.2

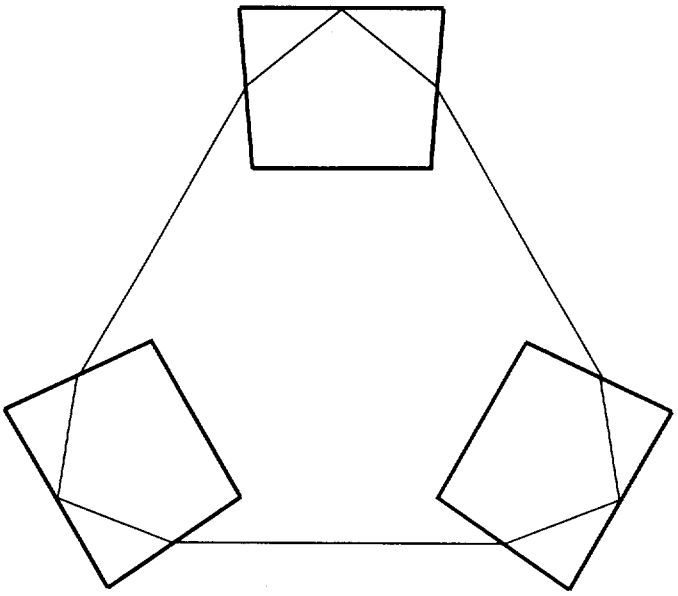


Fig. 6.2e

Three-prism Cavity

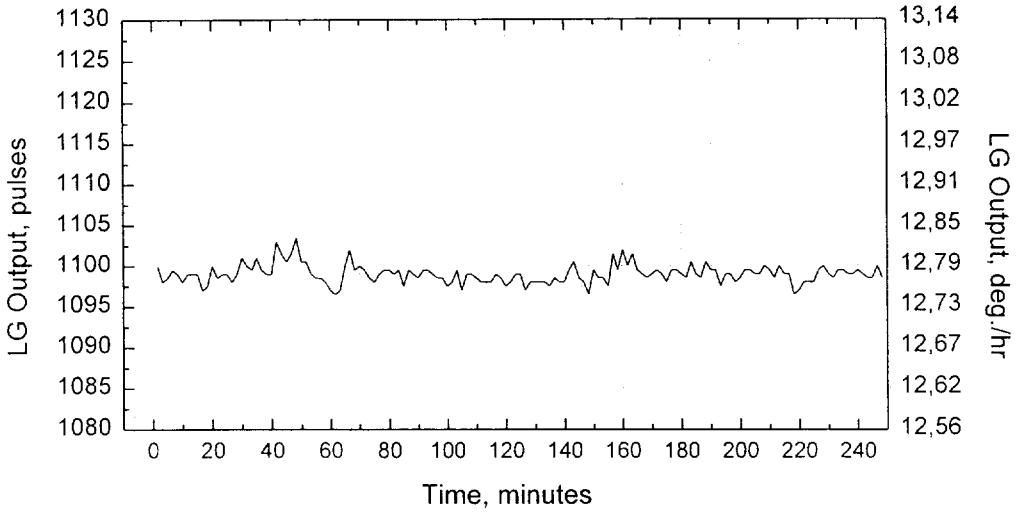
On the base of four-prism cavity with reversed second prisms (Fig.6.2a) that has total losses less than 50 ppm in 1990...1995 the mechanically dithered laser gyros with cavity path-length values of 0.45m and 0.28m were developed. Their technical specifications are given below .

Table 6.1. **Specifications of the TRP Laser Gyros.**

Cavity Path-Length	0,45 m	0,28 m
Scale factor	1,31"/pulse	2,30"/imp
Scale factor stability & linearity	10 ppm	30 ppm
Random walk	0.001 °/hr$^{1/2}$	0.003 °/hr$^{1/2}$
Day-to-Day drift (1σ)	0,003 deg/hr	0,01 deg/hr
Operating temperature span	-40...+60 °C	-40...+70 °C

To illustrate above specifications results of static tests under exposure of different temperatures and external magnetic field are given below.

On the Fig. 6.3 results of the TRP Laser Gyro static test are given. LG Cavity Path-Length is 0.28 m. In the Table the statistic data of the run are given - the mean value and the Standard Deviations for different average intervals (100 s , 10 min., 20 min and 1 hour).



Duration of the run, minutes	Mean of the run		σ 100 SEC.	σ 10 MIN.	σ 20 MIN.	σ 1 HOUR
	pulses	°/hr	°/hr	°/hr	°/hr	°/hr
248	1089,99	12,675	0,013	0,009	0,006	0,003

Fig. 6.3. The TRP Laser Gyro with the Cavity Path-Length L=0.28 m.
Static Test under t = +25°C.

Results of the Run-to-Run & Day-to-Day repeatability tests are represented in the following Table. All the runs were made as «cold», one run a day. During the test LG was fixed on the special frame and kept unmovable between the runs.

	Duration of the run, minutes	Mean of the run		$\sigma_{100\ SEC.}$		$\sigma_{10\ MIN.}$		$\sigma_{20\ MIN.}$	
		pulses	°/hr	pulses	°/hr	pulses	°/hr	pulses	°/hr
Run 1	62	402,05	4,675	1,18	0,013	0,70	0,008	0,53	0,006
Run 2	63	402,05	4,675	0,89	0,010	0,57	0,006	0,57	0,006
Run 3	62	402,22	4,677	1,18	0,014	0,65	0,008	0,19	0,002
Run 4	68	401,80	4,672	1,31	0,015	1,12	0,013	0,90	0,010
		Mean of the Test		**Run-to-Run & Day-to-Day Drift**					
		pulses	°/hr	pulses	°/hr				
		402,33	**4,675**	**0,17**	**0,002**				

Fig. 6.4 illustrates the temperature drift of the TRP LG with L =0.45 m. It is obvious, that random noise of the LG does not depend on the environment temperature. For this particular LG the temperature drift is less than 0.01°/hr : Ω^{-40}_{MEAN}=12,222°/hr, Ω^{+25}_{MEAN}=12,229°/hr, Ω^{+60}_{MEAN}=12,219°/hr. For other prototypes with higher temperature drift the temperature compensation may be used. The temperature compensation algorithm and hardware will be further discussed.

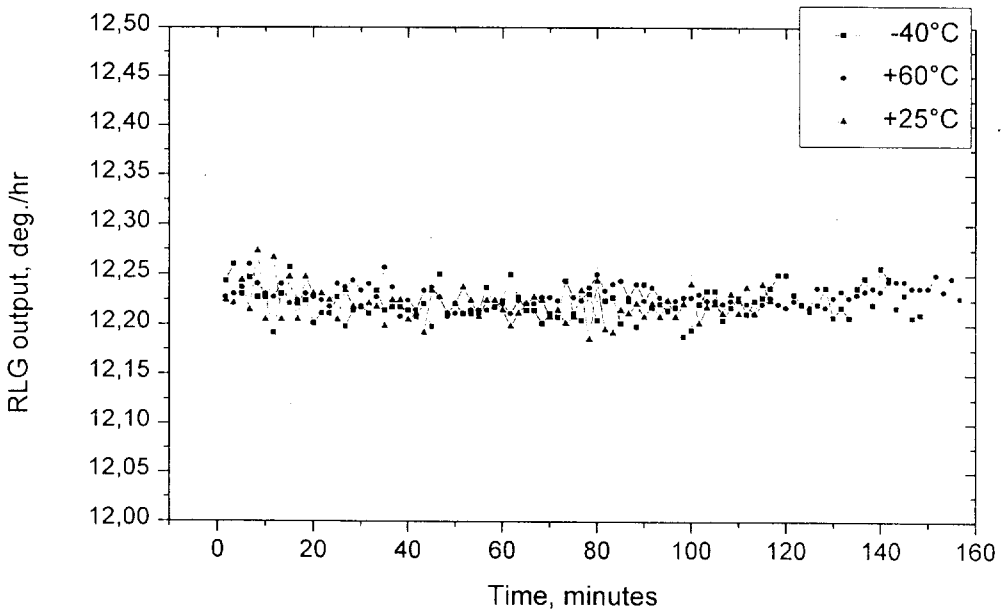
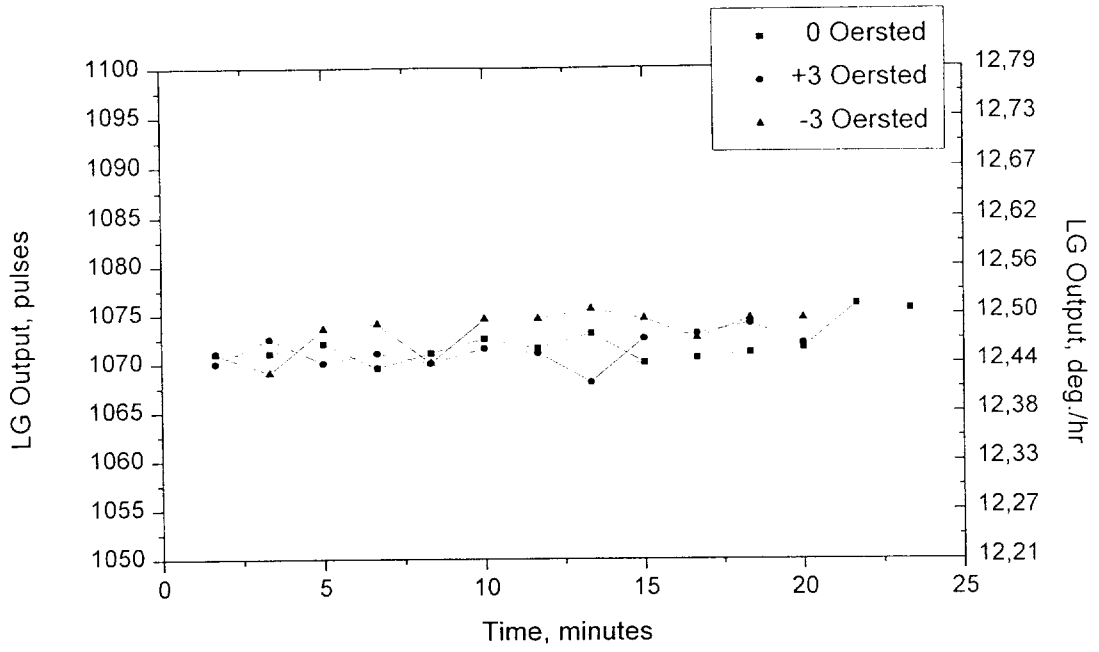
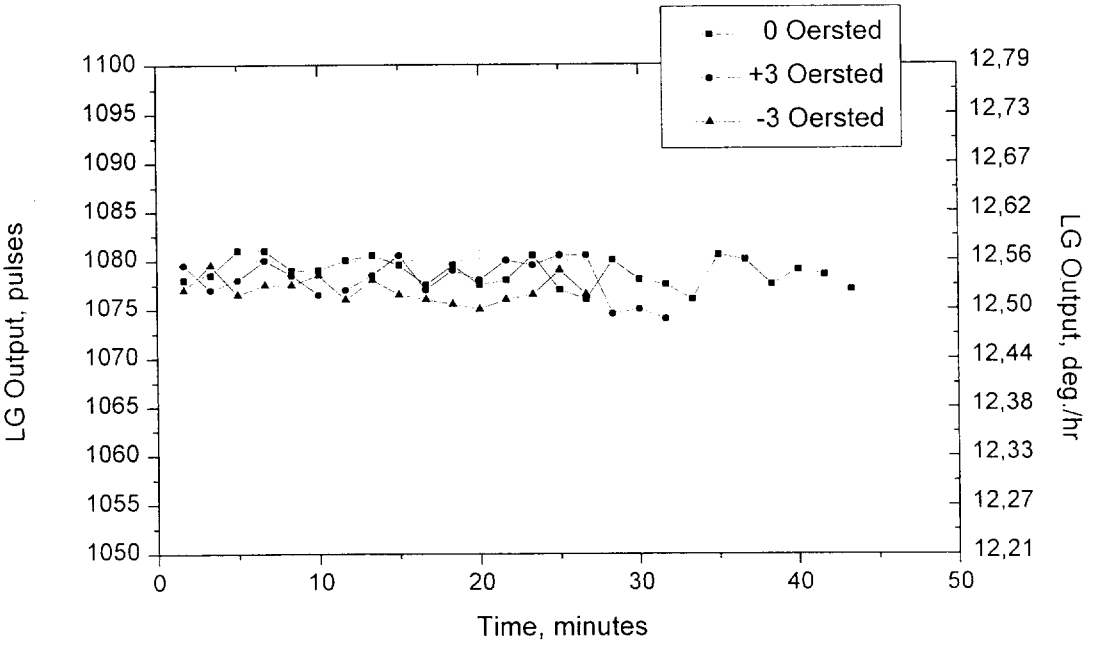


Fig. 6.4. The TRP Laser Gyro with the Cavity Path-Length L=0.45 m.
Static Test under t = -40°C, +25°C, +60°C.

Data plots on Fig. 6.5a, 6.5b are give to estimate magnetic sensitivity of the TRP LG.



a.



b.

Fig. 6.5. The TRP LG Runs under External Magnetic Field.

The gyro with L =0.28m was tested. The Helmholtz coils were used as a source of magnetic field. The LG was mounted between the coils near their axis. To have a reference level a short run without magnetic field was made, then two runs under magnetic field of +3 Oersted

and -3 Oersted were made. Then the LG was 90° rotated, and above test was repeated. As it is obvious, the LG magnetic drift is of the same order as the noise. Numerical analysis gives a value less than 0.01(°/hr)/Oersted. The same value is obtained in the LG with L =0.45m, and in quantity production it is a subject of special acceptance test.

In TRP LG based measuring units this magnetic sensitivity may be actually cancelled by additional magnetic shielding of the whole unit.

In 1995 the TRP LG with cavity path-length of 0.45 m passed the flight test. Since then it is under serial production as the basic element of INS for the civil aircrafts TU-204 and IL-96-300. Prism LG of the second type is lot job produced according to particular orders.

A lot of special design and technology solutions were applied to decrease permanent stresses inside the prisms material to overcome by this way the main problem of the prism cavities. It made the prism LG competitive with mirror LG and sometimes even more preferable because the absence of any coatings on the cavity reflective elements and inside the laser allows prism LG to keep essentially better stability of main specifications and longer life time and storage time.

• 6.2. STRUCTURE OF PRISM LG, PRINCIPLE OF OPERATION

On the Fig. 6.6 the structure diagram of mechanically dithered prism LG is given.

The main element is a ring helium-neon laser. The ring laser design details are represented on Fig. 6.7, 6.8.

Ring laser optical cavity is formed by four TRP. To provide cavity stability two of them - herein after referred to as first prisms (2, Fig. 6.6) - have spherical small leg surface. The other two prisms, second (3, Fig. 6.6), may have one spherical surface (small leg or hypotenuse) or all surfaces may be flat. The Combining prism 4 is used to output from the cavity and to mix a part of counter-propagating waves (CW and CCW) energy. This mixed beam is directed onto the two-plate photodetector. In the plane of the photodetector's sensitive plates interference picture (fringe pattern) is observed.
The TRP are fixed by the optical contact on the sitall (glass-ceramics like cervit or zerodur) body (1, Fig. 6.6). One channel of the body (one arm of the ring cavity) that is vacuum-tightly closed by first prisms, is filled with active helium-neon mixture. The TRP are air-tightly closed by special sitall covers (caps) to keep their surfaces cleanness.

The channel of the body that is located opposite to the active one, is filled with the dried air and the ring laser generation frequency control (cavity path-length control) is carried out in it.

To provide normal operation of the LG a lot of elements and sub-systems besides ring laser are included in it.

The active helium-neon medium is excited with a high frequency gas discharge controlled by the special oscillator (HFO). HFO is supplied through the output signals amplitudes stabilizer.

To ignite (light-up) a gas discharge in the helium-neon medium a light-up high-frequency transformer is included in the structure of the LG. Just after power to be applied to the HFO

eight pulses of 330...350V amplitude and 2...3 μs duration are put to the light-up transformer. Intervals between pulses should be 0.5 seconds minimum.

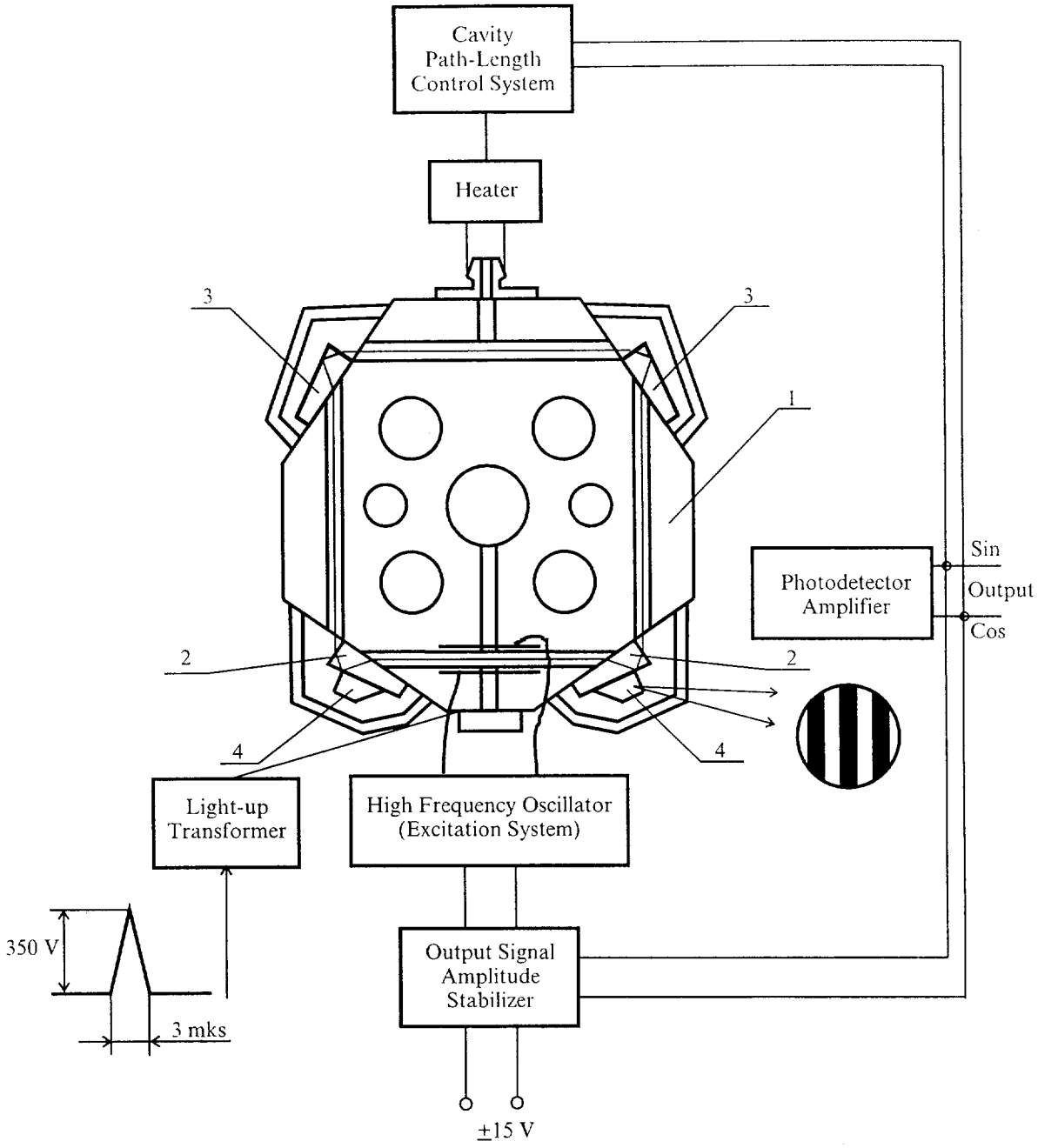


Fig.6.6.
TRP LG structure diagram.

To tune the laser generation frequency at the centre of helium-neon active medium gain line the LG includes a serving element (heater) (Fig. 6.6) which controls a cavity path-length. Main principles of the cavity path-length control are given below.

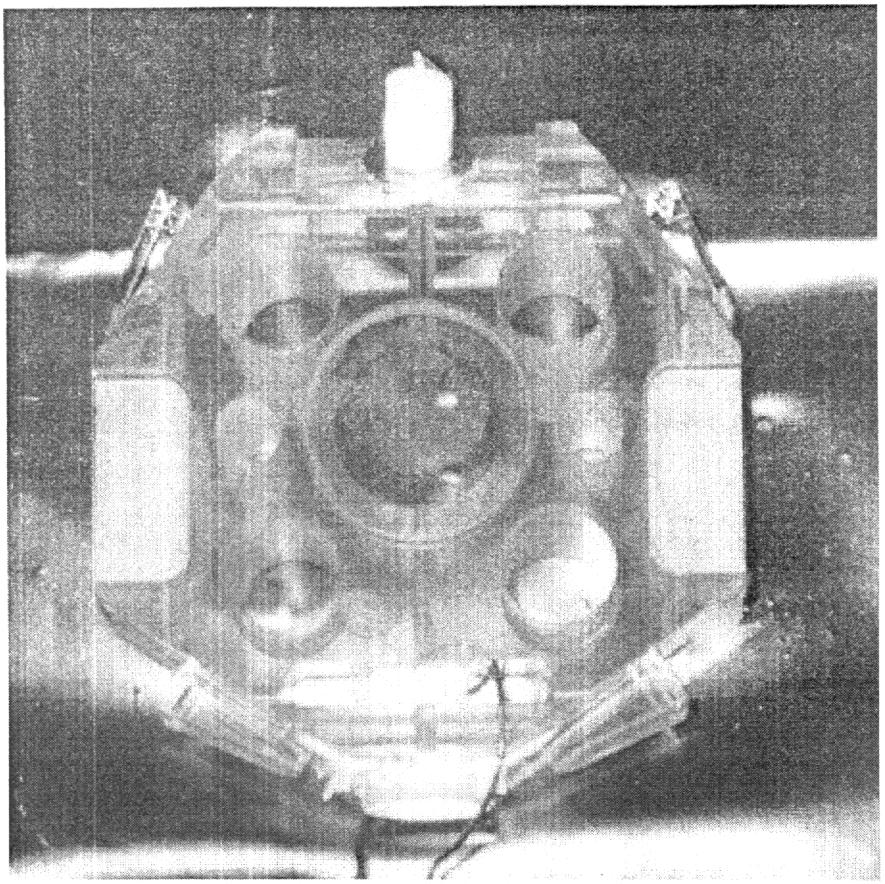


Fig. 6. 7. The Ring Laser with TRP Cavity.

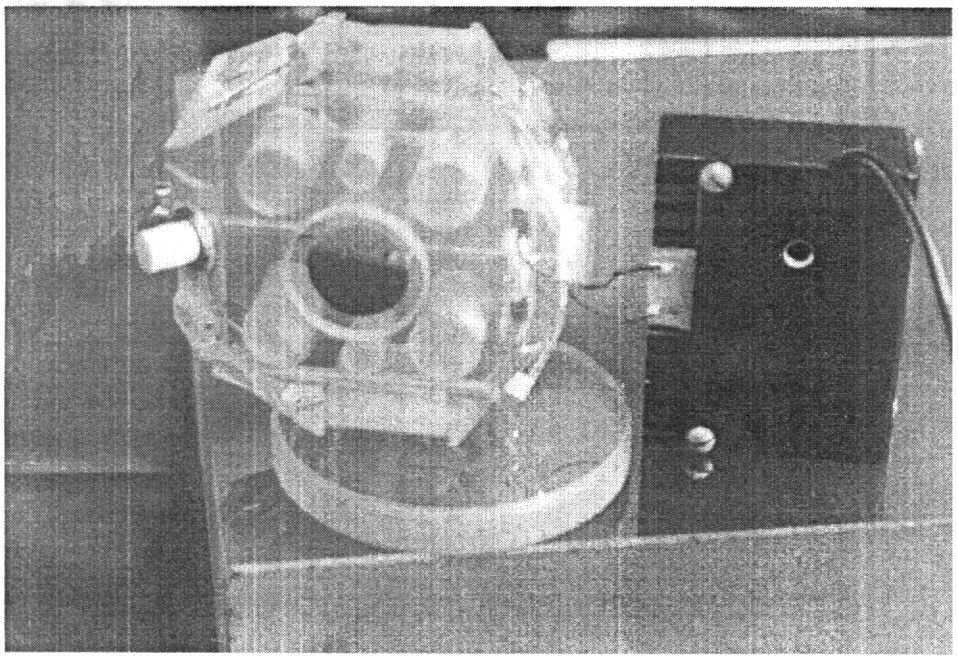


Fig. 6. 8. The Ring Laser with TRP Cavity.

To increase accuracy of the instrument a temperature sensor - a wire thermal resistor - is incorporated in it. With the use of this sensor an algorithmic compensation of the ring laser

temperature error in the external computer may be carried out. Details of the TRP LG temperature model are given in the next Chapter.

• 6.3. RING CAVITY WITH TOTAL REFLECTION PRISMS, TEMPERATURE STABILITY, CONTROL OF THE SCALE FACTOR, TEMPERATURE COMPENSATION OF THE TRP LG ERROR.

During the design of the ring lasers with the TRP cavities it is important to know the variation of the beam path position under exposure of the environment. Beam shifts occur under temperature variations because the prisms material refractive index depends on the temperature. Besides this, density and refractive index of the dried air that fills the channel of the body, being connected with the cavity path-length control system, also change.

It is in essence not possible to prevent this beam shift in the prisms cavity. It leads either to variations of the geometric parameters of the cavity (perimeter and square), that may be accompanied by LG scale factor variation, or to variation of diffraction losses and diffraction non-reciprocity and polarization effects (in the presence of stresses into prisms and under external magnetic field), that is a matter of the zero drift of LG.

This beam shift may be found out by the extended beam matrix (3×3) method [6.4, 6.5].

The beam matrixes of prism cavity elements for meridional plane are given in Table 6.2. Matrix elements are found on the assumption that the beam falls on all refractive surfaces at the Brewster angle.

Table 6.2.

Air interval matrix	$\begin{pmatrix} 1 & L(1+\alpha T_b) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Leg surface refraction matrix for beam traveling from prism into air.	$\begin{pmatrix} \frac{1}{n} & 0 & 0 \\ -\frac{(n^2-1)\sqrt{n^2+1}}{nR_c} & n^2 & -\Delta n_Q + n\Delta n_A \\ 0 & 0 & 1 \end{pmatrix}$
Leg surface refraction matrix for beam traveling from air into prism.	$\begin{pmatrix} n & 0 & 0 \\ -\frac{(n^2-1)\sqrt{n^2+1}}{n^2 R_c} & \frac{1}{n^2} & \frac{\Delta n_Q}{n^2} - \frac{\Delta n_A}{n} \\ 0 & 0 & 1 \end{pmatrix}$
Hypotenuse surface matrix	$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{2\sqrt{2}}{R_h} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

The beam matrixes of the prism cavity in saggital plane are given in the Table 6.3. These matrixes are used to check up the cavity stability only.

Table 6.3.

Leg surface matrix for beam traveling from prism into air.	$\begin{pmatrix} 1 & 0 \\ -\frac{(n^2-1)}{R_c\sqrt{n^2+1}} & n \end{pmatrix}$
Leg surface matrix for beam traveling form air into prism.	$\begin{pmatrix} 1 & 0 \\ -\frac{(n^2-1)}{nR_c\sqrt{n^2+1}} & \frac{1}{n} \end{pmatrix}$
Hypotenuse surface matrix.	$\begin{pmatrix} 1 & 0 \\ -\frac{\sqrt{2}}{R_h} & 1 \end{pmatrix}$

Notations:

$n = n_Q/n_A$;

n_Q - quartz refractive index under the temperature of +25°C (298 K);

n_A - air refractive index under the temperature of +25°C (298 K);

Δn_Q - quartz refractive index variation;

Δn_A - air refractive index variation;

R_c - curvature radius of small leg surface;

R_h - curvature radius of hypotenuse surface.

The coordinates of the beam path in non-disturbed cavity are assumed as zero. In this case the beam passes through the centers of spherical surfaces.

To find beam shifts the matrix of the disturbed ring cavity is calculated relative to any design section. Then the eigenvector of this beam matrix is found out. Components of the eigenvector present the coordinates of the disturbed beam path in the chosen design section. To find the disturbed beam path coordinates on the surfaces of the prisms the eigenvector is to be step-by-step multiplied by the beam matrixes of corresponding cavity elements. If we know the beam coordinates on the refractive and reflective surfaces of prisms it is possible to find the optical cavity square variation and the scale factor variation.

The components of the eigenvector $= \begin{pmatrix} V_1 \\ V_2 \\ 1 \end{pmatrix}$ are defined by following expressions:

$$V_1 = \frac{(1 - M_{22})M_{13} + M_{12}M_{23}}{2 - M_{11} - M_{22}},$$

$$V_2 = \frac{(1 - M_{11})M_{23} + M_{21}M_{13}}{2 - M_{11} - M_{22}},$$

where M_{ij} are the elements of the cavity beam matrix

$$M = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ 0 & 0 & 1 \end{pmatrix}$$

V_1 determines the spatial (linear) beam shift, V_2 - beam rotation («angular beam shift»).

The beam shift equations in analytical form are too clumsy, so for the analysis of concrete cavity design numerical calculations were used.

Parameters of the optical elements for two above mentioned variants of the four-prism cavity with reversal second prisms are given in the Table 6.4. Big leg surface of prisms is always flat, because it is for contact with the sitall block.

Table 4. **Optical & Geometry Parameters of the TRP Cavities.**

Cavity Path-Length, m	0.45	0.28
1st prisms small leg surface curvature radius	1000 mm	
1st and 2nd prisms hypotenuse surface curvature radius	∞	
2nd prisms small leg surface curvature radius	∞	
Length of the cavity arms, mm	110.1 × 102.4	67.7 × 60.0
Radius of the laser (Gauss) beam, mm	0.264	0.225
Scale factor, Hz/(arc.sec/sec)	0.766	0.447

The values of the beam shift for these two TRP cavities under environment temperature variation of +100°C are given in Tables 6.5, 6.6. Positions of the monitoring sections are shown on Fig.6.9.

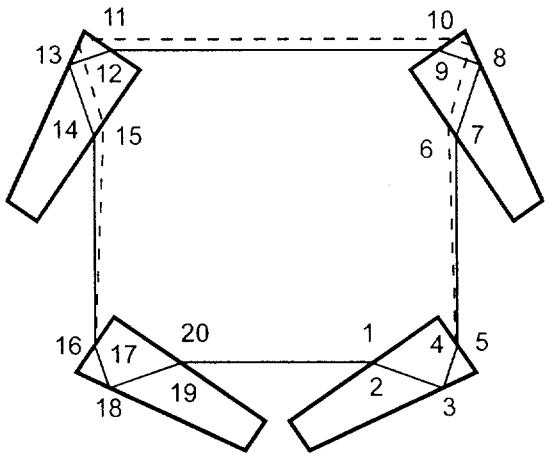


Fig. 6.9

Example of Temperature Beam Shift in the TRP Cavity.

Table 6.5. **The Beam Shift in Meridional Plane for the TRP Cavity with L = 0.45m, μm**

Prism No. 1	X_1	X_2	X_3	X_4	X_5
	3.8	5.5	2.2	0.0	0.0
Prism No. 2	X_6	X_7	X_8	X_9	X_{10}
	-185.3	-270.0	-273.4	-275.3	-188.9
Prism No. 3	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
	-188.9	-275.3	-273.4	-270.0	-185.3
Prism No. 4	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}
	0.0	0.0	2.2	5.5	3.8

Table 6.6. **The Beam Shift in Meridional Plane for the TRP Cavity with L = 0.28m, μm**

Prism No. 1	X_1	X_2	X_3	X_4	X_5
	3.8	5.5	2.2	0.0	0.0
Prism No. 2	X_6	X_7	X_8	X_9	X_{10}
	-100.2	-146.0	-149.4	-151.5	-104.0
Prism No. 3	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
	-104.0	-151.5	-149.4	-146.0	-100.2
Prism No. 4	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}
	0.0	0.0	2.2	5.5	3.8

This beam shift may cause the scale factor variation. However the beam shift leads to either the cavity path-length variation or to the beam path bounded square variation, so the scale factor (for the empty cavity it is $4S/\lambda L$) may be kept constant.

Temperature variations of the cavity path-length, beam path bounded square and the scale factor under 100° heating for two considered TRP cavities are given in the Table 6.7.

Table 6.7. **The TRP Cavities Geometry & Scale Factor variations under 100° heating.**

Exact Cavity Path-Length, mm	Cavity Path-Length Variation, mm	Beam Path Bounded Square, sq.mm	Square Variation, sq.mm	Scale Factor, pulse/arc.sec	Scale Factor Variation, Δk/k
453.289	0.050	11347.039	1.25	0.767142	$8.0\cdot10^{-08}$
283.689	0.049	4133.768	7.10×10^{-1}	0.446552	1.81×10^{-5}

Consideration of the calculations results (Tables 6.5..6.7) leads to the following conclusions :

1. In one arm of the TRP ring cavity the temperature beam shift is negligibly small. In this arm the active medium is based.
2. Both considered TRP cavities have two sections (4-5, prism No.1 and 16-17, prism No.4), where the beam shift equals to zero. In these sections beam shift sensitive optical elements (diaphragms, combining prism etc.) are arranged.
3. Geometry of the TRP cavity is optimized to provide the scale factor stability in the temperature range.

With respect to above mentioned the main temperature error of the TRP LG is the temperature drift. To cancel this error algorithmic temperature compensation with the use of the temperature sensor is to be made.

For such kind of compensation the temperature calibration of the LG is to be made. During the calibration procedure temperature measuring accuracy of 0.5°Ê is required. Temperature calibration must be realised on a special bench precisely reproducing conditions of heat incoming and removing on the object or in the system where the instrument will be finally installed.

The law of temperature drift variation vs. reading of the temperature sensor looks like:

$$\Delta\nu = \Delta\nu_0 + (T_1 - T_0)\, C_1 + (T_1 - T_0)^2 C_2,$$

where

$\Delta\nu$ - constant drift of the instrument at T_1;

$\Delta\nu_0$- constant drift of the instrument at temperature T_0;

T_1 - temperature in the temperature sensor location;

C_1, C_2 - coefficients defined during temperature calibration.

The temperature calibration should be carried out as a set of three measurements of the laser gyro constant drift at three different temperatures (T_0 and two another). At T_0 temperature $\Delta\nu_0$ is measured. Coefficients C_1, C_2 are calculated with respect to the LG performance under above two temperatures.

• 6.4. PRISM LG ERRORS SOURCES; DESIGN FEATURES OF PRISM LG.

Polarization properties.

In the ideal case (isotropic material of prisms, planar resonator) the TRP optical cavity has two eigen states of polarization, which coincide with s- and p-. Pure s-polarized radiation as a rule is not generated because of relatively great reflective losses at the Brewster surfaces of the prisms. The real ring laser has an elliptic eigen state of polarization because of following factors:

– double refraction due to mechanical stresses in the prism which appear as a result of it's fixing onto the sitall body;

– non-planarity of the optical cavity axis line.

The ellipticity is proportional to the stresses value inside the prism and to the non-planarity of the resonator. The elliptic polarization of generated radiation results in magnetic sensitivity due to Faraday effect. In the ring lasers with the TRP optical cavity this effect is amplified because the prism material has the large Verdet constant and the beam path length inside the prism is great (11.8 mm).

Method of connecting the prisms with the sitall body is the main factor the internal stresses in prism material depend on. Three main methods of this connection are known and have been used in different time - cement connection, low temperature ultra-sonic Indium soldering and optical contact.

Cement connection was initially applied before other two and was quickly abandoned because it cannot provide a reliable vacuum-tight connection, has small lifetime and modern degasation technologies cannot be used for cement-assembled units.

The ultrasonic soldering is a more perfect connection method. It has incomparably better vacuum reliability and at the same time it ensures soft connection of prisms and sitall body (block) due to low hardness of Indium. As a result stresses in the prisms due to block deformations (thermal or during its fixing) and the mismatching between the prism's large leg surface and the block's surface (for example, different non-planarity) are compensated.

However this mounting method has its own disadvantages. At first, between the surfaces of the prism large leg and the block there is a split. Because of it radiation reflection repeats many times inside a prism, that raises back-scattering. Secondly, the ultrasonic oscillations of the solder's tip and local heating of the prism periphery during the soldering create in the prism internal stresses. Indium hardness is non-zero, and under variable environment temperature deformation of the prism takes place. This deformation is bending, since the Indium seam surrounds periphery of the prism, and its middle part under effect of external temperature either departs from the block (under heating) or holds down the block stronger (under cooling). Such bending deformation generates stresses inside the whole volume of the prism, including the whole laser beam trajectory in it. Another disadvantage of Indium connection is the work effort for preliminary adjustment and assembly of the ring laser - the prisms should be at first adjusted, and then soldered precisely in the necessary position. Low melting temperature of Indium limits the degasation temperature of assembled unit - ring laser - during vacuum processing up to +140°C.

The optical contact is practically free of all above mentioned defects. It is a reliable vacuum-tight and mechanically strong connection, which accepts high-temperature degasation of the assembled ring laser. The optical contact does not create initial stresses in the prism during its connection with the block. The prism incorporates with the block by the whole surface of its large leg, therefore under effect of environment temperature only small shear stresses appear in the prism near the zone of optical contact. Contrary to [6.6, 6.7], it is experimentally confirmed within many times repeated temperature cycles, that the total stresses in the prism mounted by optical contact are small and stable. In the optical contact connection between the prism and block there is no reflective boundary of media, therefore there are no repeated reflections from the large leg of the prism and additional contribution to back-scattering is not created. Application of the special tools for ring laser assembly by optical contact allows to exclude adjustment of prisms at all. Brief description of these tools is given further.

To realize all above mentioned advantages of the optical contact the dimensions of the sitall block and shape of contacting surfaces of the block and of the prisms are to be provided precisely, the contacting surfaces are to be thoroughly cleaned because their shape mismatching, local defects and contaminations create the additional stresses inside the prisms. Note, that these requirements are to be provided in any case to produce a precise laser gyro.

Peculiarities of TRP Laser Gyro readout.

One of the problems, that is to be solved in the TRP application is how to extract a part of the laser radiation energy outside the optical cavity. It may be achieved either by disturbance of the TR at the reflective surface or by manufacturing the angle of incidence of the refractive surface non-Brewster.

To derive the transmission coefficient of the TR disturbance based combining prism the wave equation for stratified medium is to be solved with respect to the corresponding boundary conditions.

In the Figs.6.10 and 6.11 the reflective and transmission coefficients are shown as functions of a split dimension. On the right diagrams they are shown in details for the split dimension of 1.5 to 2 microns used in practice.

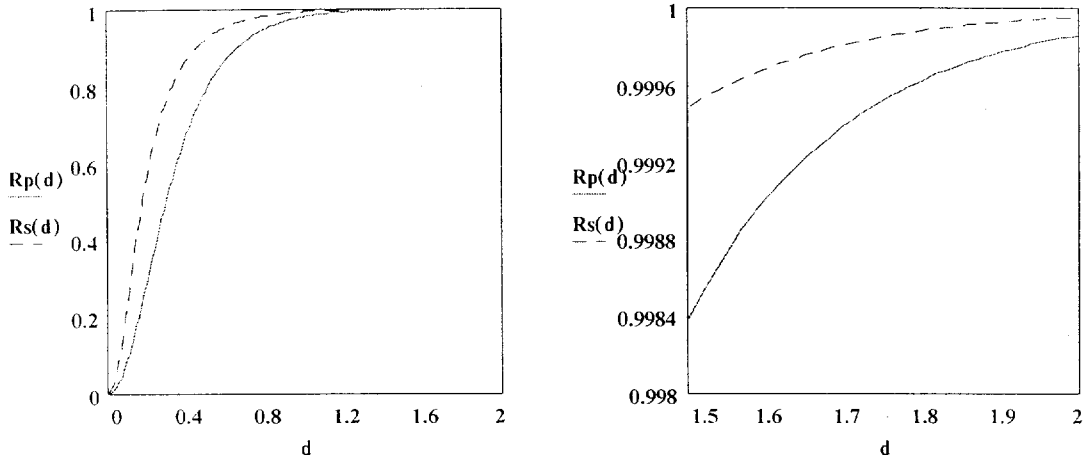


Fig. 6. 10

Dependence of the Reflective Coefficient on the Split Dimension.

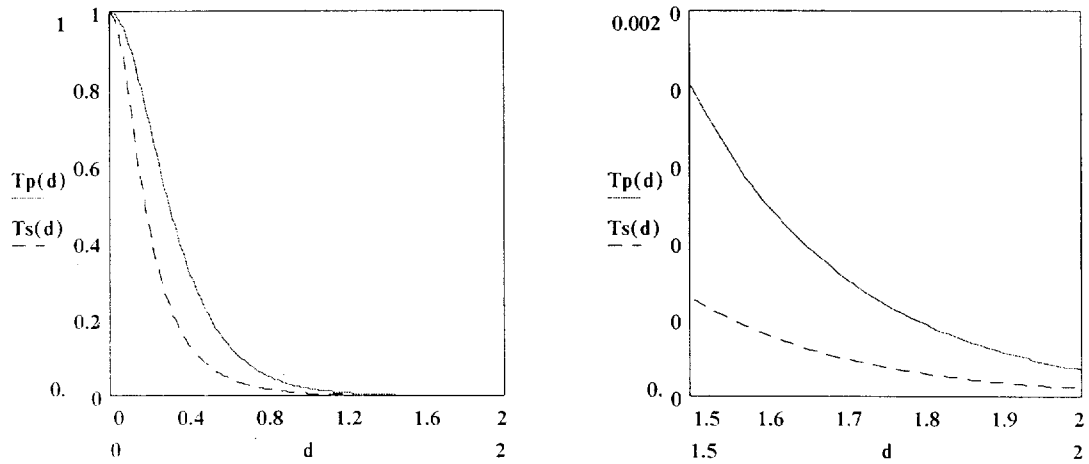


Fig. 6.11

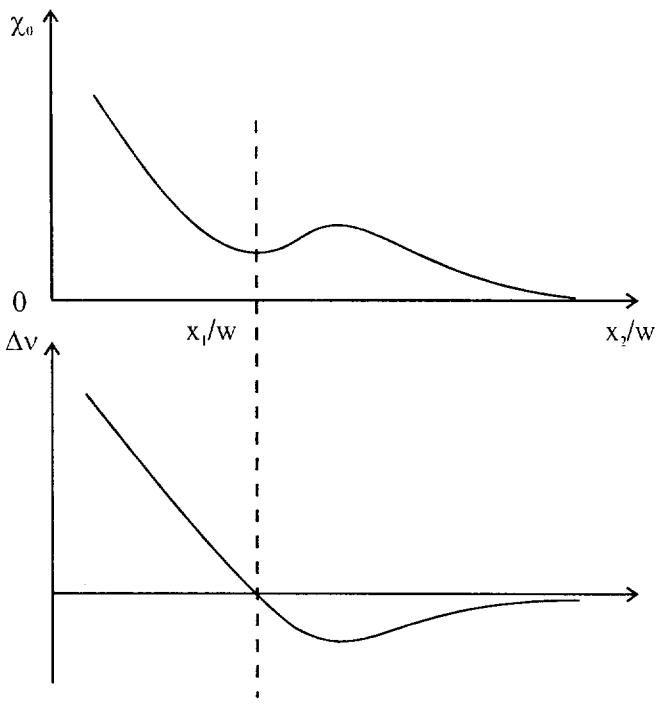
Dependence of the Transmission Coefficient on the Split Dimension.

Embodiment of Diaphragms.

Since the sitall body channels have no means for optical beam constriction as in apertures, transversal multi-mode beam is generated in the ring laser after the prisms have been adjusted. It does not seem expedient to make any apertures in the active channel of the block because they may disturb the gas discharge. The other channels of the block can not be used as apertures or diaphragms because of large temperature beam shifts in these channels. To provide a single-mode generation at the main transversal mode (TEM_{00q}) external diaphragms are applied. They are mounted on the contacting surfaces of the block near the TRP. Each diaphragm is a rectangular piece made of sitall with a polished working side of

cylindrical shape. Diaphragm is to be pressed to the contact surface of the block and adjusted in such a way to provide the working side to touch the laser beam. Single-mode generation is to be provided by moving the diaphragm in a plane normal to the working side of it. The necessary requirement of the diaphragm adjustment is to provide single-mode generation (meaning transversal modes only) with minimal losses being introduced into the cavity. The diaphragm is fixed by cement. One pair of diaphragms provide the mode selection along parallel to the cavity plane axis, another pair - along the normal axis. Diaphragms of both pairs are mounted symmetrically relative to the lateral axis of the active medium (symmetry axis of the ring laser).

An independent adjustment of the diaphragms after mounting of the TRP is made to provide a minimal diffraction non-reciprocity [6.8]. For this purpose one diaphragm is adjusted to provide TEM_{00q} generation only. The adjustment of the second diaphragm is based on new physical effect - increase of lasing intensity and synchronous decrease of diffraction non-reciprocity when it is shifting into the laser beam. Functions of TEM_{00q} losses and diffraction zero-bias vs. the second diaphragm shift into the laser beam are given on Fig. 6. 12.



χ_0 - TEM_{00q} losses;

$\Delta\nu_0$ - diffraction zero bias;

x_1, x_2 - distance between the work surface of first and second diaphragm correspondingly and the laser beam axis;

w - diameter of the laser beam at the diaphragms mounting positions.

Fig. 6.12.

As is obvious TEM_{00q} losses have a minimum (i.e. lasing intensity has a maximum) when the second diaphragm is adjusted symmetrically with the first one ($x_1/w = x_2/w$), and at the same point diffraction zero bias equals to zero.

Besides, the working surface of the diaphragm - rounded off polished edge - provides a minimal back-scattering, therefore application of such diaphragm does not cause the LG lock-in zone to increase.

Cavity Path-Length Control System.

Cavity path-length variation in the LG with the TRP depends on a variation of prisms refractive index and optical beam path-length in them rather than on sitall block temperature expansion (as it is in the mirror LG). This matter causes tens wavelengths cavity path-length variation in the operating temperature range and during the LG self-heating (0.8 $\lambda/^{\circ}C$). Since the TRP deformation leads to an increase of the double refraction and, therefore, to an increase of the polarization losses and magnetic sensitivity, the traditional cavity perimeter tuning method, based on the application of piezo-drive fixed on the reflector, is not suitable in the ring lasers with TRP cavity.

The cavity path-length is controlled by air density variation in one arm of the ring laser (in one channel of the block). The serving element is hermetically evacuated and has a piezo-controlled membrane and a nichrome spiral to heat air inside. The internal volume of the serving element is pipe connected with a channel of the sitall block. A piezoceramic plate glued on the membrane is supplied with a.c. voltage (50 V, 120..180 Hz) to induce air density oscillations within the channel. It causes a ring laser cavity path-length scanning and hence the ring laser frequency scanning over a bandwidth 2 MHz. If the cavity path-length does not fit to the centre of the gain line a synchronous variation of ring laser output signal amplitude is observed. The cavity path-length control sub-system selects the feedback signal by detecting the phase of the amplitude modulated output signal of the LG. The selected feedback signal varies the heating of the spiral in the serving element (heater). A dynamic range of the serving element does not allow to keep the optical cavity path-length stable through the whole operating temperature range. Therefore it is necessary for the cavity path-length controller to be switched back to an average level of 12±1 V when heater voltage of 4±1 V or 20±1 V is reached.

Four-wave mode in the TRP Ring Laser.

The strong dependence of the prism ring laser cavity path-length on the temperature and the limitation of the cavity path-length control system (PCS) dynamic range cause many repeated resets of the PCS even during the LG self-heating under the constant environment temperature. Thus, during the first hour of the self-heating 4 to 5 PCS resets take place.

It is obvious, that after each heater voltage reset the ring laser comes from one generated longitudinal mode to another one. According to above-mentioned principle of PCS operation, this process is sluggish and takes up to 10 s. During this period two longitudinal modes or four waves are simultaneously generated in the ring laser.

In the general case frequency differences between CW and CCW in each mode may be not the same. This results in amplitude modulation of LG output sine/cosine signals at the beat frequency of TEM_{00q} and TEM_{00q+1} CW and CCW frequency differences

$$f_b = (f_{CWq} - f_{CCWq}) - (f_{CWq+1} - f_{CCWq+1}).$$

Such a modulation may cause a glitch of pre-processing system and break in LG output data.

Initially an investigation of four-wave generation in the prisms ring lasers was begun to synchronize frequency differences in the modes and to prevent above mentioned glitches.

A distinctive feature of the four-wave mode in the TRP planar cavity is a large distance between the two longitudinal modes - 670 MHz for cavity path-length 0.45m and 1070 MHz for 0.28m - in comparison with DILAG and ZLG (100 to 150 MHz). This requires to take into account asymmetric factors of laser radiation and active medium interaction. The analysis of the four-wave generation in assumption of semi-classical Lamb theory [6.9, 6.10] regarding asymmetric factors [6.11, 6.12] demonstrated optimal for the above synchronization values of the cavity perimeter - near 0.28m and 0.42 m, optimal active medium pressures for each perimeter - 330 Pa for the cavity path-length of 0.42 m and 800 Pa for 0.28m, and monotonous evaluation of the synchronization ability upon the laser gain growing. It was found out that at ring laser perimeter of 0.33 m the four-wave generation can not be synchronized in essence, because for this cavity distance between longitudinal modes equals to the Ne^{20} and Ne^{22} isotopic shift (distance between Ne isotopes gain lines centers). This results in CW and CCW waves competition and single-directional generation.

The obtained results are experimentally confirmed in the laser gyros with cavity path-length of 0.28m, 0.36m and 0.45m. The above mentioned optimal values are used in quantity produced instruments - information loss problem has been overcome.

The analysis of the main errors of the LG in the two-wave (TWM) and the synchronized four-wave modes (SFWM) was made. It was found out that the LG zero bias caused by amplitude diffraction non-reciprocity for the SFWM is 1.5...2 times less than for the two-wave mode. Zero bias caused by phase diffraction non-reciprocity for the SFWM is 15...20% less. Diffraction zero bias dependence on the ring laser generation frequency for the SFWM is approximately 1.5 times weaker.

The experimental results confirm that the SFWM may be used as a main measuring mode of LG. Moreover, sometimes it may cause lock-in reduction also. The lock-in minimum as a rule does not coincide with the gain line center, and the center of the SFWM may be nearer to the lock-in minimum.

Numeric analysis demonstrated, as in [6.13], that in a case of asymmetric location of the longitudinal modes in the gain line the weak mode «destroys» lock-in of the strong mode, and "self-biased" LG with zero static lock-in may be obtained. Experimental investigation of these results seems very important. For this purpose we are designing a PCS to provide a reliable stabilization of the required value of cavity path-length.

Lateral High Frequency Gas Discharge with External Electrodes.

At presence there are three main methods of gas active medium excitation: direct current high voltage discharge (DCD), high frequency discharge (HFD) and combined discharge. In case of a DCD discharge source is a high voltage electric field of necessary magnitude to provide break down of the gas and sustaining of discharge. In case of a HFD a high frequency electro-magnetic field is used for the same purpose. In combined discharge principles of DCD and HFD are used together.

Let's compare the main characteristics of DCD and HFD for the excitation of the active medium in the ring lasers with TRP cavity.

DCD technique is very well investigated and widely applied in a lot of LG, so note that this method of excitation has one indubitable advantage - stable gas discharge (i.e. active medium) of arbitrary length may be provided without any problems of electro-magnetic compatibility with the other LG.

The DCD disadvantages for application in LG are also well-known - it is necessary to have high-voltage power supply and arms currents balance stabilizer. Circuitry of these devices is well-known, and the devices are quantity produced and widely applied.

Just one DCD disadvantage is very important for application in LG - practically it is difficult to create a short active medium, for example, of 5 mm length, because three DCD electrodes - cathode and two anodes - are to be mounted at such a small length, and problems of active atom flow, especially in closed-to-electrodes areas, are to be solved.

For our opinion, for application in the prism ring lasers with one active arm lateral HFD has a lot of advantages.

1. HFD allows to provide the necessary electric field strength to maintain a plasma pole with relatively low supplied voltage. This decrease of power supply voltage favours the gas discharge tube lifetime increase.

2. The HFD has 15...20% greater quantity of Helium active atoms in the meta-stable energy level 2^1S and 2^3S range in comparison with DCD [6.14].

In [6.15] there is experimentally demonstrated that in HFD direct excitation of Ne active atoms higher levels is more prevailing than in DCD - population inversion of Ne working levels pair is obtained by weak occupation of lower level. This conclusion based on the results of external walls temperatures measuring of DCD- and HFD-excited lasers. Thermocouple method of measuring demonstrated $t_{DCD}/t_{HFD}=3$.

3. Active medium of small length may be created. It allows to decrease LG zero bias components, which are proportional to the active medium length (diffraction zero bias etc.).

4. Problem of active atom flow is absent in essence. As against of axial HFD closed-to-electrodes areas are absent, and gain redistribution causes a 15% average increase of assumed discharge pole gain; since there are no concentration gradients along the discharge pole axis (ignoring boundary effects), nothing causes an active atom flow in the axial direction.

The well-known HFD disadvantage - increased walls absorption of Ne atoms and electro-magnetic compatibility with external equipment - may be overcome. During the vacuum processing the active channel of the ring laser is to be filled with Neon or a Helium-Neon mixture and to be processed by strong HFD within 8 hours. By this way two useful effects are obtained - the discharge channel walls are cleaned from impurities and saturated with Neon to prevent wall absorption in the LG working mode. The decrease of electro-magnetic interference for external devices is made by the circuit engineering when HF oscillator for excitation is designed and by matching of HFO output and its loading - discharge channel of the ring laser.

To ignite (light-up) the HF discharge the above described high-frequency light-up transformer is used. It forms a light-up pulse of 1.5...2 kV amplitude and 1.5...3 μs duration.

The application of the lateral HFD with external electrodes allows to make the vacuum volume of the ring laser free from metallic details. This makes it better for cleaning and leads to increase of LG lifetime.

Moreover, the variation of the external electrodes shape and dimensions may cause an optimization of the gas discharge characters (for example, gain distribution). Some characters of the lateral HFD are shown below.

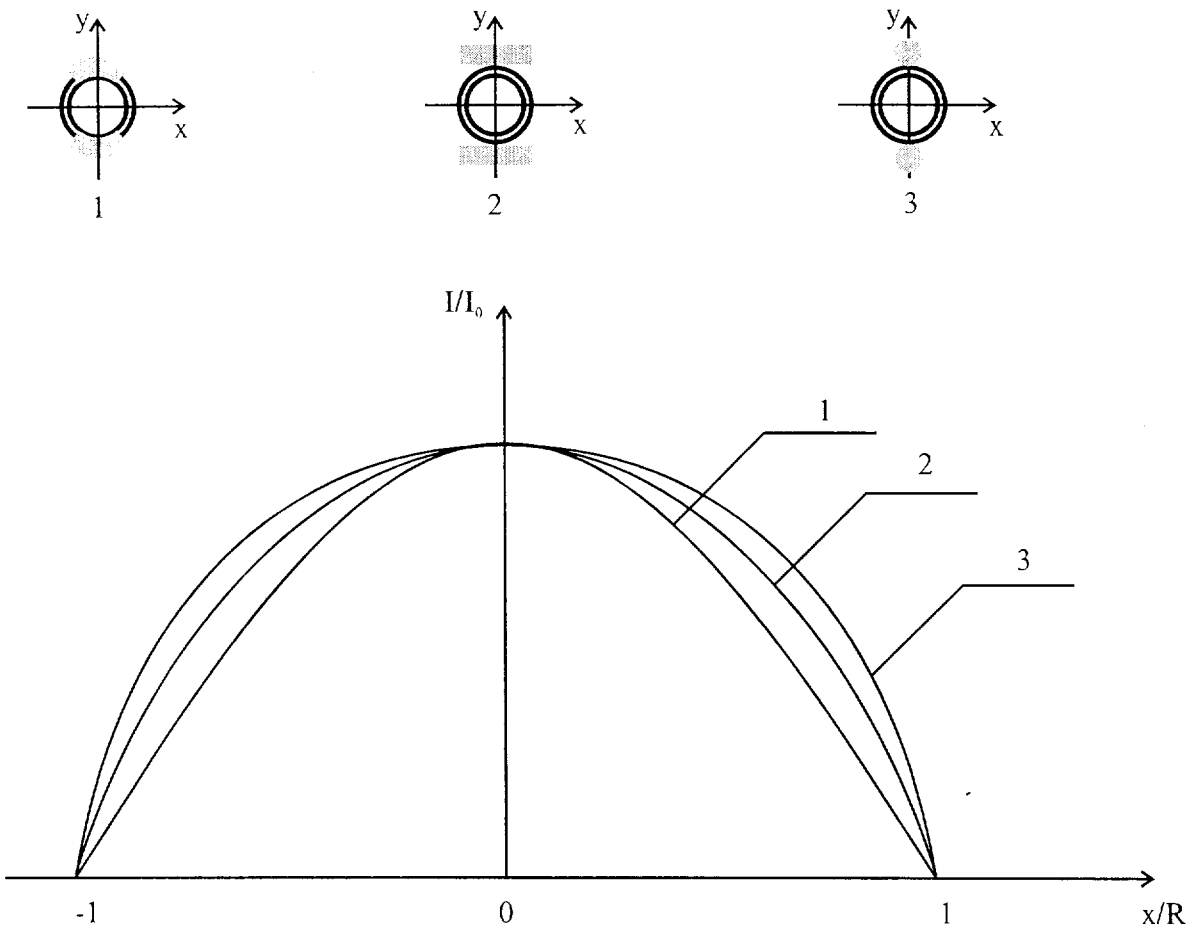


Fig. 6. 13
Radial Gain Distribution Dependence on Electrodes Shape.

Fig. 6.13 illustrates radial gain distribution dependence on electrodes shape. To provide uniformly distributed gain wire electrodes are to be used. But in this design a large part of the HF energy is radiated outside the cavity (into the ether), that cause a problem of electro-magnetic compatibility. For real applications the stripline electrodes seem optimal.

On Fig. 6.14 the radial gain distribution vs. the discharge channel diameter is shown. For channel diameters more than 5 mm gain collapse on the active channel axis is observed. For practically applied diameters (3 to 5 mm) this effect does not occur.

Fig. 6.15 illustrates the choice of HFD frequency. As it's obvious, for 150 MHz the radial gain distribution is more uniform.

As it is shown on Figs. 6.16a and 6.16b, the axial gain uniformity may be achieved by electrodes design. For the electrodes of uniform thickness gain irregularities along the HFD axis are observed (Fig. 6.16a). Special profiling of electrodes (Fig. 6.16b) results in a more uniform gain distribution.

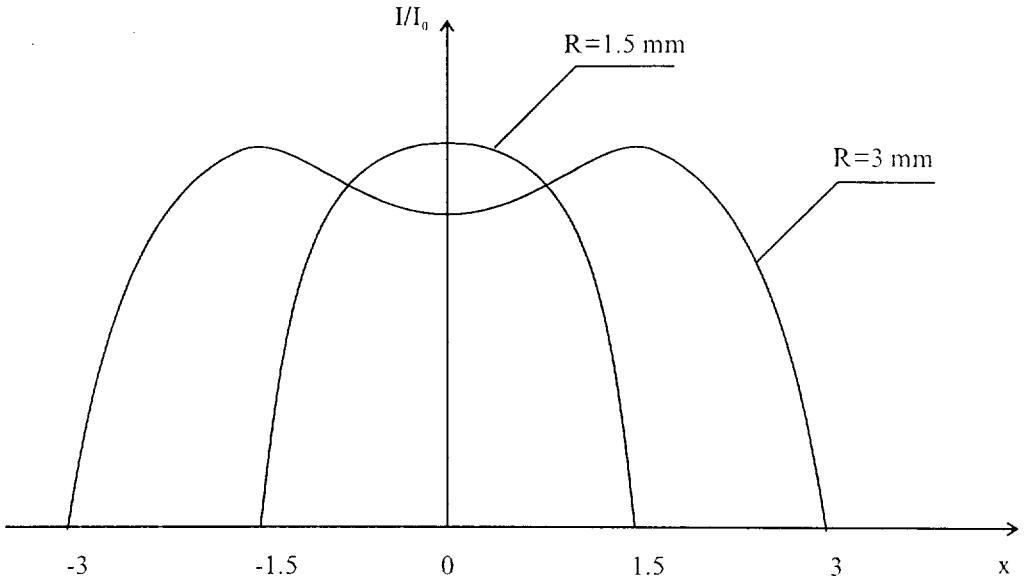


Fig. 6. 14.
Radial Gain Distribution Dependence on Discharge Tube Dameter.

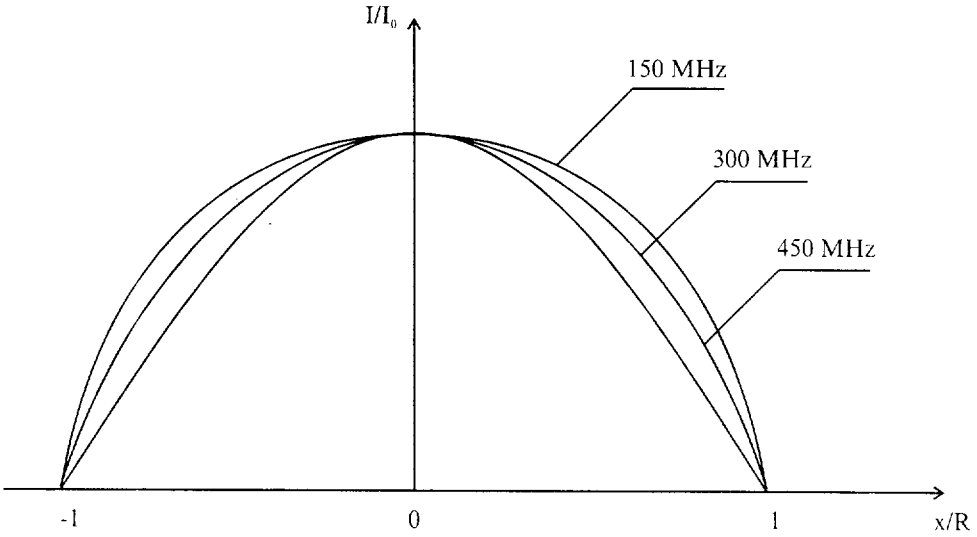


Fig. 6. 15.
Radial Gain Distribution Dependence on HFD Frequency.

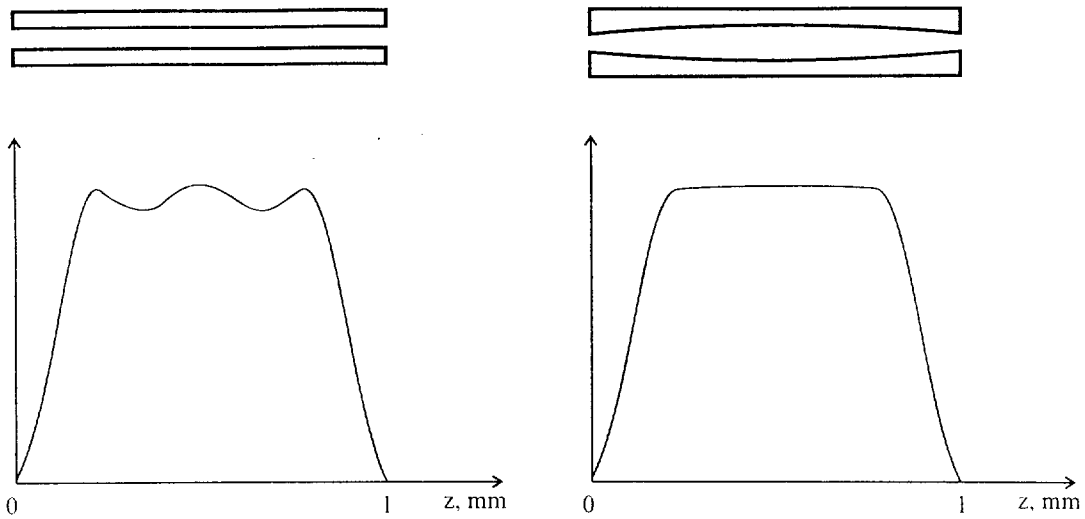


Fig. 6.16.
Axial Gain Distribution for Different Profile of Electrodes.

• 6.5. TECHNOLOGY FEATURES

Manufacturing of Precise Optical Components.

Optical components being applied in precise prisms laser gyros must provide highest requirements for surface roughness, optical cleanness (quality) class, surface flatness and, moreover, their interchangeability is to be provided by precise fabrication of linear and angular dimensions. Characteristic features of these components production technology follow from above mentioned, namely - free abrasive lapping with increased material removal (deep polishing), polishing on the pitch lap with the polishing abrasive water suspension, application of components processing in blocks based on special tools (when a lot of components are fabricated simultaneously in one process).

Highest requirements are imposed for the sitall block and TRP. Block linear dimensions fabrication accuracy is 0.02 mm, angular dimensions - 5 arc.sec., optical contacting surfaces for the TRP mounting roughness - R_z 0,063, non-flatness N=0.5, local flatness error ΔN=0.3, optical quality class PII. To provide the mentioned requirements cube technology is used for making of the sitall block. According to this technology lapping and polishing of each contacting surface of the block is made simultaneously for four blocks, which are mounted by optical contact at side surfaces of a polished cube, made of glass. Accuracy of the angles between the cube surfaces is of 2 arc.sec. Besides high productivity this technology provides the required flatness of the surfaces, because the cube with mounted blocks has a large processed surface, and guarantees precise angular dimensions of each block because the cube was made with high accuracy. Thus, the technology procedure amounts to precise mounting of the sitall blocks on to the cube, that is made with the special reference prisms and plates; lapping and polishing follows.

The required accuracy of TRP angular dimensions is also 5 arc.sec., refractive and reflective surfaces roughness - 2..5 Angstrom, non-flatness N=0.5, local flatness error ΔN=0.3, optical quality class P0-10*. Moreover, special requirement of sphere misalignment is imposed to the prisms with spherical surfaces. To provide the mentioned requirements special prism tools are used and components are processed in block. This means that one working surface of each TRP workpiece is lapped and polished, then each workpiece is to be optical contact mounted onto the prism tool and each tool is also mounted by the same method onto the

glass faceplate. On the faceplate there are based 56 to 58 workpieces. Then lapping and polishing are made. To provide optical quality class P0-10* deep lapping is made, i.e. increased material removal is made to lap off the whole crack layer, that has been originated during previous grinding and lapping. Angular dimension accuracy is provided by the prism tools, which are made with an accuracy of 2 arc.sec. TRP with spherical surfaces are made according to the same production procedure, but for them prism tools and faceplate with spherical surfaces are used. Sphere misalignment accuracy is provided by a making of the spherical surface of the tools and an accurate optical contact mounting of the TRP workpieces on the tool.

Among the optical components of the prism laser gyro there is the combinig prism with special requirement according to defects on one of its edges. No luminous defects must be seen on this edge when the laser beam is falling on it. To provide this requirement the combining prism is produced according to the block of components processing technology with application of special prism tools and deep lapping of the edge adjoining surfaces.

As it follows from above, when prism tools are applied, the angular dimension accuracy is provided «automatically» by the tools. To provide accuracy of linear dimensions during lapping and deep lapping stock removal of material is to be checked. So, during the polishing only the flatness and the optical quality of processed surfaces are to be inspected. Since dimensions of the processed surfaces are large, the final non-flatness of N=0.2 (around $\lambda/20$) may be provided.

The above mentioned technology features result in optical components production labour increase of 10%. However, improved by such a way production process allows to obtain 2 times decrease of internal stresses inside the TRP material after their optical contact mounting on the sitall block and to substantially reduce the back-scattering, that leads to the LG lock-in threshold decrease.

Vacuum Processing of the Ring Laser;
theTRP Adjustment During the Ring Laser Assembly.

Accuracy and lifetime of the LG substantially depend on the ring laser assembly technology. The ring laser consists of the sitall block, the sitall covers and four TRP, which are mounted on the block by optical contact, the diaphragms, the combining prism and the photodetector.

The assembly procedure of the ring laser is divided into three parts:

– the sitall block vacuum processing;

– the TRP adjustment and mounting;

– the diaphragms and combining prism adjustment and mounting, the mounting of the sitall caps;

– the adjustment and mounting of the photodetector.

The ring laser assembly technology must provide strict requirements for the accuracy and lifetime of the laser gyro, the accuracy and operating characteristics repeatability from one LG to another. The TRP laser gyro design allows to apply vacuum processing and assembly technologies mostly complying with the above requirement.

The fulfillment of strict requirements to non-flatness, roughness and surfaces cleanness (quality) of the laser gyro optical components allows to use for their connection a method of

optical contact. The mounting of the components by optical contact results in obtaining the guaranteed vacuum-tight connections, excludes the infiltration of any vapors into the gas mixture during the vacuum processing, the degasation and the operation, ensuring an invariance of its composition.

During the assembly the optical components are vacuum processed. The vacuum processing consists of high temperature degasation and ion-plasma cleaning of the internal surfaces of the sitall block by the powerful high-frequency gas discharge to provide consequently in the laser gyro stability of the gas mixture and the accuracy characteristics and a large lifetime.

Vacuum processing of the sitall block is made at special oil-free vacuum equipment. Its design allows to evacuate the block vacuum volume to $6.65 \cdot 10^{-6}$Pa and fill it with the necessary gas mixtures (Ne for vacuum processing or He-Ne mixture for operation).

The sitall has two peaks of gas emission : under t=+150°C and t=+250°C. In the ring laser design only two types of connection are used - optical contact for optical components connection and diffusion lead soldering to connect the sitall block and the glass exhaust tube. Both connections are resistant to the environment temperatures up to more than +350°C, so vacuum processing under both temperatures of maximal sitall gas emission may be made and the most effective degasation may be obtained.

Vacuum processing is made in two stages, which are divided by the adjustment of the first prisms (closing the gas discharge channel of the block).

At the first stage the sitall block with technological sitall plates mounted on the contacting surfaces by optical contact, is connected through the glass exhaust tube with the evacuation and gas filling vacuum equipment. The block is heated up to t = 250°C with the temperature increase speed of 4°C/min. maximum. Then it is degased under the temperature of +250°C and continuous evacuation within 6 hours. Then vacuum concavity and the active channel of the block are filled with gas mixture and are treated by the powerful high-frequency discharge within 8 hours.

The adjustment of the first TRP is one of the main ring laser assembly procedures. Prisms are to be adjusted to provide the laser beam passing precisely along the axis of the gas discharge channel of the block. Proceeding from the optical cavity analysis and the sitall block geometry design, the first prisms are to have a definite position on contacting surfaces of the block.

To adjust the TRP special tools are used, which enable the prism adjustment by simply mounting of them onto the block.

The reference sitall block is mounted onto the polished flat sitall plate «on the colour» (with a very small split to see in the block/plate contact zone a fringe pattern of 5 bars maximum). Near the contacting surfaces of the block at the side of its active channel two special props are mounted onto the plate by optical contact in such a way as to observe a fringe pattern of 5 bars maximum in the contact zones of the block and the props. The height of the props is designed to provide a coincidence of the height middle of the TRP (being placed on the prop) and the block's active channel axis. Then special stops are mounted in optical contact onto the props at S distance from the side surfaces of the block. The required S distance is computed with reference to the optical cavity design and the sitall block geometry to provide the required configuration of the optical cavity after the TRP are placed on the props in their working position and held down to the stops and to the contact surfaces of the block. The stops are adjusted by gauge blocks and 10 microns accuracy of S distance is provided.

Then the reference block is taken off the plate. The adjusted and assembled tool set is to be applied many times for the adjustment of the first TRP.

The adjustment of the TRP by means of above described tool set is carried out to following procedure. The sitall block for assembly is mounted onto the sitall plate «on the colour» and pressed against both props in such a way as to provide in the block/plate and block/props contact zones fringe patterns consisting of 5 bars maximum. The TRP for adjustment is side surface placed onto the prop «on the colour», the small leg surface held down to the stop and then the large leg surface is mounted onto the block by optical contact.

This adjustment technology is very simple, because there is no adjustment in it - the TRP are mounted in optical contact onto the sitall block in positions, which are precisely and strictly determined by the adjustment tool set. Taking into account, that the batch technology of sitall block production with application of the reference tools provides the block to block dimensions repeatability with error of 10 microns maximum, it is obvious that the above mentioned adjustment technology provides good repeatability of the LG to LG ring laser optical cavity geometry and, hence, of the LG to LG accuracy characteristics.

To achieve high accuracy characteristics of LG with above adjustment technology two requirements are to be provided:

- the TRP quality, i.e. their geometry and optical quality of the surfaces, is to comply with design requirements and the ability of the above parameter inspection during the adjustment is to be provided;
- the inspection of the optical cavity flatness (planarity) during the adjustment is to be provided.

For this purpose special inspection gadget is used during adjustment. It presents a sitall block, that has two TRPs mounted by optical contact onto the contacting surfaces and an active gas discharge channel assembled according to the vacuum processing procedure and filled with He-Ne mixture.

To inspect the TRP quality the inspection gadget is mounted into the adjustment tool set. Two TRPs for inspection are side surface placed onto the props «on the colour», the small leg surface pressed against the stops and the large leg surface held down to the two contacting surfaces of the block to provide between the block and the large leg a fringe pattern of 5 bars maximum. By this way the closed ring cavity is adjusted, and lasing appears. In the absence of diaphragms and apertures the laser beam has a large diameter. In this position it's possible to inspect

- the TRP material quality;
- the quality of the TRP surfaces optical processing;
- the quality of the TRP chemical cleaning.

On the prism surfaces in the lasing zone there must be no crack or group of small cracks or chip-out. Traces of chemical cleaning, dust particles etc. must be absent. Laser radiation intensity must provide the production procedure requirements. Weak intensity may be caused by increased radiation absorption inside the prism material, absorption or scattering on the prism surfaces connected with presence of thin oxide layers or wrong polishing quality.

Prisms with the above mentioned defects are not acceptable for the ring laser assembly. They

are to be polished or chemically cleaned again or to be rejected.

This prisms quality inspection may be arranged directly during the adjustment procedure. For this purpose during the adjustment of the first prisms the above mentioned inspection gadget is mounted on the same sitall plate near the back side of the block being assembled (at the side of PCS channel) in such a way that the prisms of the inspection gadget and prisms being adjusted form a ring cavity. Then in the inspection gadget the gas HFD is ignited. Lasing appears. In this position quality of being adjusted TRP is inspected in a view of all above mentioned characteristics and lasing threshold parameters are also inspected.

If all required above mentioned parameters are provided, the TRP are to be mounted in optical contact. In the optical contact zone non-transparent spots with character size more than 0.5 mm are not acceptable. Double refraction of prisms is measured by special polarimeter. Along the laser beam trajectory double refraction greater than 2...6 nm/sm is not acceptable.

Adjustment of the second TRP (near the PCS channel) is made after the block with the first prisms has been vacuum processed and filled with the He-Ne mixture. For this adjustment an analogous tool set is used as for the first prisms adjustment.

The tools set for the adjustment of the second TRP may be applied many times.

For the adjustment of the second prisms the block with first prisms is mounted on the sitall plate «on the colour» inside the tools set. HFD in the active channel of the block is ignited. The second TRP are side surfaces placed onto the props, the small leg surfaces are pressed against the stops and the large leg surfaces are held down to the contacting surfaces of the block to provide between the block and the large leg a fringe pattern of 3 bars maximum. Lasing appears. The second prisms are to be inspected. If their quality comply with the requirements, they are to be mounted by optical contact and covered by protective caps. The block is taken off the sitall plate.

For the assembled ring laser the following parameters are to be inspected:

– the excitation values corresponding to the lasing threshold U_{thr} and I_{thr};

– the laser radiation intensity;

– the lock-in threshold, measured in one beam (in one wave);

– the ring cavity planarity;

– the view of the optical contact of the TRP and the block;

– the double refraction inside the prisms along the laser beam trajectory.

These parameters must comply with the production procedure requirements of the LG.

The design of the combining prism is shown oh the Fig. 6.17.

The combining prism is a rectangular prism. Two surfaces of this prism enclose an angle of 179°59'. One of these surfaces is intended to let pass a part of the radiation to the outside by means of TR disturbance, another one - for the combining prism mounting onto the TRP by optical contact.

By means of a special fixing gadget the combining prism is placed on the hypotenuse surface of the first TRP. Then the combining prism is shifted to the laser beam and rotated to provide the matching of the two output beams coming off it and obtaining the contrast fringe pattern.

The position of the combining prism determines the output radiation intensity and threshold excitation voltage U_{thr}. Values of above intensity and U_{thr} are to be inspected and must comply with the ring laser assembly procedure.

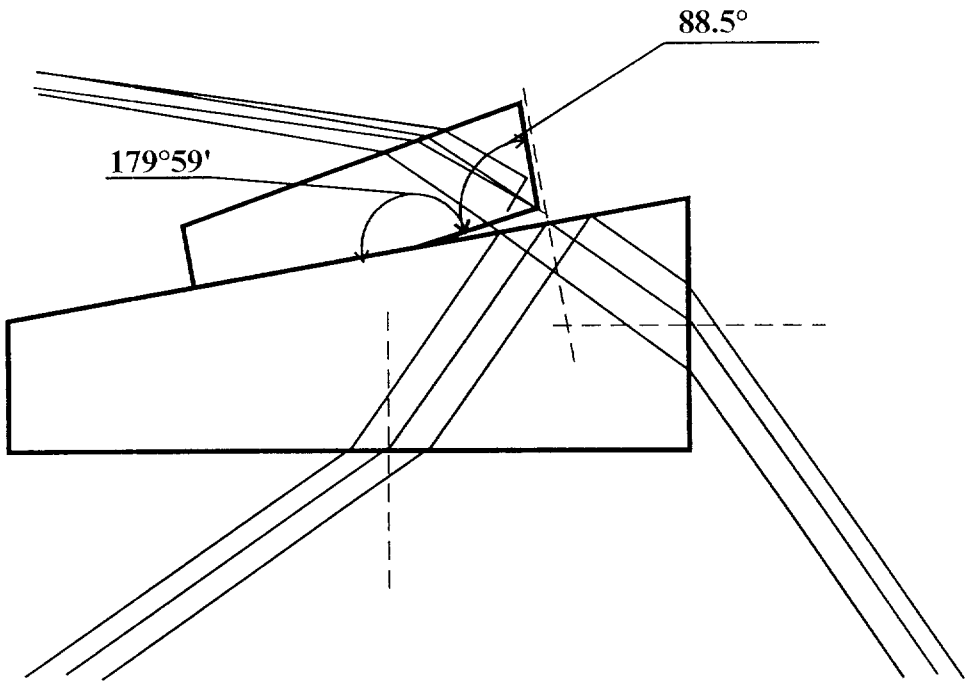


Fig. 6.17

Configuration and Mounting of the Combining Prism.

During the combining prism production much attention is paid to the quality of the edge, at which the laser beam splitting takes place. This edge must not have any spall or trace of dirt, that is visible inside the laser beam. The above edge defects result in back-scattering and lock-in threshold increase. So, during the combining prism adjustment the edge quality is to be inspected visually by means of magnifier. The combining prism is fixed on the first TRP by optical contact. The first TRP with mounted combining prism is covered by the sitall protective cap, that has a window to output the combined beam onto the photodetector.

The rest of the TRP are also covered by protective sitall caps, which are fixed on the block by ultra-sonic Indium soldering. The caps are designed in such a way to provide minimal splits between the TRP surfaces and the internal surfaces of the caps. Ultra-sonic Indium soldering application makes the connection reliable and simple for assembly and disassembly.

Two caps, which protect the first TRP, have holes, and couplings are soldered to them by means of ultra-sonic Indium soldering. Through these couplings side channels of the block are evacuated. This evacuation excludes any convection flow inside the above channels under the LG self-heating or under external temperature variations.

• 6.6. APPLICATIONS, FUTURE DIRECTIONS.

The LG with TRP cavities are applied in traditional fields. I42-1-C strapdown navigation

system based on above LG is certified and quantity produced for civilian aircraft IL-96-300 and TU-204.

The development of the TRP cavities with high Q-factor, small polarization non-reciprocity and low back-scattering allows to increase the laser gyros accuracy and opens new fields for their application - precise measuring systems for railways and for oil and gas pipe lines.

The main requirement is imposed to the measuring systems for railways - high accuracy under hard operation conditions, because the user wants to install the measuring system directly to the axle box of the vagon. Now in Russia the prism laser gyros with accuracy of 0.003°/hr are widely applied in profile measuring systems mounted inside the railway car. There is an experience of prism LG applications in the measuring systems mounted on the axle box. Obtaining of the accuracy of 0.002...0.001°/hr allows to make precise LG based systems for high-speed railway profile measurements.

Large oil and gas pipelines are laying at the depth of 4 to 5 meters under ground. During the year pipe line season shifts take place. This may cause serious accidents because of pipe line cracks etc. To prevent these accidents pipe line movements are to be monitored with accuracy of ±2 sm. GPS can not be used for this purpose, because the pipe line is under ground. To solve this problem small sized (up to 200 mm diameter) precise navigation system moving within the pipe is to be used. Required accuracy is provided by means of a combination of a navigation system and a distance sensor outputs and the precise positions of the pump-stations measured by the GPS, in the post-mission processing mode.

To provide more wide application of the prism LG the following development directions seem very important:

- improvement of the optical technology to provide the sitall block contacting surfaces and the TRP contacting surfaces non-flatness of $\lambda/100$ in quantity production to decrease stresses inside the prisms material in optical contact;
- providing of the LG drift repeatability of 0.001°/hr in quantity production to expand their application in gas and oil pipe lines monitoring systems and railway profile measuring systems;
- development of the four-wave generation in the prisms LG to create a self-biased LG without dither. Since there are no magneto-optic non-reciprocal elements, the cavity of this LG has the total losses of 50 ppm level. The level of these losses in combination with a decrease of random walk because of dither absence allows in essence to obtain an accuracy of 0.001°/hr in a small-sized laser gyro.

References

[6.1] Аблеков В.К., Соколов С.А. «Кольцевой четырехугольный оптический квантовый генератор» Авторское свидетельство СССР № 204456, кл. H01 S 3/08, 1967.

[6.2] Шафер Ф.П., Мюллер Г. Кольцевой лазер с регулируемым цветом луча.- Optics Communications, v.2, No.8, 1971.

[6.3] Aronowitz, F. Chapter «The Laser Gyro» / Laser Applications. Edited by Monte Ross, New Yorrand, London, 1971.

[6.4] Gerrard, A., Burch, J.M., "Introduction to matrix methods in optics". - Wiley-Interscience Publication, John Wiley & Sons, London, New York, Sydney, Toronto

[6.5] Ищенко Е. Ф."Открытые оптические резонаторы: Некоторые вопросы теории и расчета". - М.: Сов. радио,1980.

[6.6] Индисов В.О., Курятов В.Н., Семенов Б.Н.и др. "Исследование поляризационных характеристик лазерных призм полного внутреннего отражения". - Оптика и спектроскопия (Optics & Spectroscopy), v.75, No.1, 1993, pp.201-215.

[6.7] Индисов В.О., Курятов В.Н., Семенов Б.Н.и др. "Исследование поляризационных характеристик лазерных призм полного внутреннего отражения". - Оптика и спектроскопия (Optics & Spectroscopy), v.75, No.2, 1993, pp.451-460.

[6.8] Бакин Ю.В., Зюзев Г.Н., Людомирский М.Б.и др. "Способ селекции основной поперечной моды кольцевого лазера". - Авторское свидетельство СССР № SU 1829837, H01 S 3/083, 12.1990.

[6.9] Lamb, W.E., Jr., "Theory of an optical maser", Phys.Rev., vol.131, p.p.A1420-A1450, June 15, 1964.

[6.10] Волновые и флуктуационные процессы в лазерах / Под ред. Климонтовича Ю.Л., Москва, Наука, 1974.

[6.11] Кристаллов А.Р., Мельников Л.А., Тучин В.В."Асимметрия выходных характеристик газового лазера, вызванная столкновениями".- Опт. и спектроскопия (Optics & Spectroscopy), v.48, No.1, 1980, pp.89-92.

[6.12] Дорофеев И.Б., Зюзев Г.Н., Соколов В.А., Курятов В.Н., Фрадкин Э.Е."Асимметрия взаимодействия волн кольцевого лазера в двухмодовом режиме генерации".- Опт.и спектроскопия (Optics & Spectroscopy), , v.60, No.11, 986.

[6.13] Chow ,W., Gea-Banacloche, J., Pedrotti, L., Sanders, V., Schleich, W.and Scully, M., «The ring laser gyro», Reviews of Modern Physics, vol.57, no.1, pp61-104, January, 1985.

[6.14] Муллер Я.Н., Геллер В.М. - Квантовая электроника, 1977, т.4, №7, с.1013.

[6.15] Юдин В.Н. - Квантовая электроника, 1973, №3(15), с.131.

7. APPLICATION OF THE SAGNAC EFFECT IN THE INTERFEROMETRIC FIBER-OPTIC GYROSCOPE

Hervé C. Lefèvre

EuroFOG - PHOTONETICS

52, Avenue de l'Europe, BP 39

78160 Marly-le-Roi - France

ABSTRACT

This chapter reviews the basics of the interferometric fiber-optic gyroscope (I-FOG) : fundamental principle based on Sagnac effect, importance of reciprocity and single-mode propagation, analysis of coherence and polarization problems, signal processing techniques. It also describes the technological progresses of guided-wave components (single-mode optical fiber, semi-conductor diode source, integrated optics, in-line fiber components). Finally, recent trends like multi-axis configurations and rare-earth doped fiber source are presented.

1. INTRODUCTION

Over the last twenty years, the interferometric fiber-optic gyroscope (I-FOG) has evolved from a pioneering physics experiment [1] to a practical device that is now in production. This has been made possible by a refined analysis of the system architecture and of the optimal signal processing techniques, but also by fundamental progresses of the various guided-wave technologies driven by the very important development of fiber-optic communications : single-mode and polarization-maintaining optical fibers, semi-conductor diode sources, integrated optics, in-line fiber components and more recently rare-earth-doped amplifying fibers.

Today, it is now accepted among inertial guidance and navigation specialists that I-FOG is a strong contender for many civilian or military applications. It brings the advantages of solid-state technology (guided-wave optics and low-voltage low-power electronics) with a potential of cost reduction which will enlarge the domain of application of inertial guidance techniques.

2. BASIC PRINCIPLE OF THE I-FOG

The I-FOG is based on Sagnac effect [2] which produces, in a ring interferometer, a phase difference ϕ_s proportional to the dot product of the rotation rate vector $\vec{\Omega}$ by the area vector $\vec{A}$ enclosed by the closed optical path :

$$\phi_s = \frac{4\,\omega}{c^2}\vec{A}\,.\,\vec{\Omega}$$

where ω is the angular frequency of the light source and c is light velocity in a vacuum. To understand simply Sagnac effect, it is possible to consider the case of an "ideal" circular path in a vacuum. Light entering the interferometer is divided into two counter-propagating waves which return perfectly in phase after having traveled along the same path in opposite directions [Figure 1-a]. Now, when the interferometer is rotating, an observer at rest in the inertial frame of reference sees the light traveling with the same vacuum velocity c in opposite directions, but

during the transit time through the loop, the beamsplitter has moved to M' [Figure 1-b]. Therefore our observer sees that the co-rotating wave has to propagate over more than one complete turn (its path increases of Δl_v) while the counter-rotating wave has to propagate over less than one complete turn (its path decreases of Δl_v). This difference of transit time may be measured by interferometric means.

Sagnac effect in matter is more delicate to explain, but it may be shown that it is completely independent of the indices of refraction or of the guidance conditions, and that is keeps the same value as in a vacuum [3,4]. With the I-FOG, the sensitivity is enhanced by the use of a multi-turn fiber coil which multiplies the enclosed area A by the number of turns N [Figure 2], just as the flux of a magnetic field is enhanced in a multi-turn inductance coil. Sagnac effect is a pure temporal delay independent of matter which is measured with a "clock", the frequency of the light source, but since the wavelength λ in a vacuum is a more familiar quantity than the frequency for an optical wave, the effect is usually written :

$$\phi_s = \frac{2\pi\ LD}{\lambda\ c}\ \Omega_{//}$$

where D is the mean coil diameter, L is the total coil length equal to $2\Pi.N.D$ and $\Omega_{//}$ the rotation rate component parallel to the coil axis.

The basic response of the I-FOG is a raised cosine as for any two-wave interferometer [Figure 3] with intensity I given by :

$$I = I_1\ [1 + \cos\ \phi_s]$$

where I_1 is the mean value of the intensity. The highest sensitivity is obtained around the $\Pi/2$ bias points and there is an unambiguous range of measurement of $\pm\pi$ radian around zero for the phase difference which corresponds to an unambiguous operating range of $\pm\Omega_\pi$ for the rotation rate. The sensitivity $\phi_s / \Omega_{//}$ has the dimension of time and is often expressed in seconds, and the

rate Ω_μ which corresponds to a phase difference of a micro-radian is a good order of magnitude for measurement noise limit, and drift which usually correspond to 0.1 to 10 μrad.

It is possible to scale up or down the sensitivity by changing the area of the sensing coil. This geometrical flexibility is an important advantage of I-FOG technology since the same components and assembly techniques may be used for various devices without a complete redesign. Two typical examples may be given :

	L, D, λ	ϕ_s / $\Omega_{//}$	Ω_π	Ω_μ
High sensitivity	1km, 10 cm, 1550 nm	1.35 s	133°/s	0.15°/h
Medium sensitivity	200 m, 3 cm, 850 nm	0.15 s	1220°/s	1.4°/h

3. RECIPROCITY AND SINGLE-MODE GUIDED PROPAGATION

As in any passive optical system, the theoretical sensitivity of the I-FOG is limited by photon shot noise, which states that the counting of **N** uncorrelated discrete particles like photons is a random process with a standard deviation of $\sqrt{N}$. Typically, with a returning power in the 1 to 100 μW range, there is a flow of 10^{12} to 10^{14} photons per second, i.e. over one second these 10^{12} to 10^{14} photons are measured with a respective standard deviation of 10^6 to 10^7 photons, which yields a white relative noise of 10^{-6} to 10^{-7} / $\sqrt{Hz}$ corresponding to 10^{-6} to 10^{-7} radian / $\sqrt{Hz}$ when the interferometer is operated about $\pi/2$ bias at the inflexion point of the raised cosine response where the slope is 1. Through time averaging, a phase difference of 10^{-7} to 10^{-8} radian should be measurable while the absolute phase accumulated by the wave along 100 m to 1 km is 10^9 to 10^{10} radians. This sensitivity limit is 16 to 18 orders of magnitude below the total propagation path when the temperature dependence is already on the order of 10^{-5}/K. Such a performance could look unrealistic but the fundamental principle of reciprocity of light propagation in a linear medium is the key to the solution of this problem.

Reciprocity states that for any optical wave in a linear medium, there is an opposite wave which propagates with exactly the same accumulated phase and the same attenuation. In free space, difficult alignments are required to excite both reciprocal waves and it is never perfect. However with single-mode (i.e. "single-solution") guided propagation, alignments are only needed to optimize the throughput power, but once they are coupled, both opposite waves are automatically reciprocal. The single-mode waveguide filters the exact part of both opposite guided waves corresponding to the same unique propagating solution. The nonreciprocal Sagnac effect is a very small first order effect (in $D\Omega/c$) which should be buried in the changes of the zero order, the absolute phase accumulated in the propagation, but single-mode guided propagation and reciprocity provide perfect "common-mode rejection" of this absolute phase.

The first experimental demonstration of a fiber gyro showed however that a simple fiber ring interferometer is not intrinsically reciprocal. Complementary interferences fringes were observed at both ports depending on the alignments of the fiber ends [1]. This problem could have been a severe limitation to high performance, but it was shown [5] that it can be solved very simply with a so-called reciprocal (or minimum) configuration [Figure 4] where light is fed in the interferometer through a single-mode waveguide filter and where the returning signal wave is filtered through this same waveguide filter at the output. This ensures that both returning waves have experienced exactly the same total phase despite potential defects in the splitting. A reciprocal operation of the I-FOG does not require a continuous single-mode propagation but merely a single-mode filtering at the common input-output port called the reciprocal port. This single-mode filtering comprises a single-spatial-mode filtering but also a polarization filtering because the so-called single-mode fibers are actually dual-polarization-mode fibers. The use of polarization-maintaining (or holding or preserving) fiber [6] improves the performance of the device and avoids signal fading but is not sufficient to avoid the use of a polarizer. It is also possible to use a depolarized approach [7] with a Lyot depolarizer and an ordinary fiber coil. In addition of the so-called coil splitter of the interferometer this reciprocal configuration requires a second splitter (so-called source splitter), to tap off part of the returning wave and to send it to a detector. Note that this source splitter might be replaced by a Faraday circulator [8] which avoids

the 6 dB-slitting-loss of the coupler, or even suppressed since the source may be used as an optical amplifier for the returning wave and the detector placed at its rear end [9].

To avoid the difficulties of the coupling stability of free-space waves into single-mode fiber it is desirable to use an all-guided scheme which improves ruggedness. This requires to duplicate in a guided form the various functions required in the interferometer. In particular the splitters may be replaced by evanescent-field couplers [10] or by Y junctions [11] using in-line fiber components or an integrated-optic circuit. Even with these guided-wave components, the reciprocal configuration must be used to ensure the best possible performances. An all-fiber approach using fiber couplers may yield very good performances [10] but today an hybrid solution [12-13] using a multifunction integrated-optic circuit is preferred. Integrated optics on an electro-optic substrate like lithium niobate ($LiNbO_3$) provides efficient low-voltage wide-band phase modulators which allow one to implement powerful signal processing techniques. This so-called Y-coupler configuration [Figure 5-a] is composed of a multifunction integrated-optic circuit [Figure 5-b] comprising a Y junction as the coil splitter, phase modulators on the junction branches, and a polarizer, the source splitter being an in-line fiber coupler. The polarizer may be fabricated with an absorbing metallic overlay, but $LiNbO_3$ allows one to fabricate true single-polarization waveguides with proton exchange (hydrogen ion, i.e. proton, replacing lithium ion in the crystal) [14] which provides simply a very high extinction polarizer.

4. BROADBAND SOURCE

When a broad-spectrum source of light is used in an interferometer, it is observed that there is good interference contrast about a null path difference, and as this difference increases, the fringe visibility decreases and finally vanishes completely [Figure 6]. When the path difference is longer than a given correlation length, called coherence length of the source in optics, both interfering waves are not correlated anymore which averages out the interference. This coherence length is inversely proportionnal to the spectrum width, and another way to understand the phenomenon is to decompose the broad spectrum into separate wavelengths. Each wavelength λ_i creates its own

interference pattern with a period λ_i [Figure 7]. About zero path difference, all the interference patterns coincide ; but as the path difference increases, they loose their coincidence because they have slightly different periods. On average, the total intensity becomes constant and the contrast disappears. Since an I-FOG operates about a null path difference, it may use a broadband source without degradation of the signal.

Now, in addition of both primary reciprocal waves which interfere about this null path difference, there are many parasitic waves which could yield unstable spurious interferences. They are due to backreflections [15], backscattering [16] or unperfectly filtered cross-polarized waves [17]. With a coherent source, their residual contrasts are proportional to the amplitude ratio between these parasitic waves and the primary waves, i.e. the square-root of the intensity ratio. To limit them below 10^{-6} - 10^{-7} and to ensure that the equivalent phase difference error in the main interferometer is less than 10^{-6} - 10^{-7} radian, it would require a ratio as small as -120 to -140 dB, which could have been a drastic limitation to high performances. However the use of a broadband source destroys the contrast of these spurious interferences related to long path differences while preserving the main signal due to interferences about null path difference. This is very useful for reducing backscattering noise [18] and also for solving the difficult problem of polarization non-reciprocities [12,19]. The analysis of these problems of polarization is very important in a gyroscope. The residual bias drift is due to a complex combination of polarizer rejection, crossed-polarization coupling and depolarization induced by the propagation of broad-spectrum light along birefringent waveguides. Note that optical coherence domain polarimetry based on white-light interferometry [20] is a very efficient tool to perform the analysis of these problems.

The low temporal coherence of broadband sources makes them very suitable for the I-FOG, but they have to be coupled efficiently in a single-mode fiber which requires a high spatial coherence. Both criteria (low temporal coherence and high spatial coherence) are not antagonistic, but complementary development was necessary to optimize adequate sources. These sources are based on the superluminescence (or superradiance or superfluorescence) of high-gain single-mode-

waveguide amplifiers : by spontaneous emission photons are emitted over a broad spectrum and some of them, captured within the guidance aperture of the waveguide, are then amplified by stimulated emission. Because of the high amplification gain, large power is obtained over a single pass which yields a strong single-spatial-mode emission preserving the broad spectrum of the seminal spontaneous emission. These sources were originally based on semi-conductor diode lasers where the resonant Fabry-Perot cavity is suppressed by anti-reflection coating, absorber or slant edges. Such sources called superluminescent diodes (SLD) work in the 850 nm window with GaAlAs diodes or the 1300 nm or 1550 nm windows with InP diodes. They bring a high reduction of bias noise but they suffer from a poor spectrum stability (400 ppm/°C) which yields scale factor drift since as we saw the sensitivity ($\phi_s/\Omega_{//}$) is inversely proportional to the wavelength.

For high-performance applications, the trend is now to use fiber sources in the 1550 nm window which are derived from erbium-doped-fiber amplifier (EDFA) technology [21,22,23] that is revolutionizing optical fiber communication. They are also based on superluminescence with a doped fiber excited by a pump laser diode [Figure 8]. This superluminescence is also called ASE for Amplified (by stimulated emission) Spontaneous Emission. Energy states being more stable with rare-earths than with semi-conductor diodes, it was expected a better stability of the spectrum but this problem is delicate since erbium ASE is composed of a 1530 nm peak and a 1550-1560 nm shoulder [Figure 9] which must be controlled to ensure a good behaviour. Note that it is possible to stabilize simply the spectrum by inserting a stable band-limited reflector (like in-fiber Bragg grating) in the source [24].

Er-doped-fiber source brings in addition a high power (more than 10 mW) which improves the signal over noise ratio and an unpolarized emission which is useful to reduce polarization non-reciprocities [25], and allows one to use ordinary (non-polarization preserving) coupler for the source splitter. However its spectrum is not very broad and one has to deal with excess random intensity noise (excess RIN) which has an rms value inversely proportional to the square-root of the spectrum width [26]. This RIN can be compensated for [27] or its actual effect reduced by operating the FOG near π radian instead of about $\pi/2$ [28].

Note that these Er-sources might have been considered as a luxurious solution, but today with the very broad development of EDFA technology for communications, they have also an economic advantage in particular with multi-axis operation since their high power may be shared with very low penalty on noise.

5. SIGNAL PROCESSING TECHNIQUE

The reciprocal configuration provides an interference signal I which is a raised cosine maximum at zero [Figure 3]. To get high sensitivity this signal must be biased about an operating point with a non-zero response slope. This is efficiently obtained by the use of a reciprocal phase modulator (or a push-pull pair) placed at the end of the coil that acts as a delay line [29] [Figure 10]. A phase modulation $\phi_m(t)$ being applied with an electrical voltage, this yields a modulation of the phase difference :

$$\Delta\phi_m(t) = \phi_m(t) - \phi_m(t-\Delta\tau_g)$$

where $\Delta\tau_g$ is the difference of group transit time between the long and short paths that connect the modulator to the splitter.

This technique may be implemented with a square-wave modulation where the half-period is equal to $\Delta\tau_g$ [29]. This corresponds to the so-called proper (or eigen) frequency $f_p = 1 / (2\Delta\tau_g)$ of the coil (the product coil length x proper frequency being about 100 m x MHz with silica fiber). At rest the biasing modulation is centered but in rotation the signal becomes unbalanced [Figure 11] which yields a modulated output which can be demodulated to give a signal S which is a biased sine response with a very stable zero since the optical intensity and modulation amplitude dependence are multiplicative [Figure 12].

The use of a reciprocal modulator is fundamental since the mean value of a periodic modulation of the phase difference is perfectly centered around zero without any spurious offset susceptible of drift :

$$< \Delta\phi_m(t) > = < \phi_m(t) - \phi_m(t-\Delta\tau_g) > = < \phi_m > - < \phi_m > = 0$$

It is also very important to operate at the proper frequency to avoid parasitic errors due to defects of the modulation. With square wave, defects arise from the residual asymmetry of the duty cycle [20] and with sine wave, which can be alternatively used [5], defects arise because of non-linearities inducing harmonic components in the modulation as well as because of residual intensity modulation [30].

This modulation-demodulation scheme provides a very stable bias since it preserves the reciprocity of the ring interferometer. However, a high-performance gyro requires also good accuracy over the whole dynamic range and not only about zero. The measurement of interest is the integrated angle of rotation and any past error will affect the future information. This implies the need for an accurate measurement at any rate, i.e. an accurate scale factor. Furthermore the intrinsic response S of the interferometer is sinusoidal while the desired rate signal should be linear.

This problem is solved with a closed-loop signal processing approach [31,32]. The demodulated biased signals (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 maintained opposite to the Sagnac phase ϕ_s [Figure 13]. With such a closed-loop scheme, the new measurement is the value of $\Delta\phi_{FB}$. This yields a linear response with good stability since this feedback is independent of the returning optical power and of the gain of the detection chain.

The first proposed closed-loop arrangements used acouto-optic frequency shifters [30,31] but this destroys the full reciprocity of the interferometer. This is now overcome by the use of analog

phase ramp ϕ_{PR} which generates a feedback phase difference $\Delta\phi_{FB} = \dot{\phi}.\Delta\tau_G$ where $\dot{\phi}$ is the time derivative (i.e. the slope) of the phase ramp and $\Delta\tau_G$ is the difference of group transit time through the coil [figure 14] [33,34]. Note that the reset ϕ_{RS} has to be controlled to 2π radian to avoid defects. Important improvements are obtained with a digital version generating phase steps ϕ_S equal to $\Delta\phi_{FB}$ and which allows one to synchronise naturally the ramp resets which must be equal to 2π radian with the biasing modulation [Figure 15], the response of the phase modulator being controlled with a second feeback loop checking the error signal at these resets [12,13,33].

Associated with a digital demodulation, this yields an all-digital closed-loop processing [35] method which is intrinsically free of electronic source of bias drift. Such an electronics working at a frequency of 0.1 to 1 MHz for a coil length from 1000 to 100 meters, is composed of a detector, filtering means, direct A/D conversion of the modulated signal, a digital processing circuit (demodulation, integration filter for closing the loop, generation of feedback ramp and biasing modulation), D/A conversion of the modulator driving signal and a buffer amplifier. One has to be very careful of ground loops and electromagnetic interferences between the phase modulators which are driven with few volts and the detector where the sensitivity limit corresponds to a primary current smaller than 10^{-12} ampere which will be demodulated at the same frequency. There is in addition a second D/A converter for the additional control loop of the gain of the modulation chain [Figure 16]. This second loop controls that the ramp resets are equal to 2π radian (the period of the interference response) or may alternatively work with a so-called "4-state" modulation [36]. Figure 17 summarizes the optimal architecture of the I-FOG with a Y-coupler optical configuration, a broadband source and an all-digital closed-loop processing electronics.

Note finally that once a perfect phase measurement is performed with a closed-loop scheme, the scale factor remains proportional to the equivalent area of the coil and to the mean wavelength of the detected signal. The coil area exhibits a typical variation of several ppm per kelvin because of the thermal expansion of the polymer coating of the fiber, which necessitates some modeling for high-grade applications. Wavelength stability is also difficult to solve and requires a control

system to get the best possible performances.

6. OTHER PARASITIC EFFECTS

Because of reciprocity, the two single-mode counter-propagating paths are equalized but this is valid only if the system is time-invariant. We have already seen that biasing phase modulation can be generated with a reciprocal modulator placed at the end of the coil. Parasitic phase shifts generated by non-uniform temperature changes yields a similar effect (so-called Shupe effect) [37]. However, if the same temperature change is experienced by fiber segments at equal distance from the middle of the coil, the effect is canceled out. This is obtained with symmetrical winding where the fiber is wound from the middle and by alternating layers coming from each half-coil length with a so-called "dipolar" winding [37]. An even better compensation is obtained with a so-called "quadrupolar" winding [38] [Figure 18].

Acoustic noise and vibrations are also a potential source of parasitic effects related to transient non-reciprocity. Symmetrical winding and adequate holding of the coil are necessary to minimize the effect. In principle, these AC perturbations should not affect the mean DC bias but one has to be careful to avoid rectification in the processing electronics.

In addition to these transient non-reciprocities there are two additional sources of true non-reciprocity : the magneto-optic Faraday effect [39,40] and the non-linear Kerr effect [41]. Faraday non reciprocity is induced with polarization-maintaining fiber but it remains in the typical range of 1 microradian per gauss (earth magnetic field being typically 0.5 gauss). This residual dependence is due to the random rotation of the birefringence axes of practical polarization-maintaining fibers which yields a state of polarization with a small random ellipticity along the coil [42]. For a known fiber the Faraday phase error should increases as the square root of its length [42,43]. If the application requires a very low magnetic dependence, it is useful to shield the coil with μ-metal.

The second truly non-reciprocal effect is Kerr non-linearity [42], since reciprocity is indeed based on the linearity of propagation equation. An imbalance in the power levels of the counterpropagating waves can produce a small phase difference because of non-linearity of the index of refraction induced by the high power density in the very small fiber core. Slow variations in the splitting ratio in the sensing coil may therefore translate directly into bias drift. However, this Kerr-induced error results from a complex four-wave mixing process between the opposite waves and not simply from an intensity self-dependence. It was shown that a broadband source equalizes the Kerr non-linearity for both opposite waves, and suppresses completely the phase-difference error [44,45].

7. MULTI-AXIS ARCHITECTURES

To reduce the size and the cost of FOG-based inertial measurement unit (IMU), there is a trend to use multi-axis fiber gyro where a single source is shared [11] [Figure 19], in particular by a triad of ring interferometers measuring the three orthogonal rotation-rate components [46,47]. Note that an Er-doped fiber source is particularly useful in this case because its high power may be divided with low penalty on noise and it is possible to use a cluster of ordinary couplers since it is unpolarized.

A further improvement is possible with multiplexing. The source is shared between several interferometers composed of a sensing coil and a multi-function integrated-optic circuit, and the signals return on a single detector [Figure 20]. Multiplexing is performed with the phase modulation signal applied on the various interferometers. A first proposed solution is to apply, in a sequence, the biasing modulation on one interferometer while there is no modulation on the other ones [48]. It is also possible to modulate the other ones between $\pm\pi$ radian which nulls out their sensitivity [49]. This improves the signal over noise ratio because the unused axis are "turned off" which reduce the detected power. However, with these two approaches the detected power may vary significantly during the switching between two axis if their respective returning powers are not perfectly balanced. These power changes may yield defects in the high gain detection chain. To avoid

this problem, a multiplexed time-shifted phase biasing modulation has been proposed [50].

8. CONCLUSION

This chapter has reviewed the fundamentals of the I-FOG which is based on enhanced Sagnac effect in a multi-turn ring fiber interferometer. Its optimal conditions of operation may be summarized as follows :

- Use of a single-mode reciprocal configuration to ensure perfect common mode rejection of environmental perturbations,
- Use of a modulation-demodulation biasing scheme with phase modulators at the ends of the coil,
- Use of a broadband source to suppress coherence-related sources of noise and drift,
- Use of a closed-loop processing scheme to linearize and stabilize the scale factor. Among the possible techniques, the all-digital closed-loop approach with phase ramp feedback provides fundamental advantages,
- Use of wavelength control,
- The optimal practical configuration uses a muti-function integrated optic circuit and a polarization-maintaining sensing coil.

After more than twenty years of research and development, the I-FOG is now recognized as a crucial technology for inertial guidance and navigation. It covers medium accuracy range (0.1 to 10°/h) with compact sensing coils and it becomes also a significant competitor for high-performance navigation grade (0.01°/h) with larger coils. Using mass-producible components, it brings great expectation of significant cost reduction which will further extend the field of use of inertial guidance techniques.

REFERENCES

1) Vali V. and Shorthill R.W., *"Fiber ring interferometer"*, Applied Optics, Vol. 15, 1099-1100, (1976).

2) Sagnac G., *"L'éther lumineux démontré par l'effet du vent relatif d'éther dans un interféromètre en rotation uniforme"*, Comptes rendus de l'Académie des Sciences, Vol. 95, 708-710, (1913).

3) Post E.J., *""Interferometric path-length changes due to motion"*, J.O.S.A., Vol. 62, 234-239, (1972).

4) Arditty H.J. and Lefèvre H.C., *"Sagnac effect in fiber gyroscopes"*, Optics Letters, Vol. 6, 401-403, (1981).

5) Ulrich R., *"Fiber optic rotation sensing with low drift"*, Optics Letters, Vol. 5, 173-175, (1980).

6) Burns W.K., Moeller R.P., Villaruel C.A. and Abebe M., *"Fiber-optic gyroscope with polarization-holding fiber"*, Optics Letters, Vol. 8, 540-542, (1983).

7) Böhm K., Marten P., Petermann K., Weidel E. and Ulrich R., *"Low-drift fibre gyro using a superluminescent diode"*, Electronics Letters, Vol. 17, 352-353, (1981).

8) Lefèvre H., *"Contribution à la géométrie à fibre optique de classe inertielle"*, Thèse de docteur ès-Sciences, University of Nice, (1982).

9) Arditty H. and Lefèvre H., *"Monomode optical fiber ring interferometer device with semi-conductor diode as light energy emission reception-amplification means"*, US patent 4,842,409, (1989).

10) Lefèvre H.C., Bergh R.A. and Shaw H.J., *"All-fiber gyroscope with inertial-navigation short-term sensitivity"*, Optics Letters, Vol. 7, 454-456, (1982).

11) Arditty H.J., Papuchon M., and Puech C., *"Reciprocity properties of a branching waveguide"*, Fiber-Optic Rotation Sensors, Springer Series in Optical Sciences, Vol. 32, 102-110, (1982).

12) Lefèvre H.C., Bettini J.P., Vatoux S. and Papuchon M., *"Progress in optical fiber gyroscopes using integrated optics"*, NATO/AGARD Conference Proceedings, Vol. 383, 9A-1, 9A-13, (1985).

13) Lefèvre H.C., Vatoux S., Papuchon M. and Puech C, *"Integrated optics : a practical solution for the fiber-optic gyroscope"*, SPIE Proceedings, Vol. 719, 101-112, (1986).

14) Suchoski P.G., Findakly T.K. and Leonberger F.L., *"$LiNbO_3$ integrated optical components for fiber optic gyroscopes"*, SPIE Proceedings, Vol.993, 240-244, (1988).

15) Arditty H.J., Shaw H.J., Chodorow M. and Kompfner R.K., *"Re-entrant fiber-optic approach to rotation sensing"*, SPIE Proceedings, Vol. 157, 138-148, (1978).

16) Cutler C.C., Newton S.A. and Shaw H.J., *"Limitation of rotation sensing in scattering"*, Optics Letters, Vol. 5, 488-490, (1980).

17) Kintner E.C., *"Polarization control in optical-fiber gyroscopes"*, Optics Letters, Vol. 6, 154-156, (1981).

18) Böhm K., Russer P., Weidel E. and Ulrich R., *"Low-noise fiber-optic rotation sensing"*, Optics Letters, Vol. 6, 64-66, (1981).

19) Jones E. and Parker J.W., *"Bias reduction by polarisation dispersion in the fibre-optic gyroscope"*, Electronics Letters, Vol. 22, 54-56, (1986).

20) Lefèvre H.C., *"Comments about fiber-optic gyroscopes"*, SPIE Proceedings, Vol. 838, 86-97, (1987).

21) Digonnet M.J.F., *"Status of broadband rare-earth doped fiber sources for FOG applications"*, SPIE Proceedings, Vol. 2070, 113-131, (1993).

22) Iwatsuki K., *"Long-term bias stability of all-panda fibre gyroscope with Er-doped superfluorescent fibre laser"*, Electronics Letters, Vol. 27, 1092-1093, (1991).

23) Hall D.C., Burns W.K. and Moeller R.P., *"High-stability Er^{3+}-doped superfluorescent fiber sources"*, Journal of Lightwave Technology, Vol. 13, 1452-1460, (1995).

24) Gaiffe T., Simonpiétri P., Morisse J., Cerre N., Taufflieb E.M. and Lefèvre H.C., *"Wavelength stabilization of an erbium-doped fiber source with a fiber Bragg grating for high-accuracy FOG"*, To be published in SPIE Proceedings, Vol. 2837, (1996).

25) Pavlath G.A. and Shaw H.J., *"Birefringence and polarization effects in fiber gyroscopes"*, Applied Optics, Vol. 21, 1752-1757, (1982).

26) Morkel P.R., Laming R.I. and Payne D.N., *"Noise characteristics of high-power doped-fiber superluminescent sources"*, Electronics Letters, Vol. 26, 96-98, (1990).

27) Killian K., Burmenko M. and Hollinger W., *"High performance fiber optic gyroscope with noise reduction"*, SPIE Proceedings, Vol. 2292, 255-263, (1994).

28) Lefèvre H. *"The fiber optic-gyroscope"*, Artech House, Boston-London, Sections 2.3.2 and 9.2.2, (1993).

29) Martin J.M. and Winkler J.T., *"Fiber-optic laser gyro signal detection and processing technique"*, SPIE Proceedings, Vol. 139, 98-102, (1978).

30) Bergh R.A., Lefèvre H.C. and Shaw H.J., *"All-single-mode fiber-optic gyroscope with long-term stability"*, Optics Letters, Vol. 6, 502-504, (1981).

31) Davis J.L. and Ezekiel S., *"Techniques for shot-noise-limited inertial rotation measurement using a multi-turn fiber Sagnac interferometer"*, SPIE Proceedings, Vol. 157, 131-136, (1978).

32) Cahill R.F. and Udd E., *"Phase-nulling fiber-optic laser gyro"*, Optics Letters, Vol. 4, 93-95, (1979).

33) Lefèvre H.C., Graindorge Ph., Arditty H.J., Vatoux S. and Papuchon M., *"Double closed-loop hybrid fiber gyroscope using digital phase ramp"*, Proceedings of OFS 3, OSA/IEEE, Post-deadline paper 7, (1985).

34) Kay C.J., *"Serrodyne modulator in a fibre-optic gyroscope"*, IEE Proceedings, Vol. 132, 259-264, (1985).

35) Lefèvre H.C., Martin P., Morisse J., Simonpiétri P., Vivenot P. and Arditty H.J., *"High dynamic range fiber gyro with all-digital processing"*, SPIE Proceedings, Vol. 1367, 72-80, (1990).

36) Lefèvre H.C., Martin Ph., Gaiffe T., Graindorge Ph., Le Boudec G., Morisse J., Simonpiétri P., Taufflieb E., Vivenot Ph. and Arditty H.J., *"Latest advances in fiber-optic gyroscope technology at Photonetics"*, SPIE Proceedings, Vol. 2292, 156-165, (1994).

37) Shupe D.M., *"Thermally induced nonreciprocity in the fiber-optic interferometer"*, Applied Optics, Vol. 9, 654-655, (1980).

38) Frigo N.J., *"Compensated of linear sources of non-reciprocity in Sagnac interferometers"*, SPIE Proceedings, Vol. 412, 268-271, (1983).

39) Böhm K., Petermann K. and Weidel E., *"Sensitivity of a fiber-gyroscope to environmental magnetic fields"*, Optics Letters, Vol. 7, 180-182, (1982).

40) Bergh R.A., Lefèvre H.C. and Shaw H.J., *"Geometrical fiber configuration for isolators and magnetometers"*, Springer Series in Optics Sciences, Vol. 32, 400-405, (1982).

41) Ezekiel S., Davis J.L. and Hellwarth R.W., *"Intensity dependent non-reciprocal phase shift in a fiber-optic gyroscope"*, Springer Series in Optical Sciences, Vol. 32, 332-336, (1982).

42) Hotate K. and Tabe K., *"Drift of an optical fiber gyroscope caused by the Faraday effect : influence of the earth's magnetic field"*, Applied Optics, Vol. 25, 1086-1092, (1986).

43) Lefèvre H., *"The fiber-optic gyroscope"*, Artech House, Boston-London, Section 7.1, (1993).

44) Bergh R.A., Culshaw B., Cutler C.C., Lefèvre H.C. and Shaw H.J., *"Source statistics and the Kerr effect in fiber-optic gyroscopes"*, Optics Letters, Vol. 7, 563-565, (1982).

45) Petermann K., *"Intensity-dependent nonreciprocal phase shift in fiber-optic gyroscopes for high sources with low coherence"*, Optics Letters, Vol. 7, 623-625, (1982).

46) Pavlath G.A., *"Production of fiber gyros at Litton Guidance and Control Systems"*, SPIE Proceedings, Vol. 1585, 2-6, (1991).

47) Büshelberger H.J. and Kemmler M., *"Closed-loop fiber-optic gyro triad"*, SPIE Proceedings, Vol. 1585, 42-47, (1991).

48) Auch W., *"Progress in fiber-optic gyro development and application"*, SPIE Proceedings, Vol. 2070, 104-112, (1993).

49) Goldner E.L., *"Triaxial fiber optic Sagnac interferometer with single source and detector"*, US Patent 5,184,195, (1993).

50) Gaiffe T., Morisse J., Simonpiétri P., Martin P. and Lefèvre H.C., *"Fog multiplexing by time-shifted phase modulation"*, to be published in SPIE Proceedings, Vol. 2837, (1996).

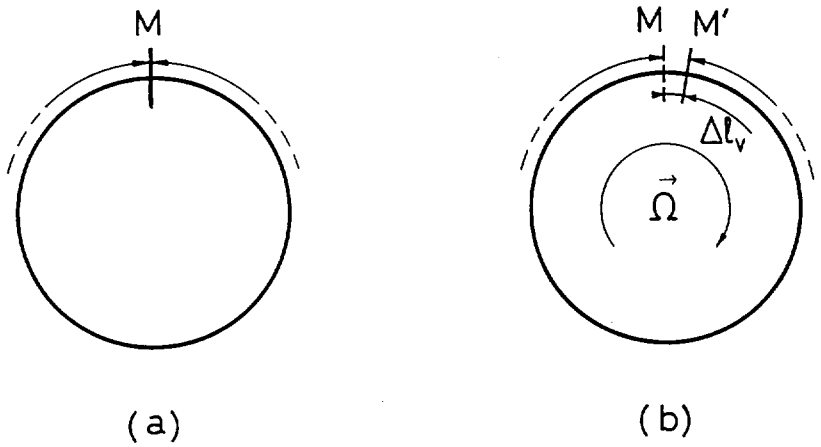


Figure 1

Sagnac effect in a vacuum considering an "ideal"circular path :

(a) interferometer at rest

(b) interferometer rotating

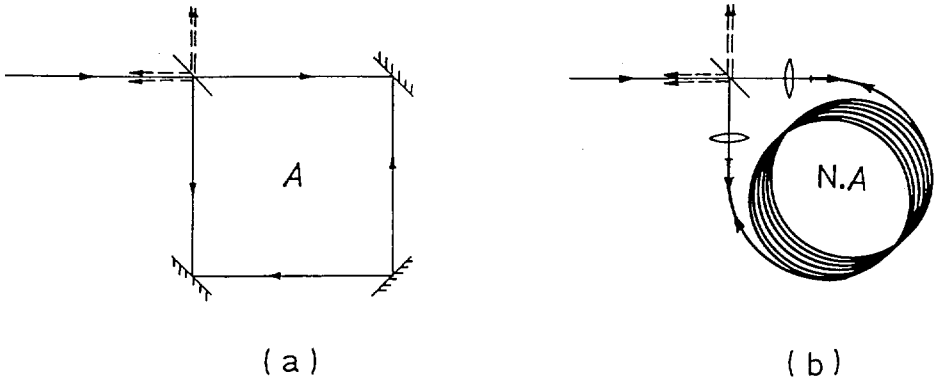


Figure 2

Ring interferometer :

(a) bulk form with an enclosed area A,

(b) enhanced sensitivity with a multiturn fiber coil that has an enclosed area N.A

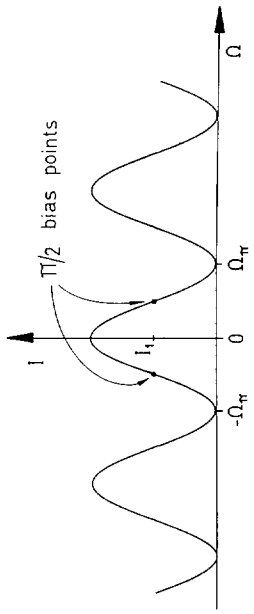


Figure 3

Basic response of an I-FOG

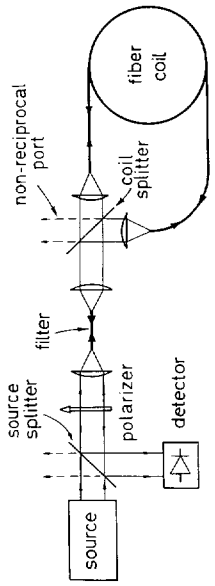
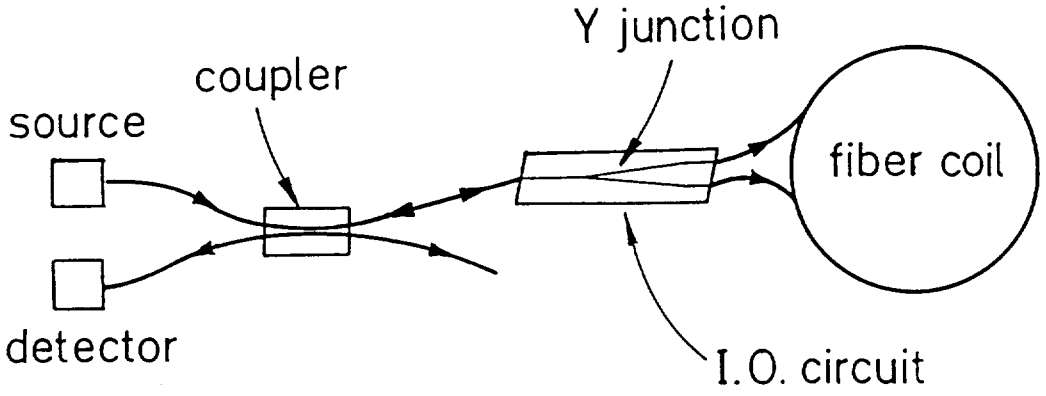
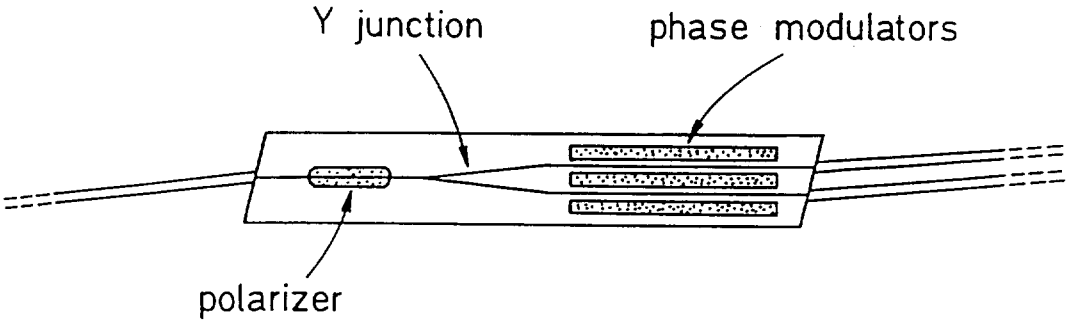


Figure 4

Reciprocal (or minimum) configuration of the I-FOG



(a)



(b)

Figure 5

Optimal hybrid configuration of the I-FOG :

a) Y-coupler configuration

b) multifunction integrated-optic (IO) circuit

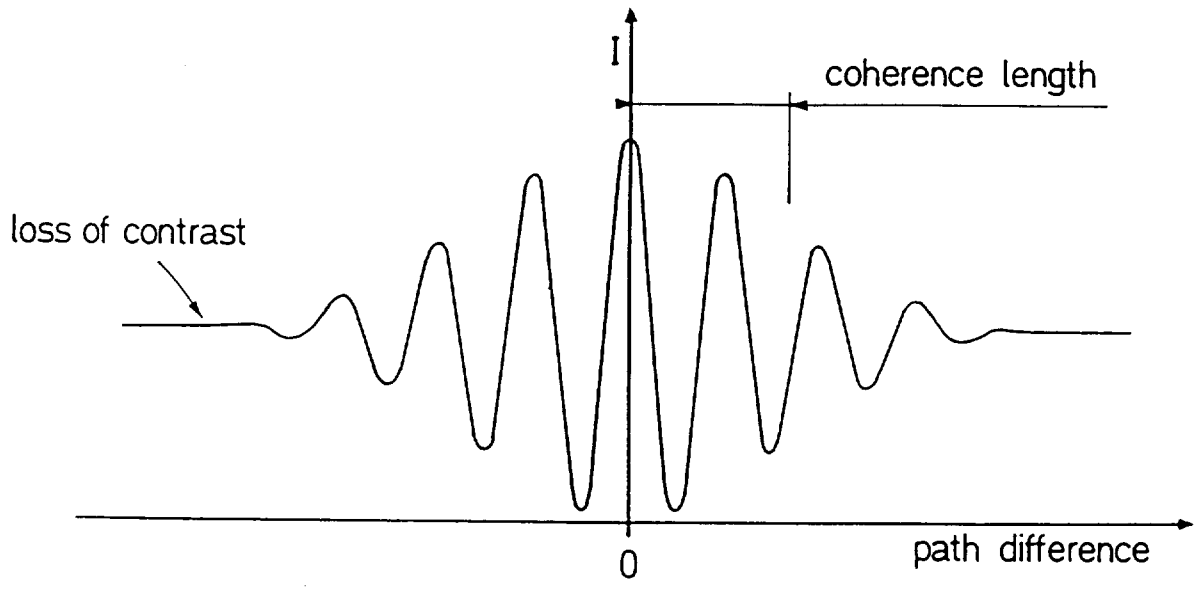


Figure 6

Interference signal with a broad-spectrum source

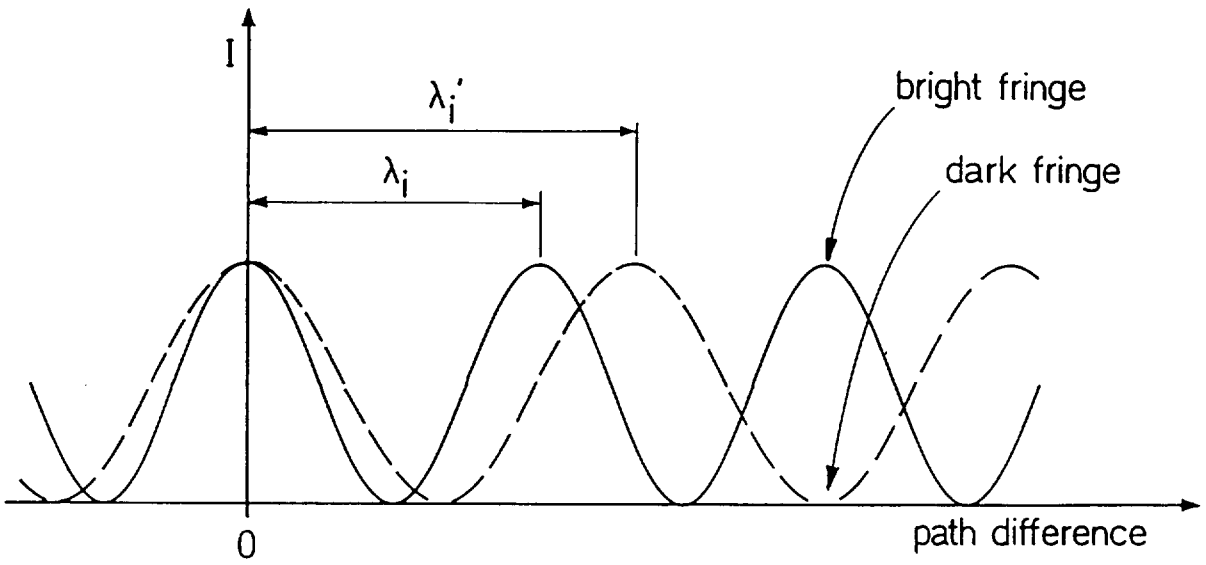


Figure 7

Interference signals with different wavelengths

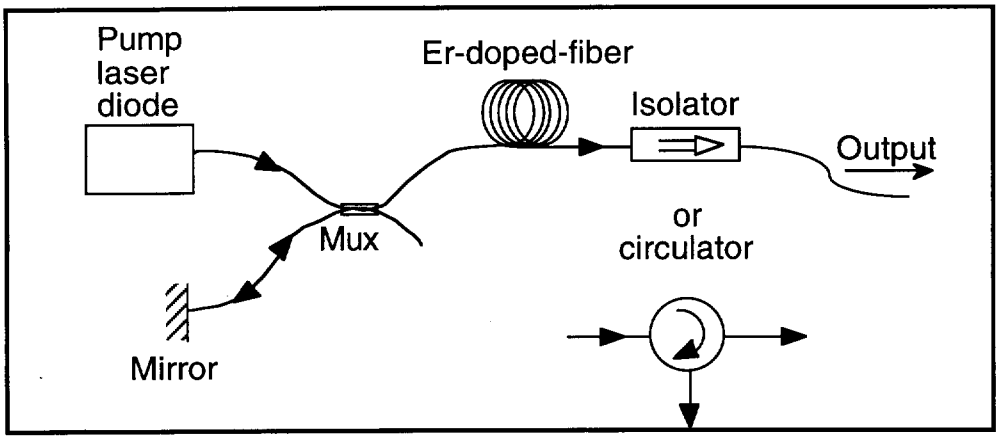


Figure 8

Configuration of a superluminescent (or ASE) Er-doped fiber source (for higher power a mirror is placed at one end and an isolator or a circulator is needed to avoid lasing)

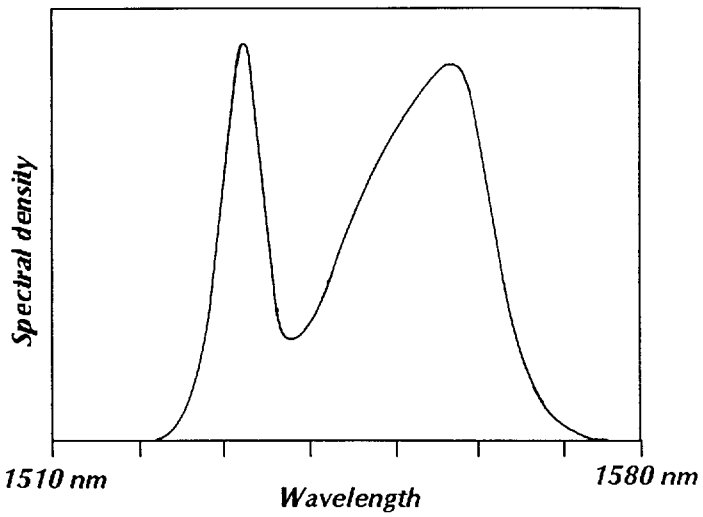


Figure 9

Typical ASE spectrum of an Er-doped fiber source with a 1530 nm peak and a 1550-1560 nm shoulder

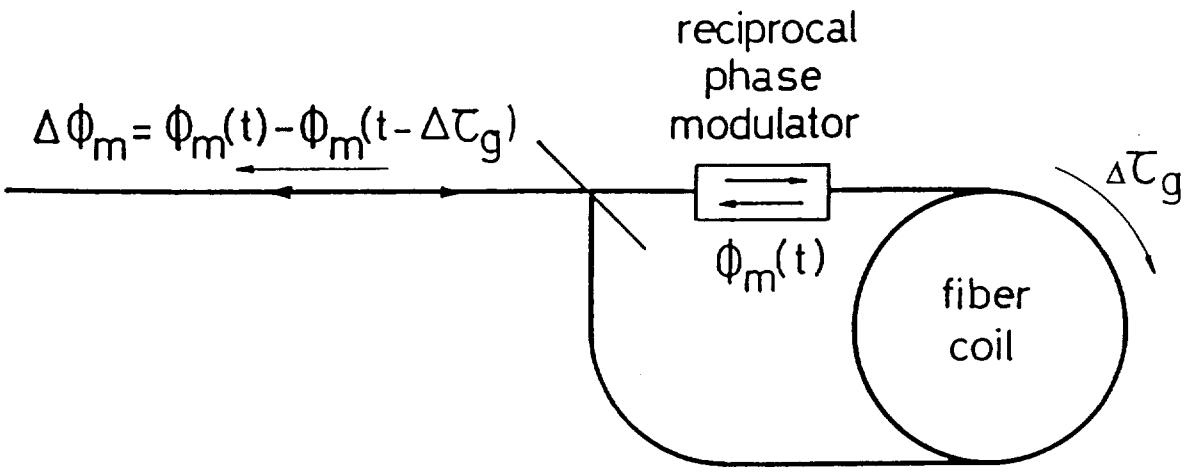


Figure 10

Generation of a modulation of the

phase difference using the delay through the coil

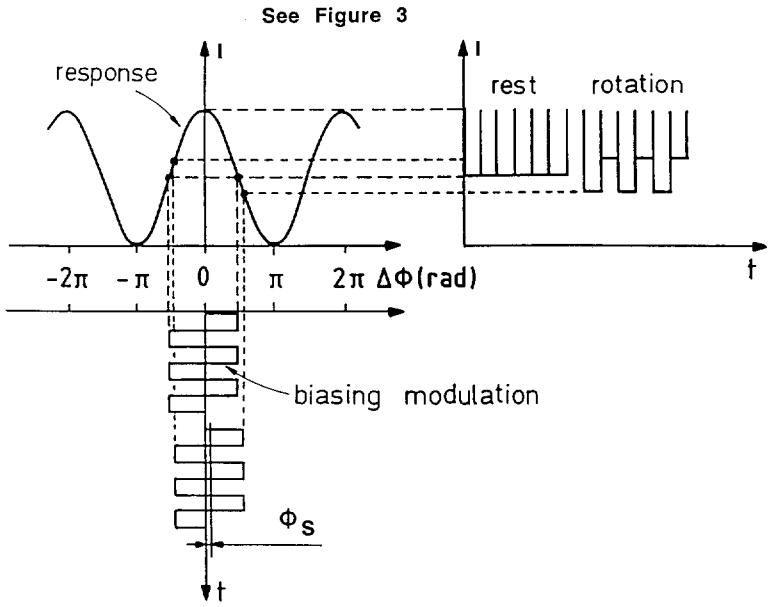


Figure 11

Square-wave biasing modulation

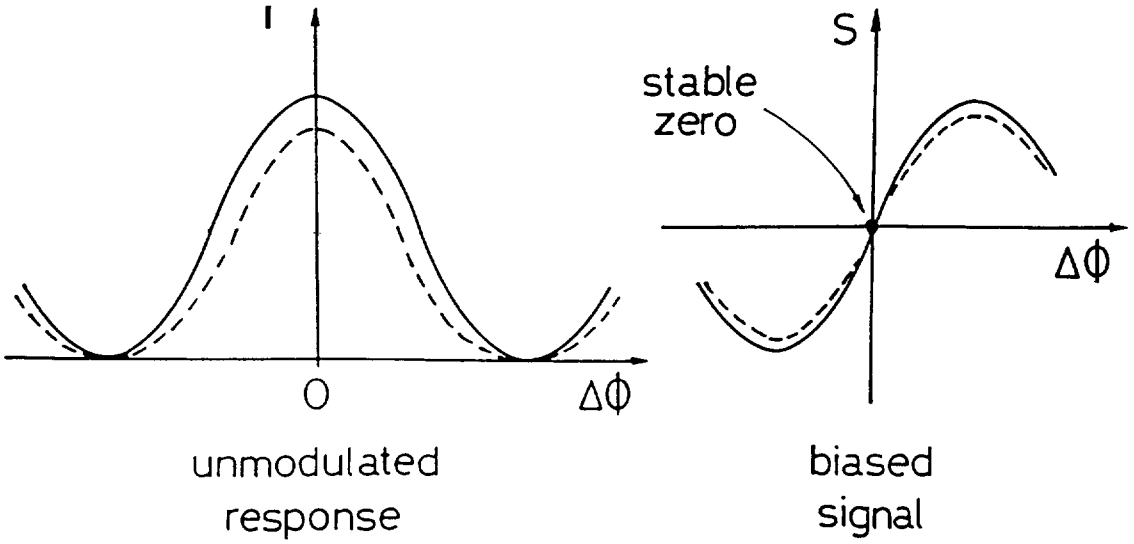


Figure 12

Unmodulated response and demodulated biased signal of an I-FOG

(dashed lines indicate a change of optical intensity)

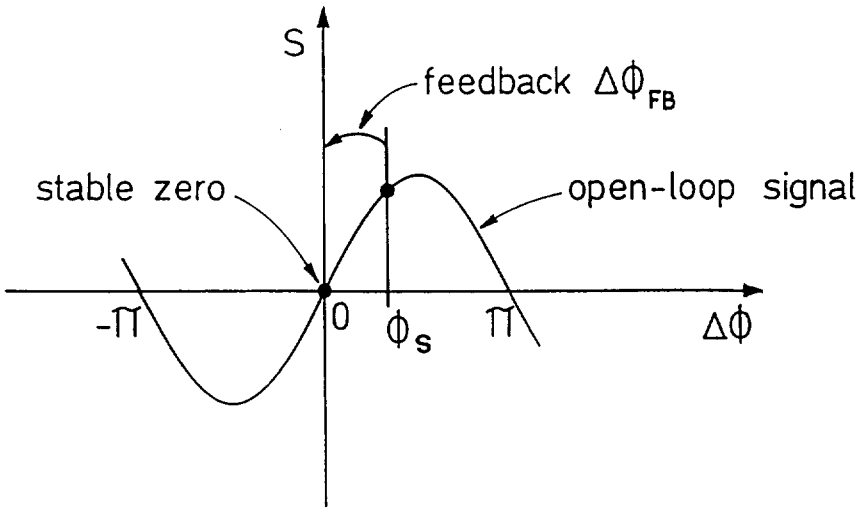


Figure 13

Principle of closed-loop operation of an I-FOG

with a feedback phase difference $\Delta\phi_{FB}$

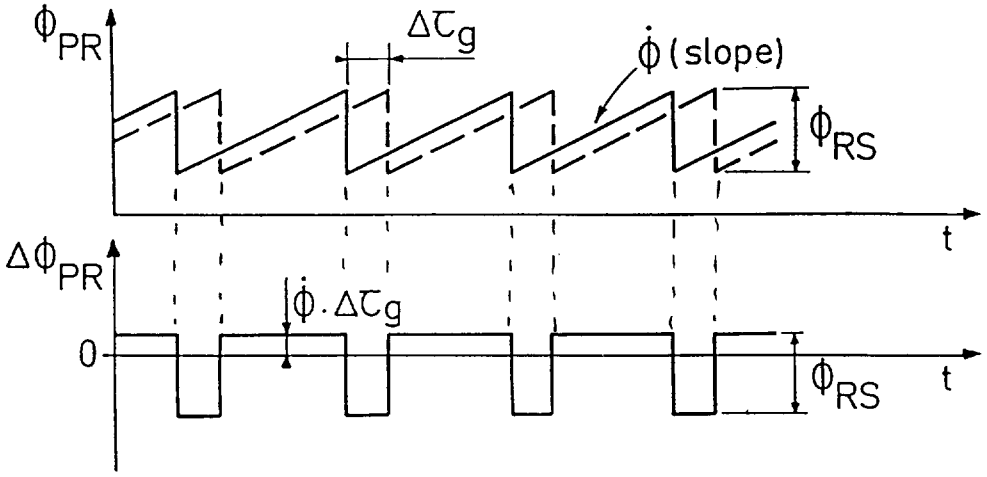


Figure 14

Analog phase ramp ϕ_{PR} and

induced feedback phase difference $\Delta\phi_{PR}$

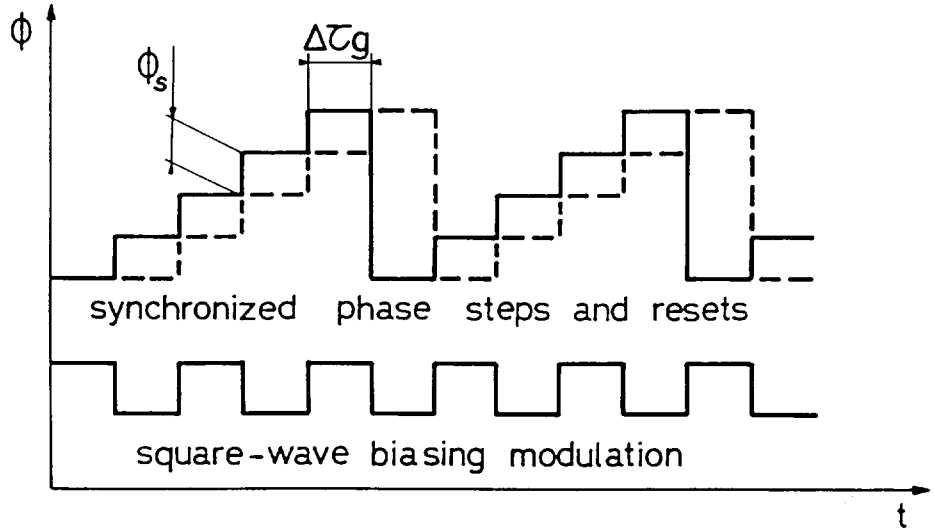


Figure 15

Digital phase ramp with steps and resets

synchronized with square-wave biasing modulation

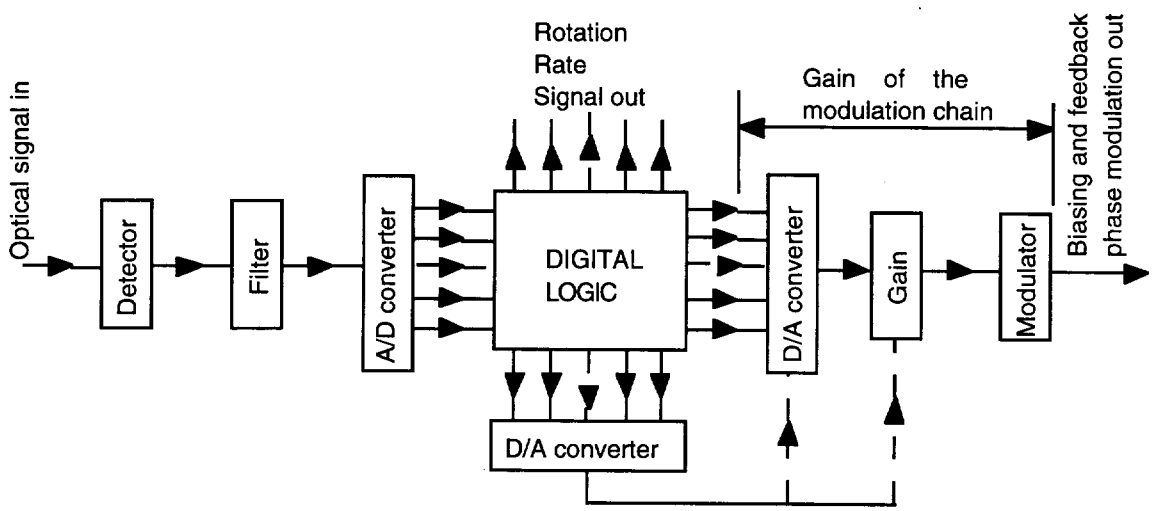


Figure 16

Functional schematics of all-digital closed-loop processing electronics

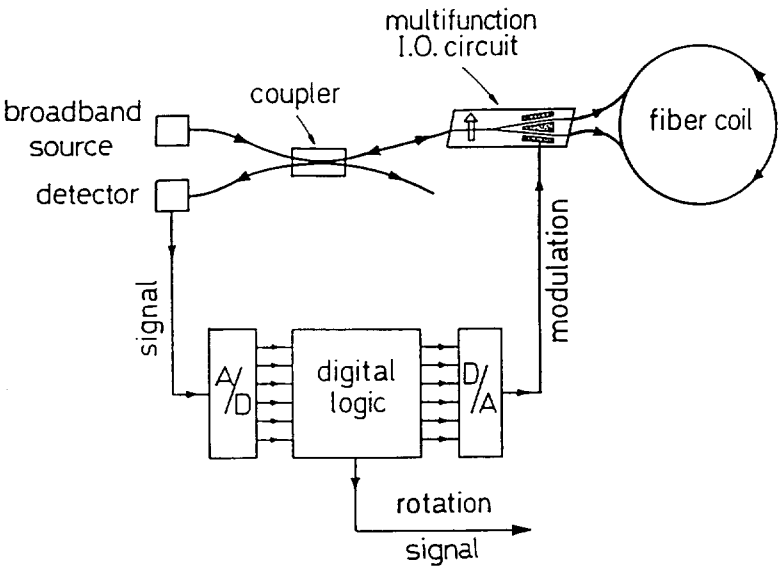


Figure 17

Optimal architecture of the I-FOG with Y-coupler optical configuration and all-digital closed-loop processing electronics

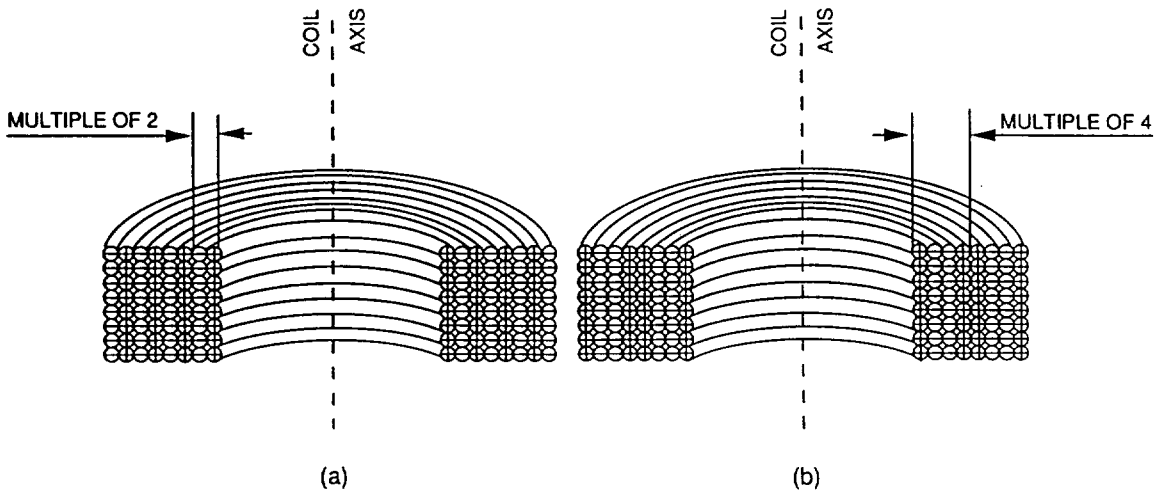


Figure 18

Section view of symmetrical windings

(one half-coil length is represented by + and the other by -) :

a) dipolar, b) quadrupolar

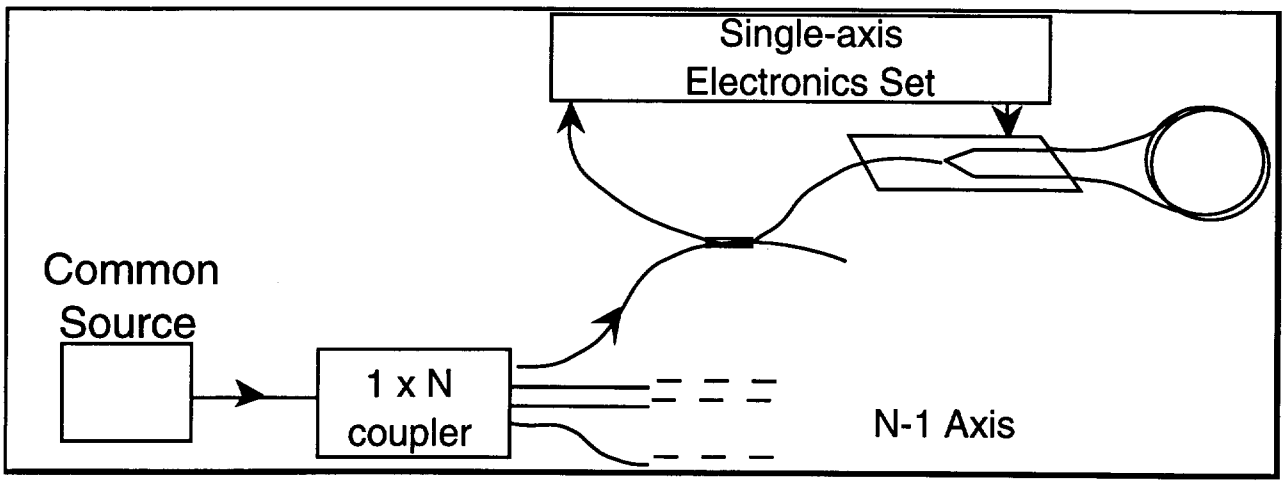


Figure 19

Source sharing in a multi-axis FOG

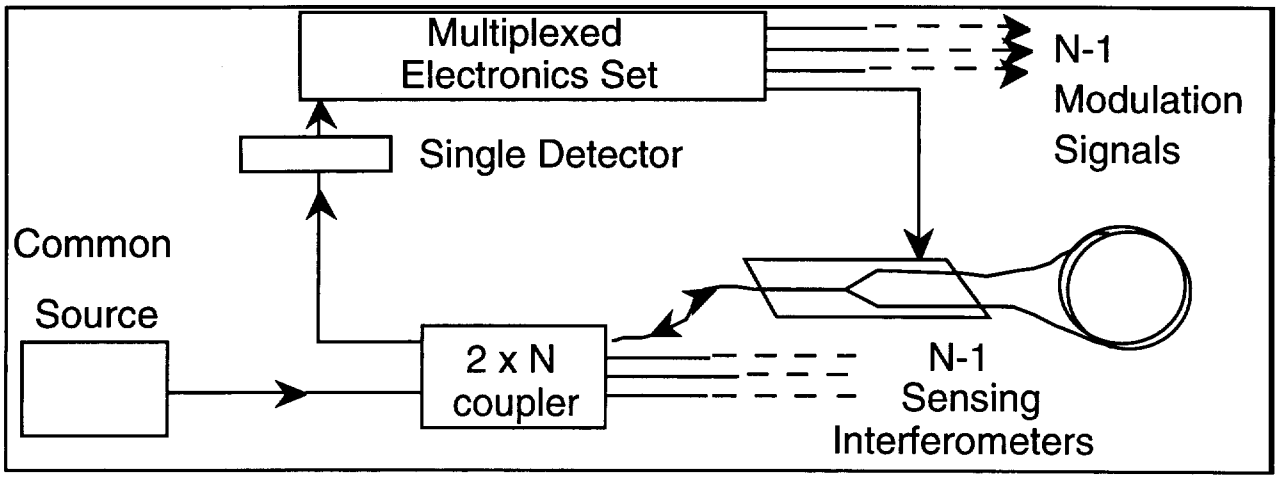


Figure 20

Multiplexing in a multi-axis FOG

8. Passive All-Fiber Open Loop Gyroscope

8.1. Introduction

For years research and development in the fiber optic gyro area have concentrated on better performances, touching the physical limits of the respective concepts. Fiber optic gyroscopes (FOG) with sophisticated signal recovery schemes were generated and started to compete with well-established mechanical gyros or laser gyros.

The other part of the gyro market concerning lower performance, but essentially lower priced gyros for stabilization or short time navigation, has received less attention. To realize low prices, the standard solution of the so-called minimum configuration with nonreciprocal phase modulation seems to be inadequate, requiring too complex photonics and fast electronics circuits.

Here a different approach is presented, basing on a passive all-fiber open loop gyroscope. This new FOG concept is mainly directed to the low cost and low-to-medium performance sector.

8.2. Concept

One of the fundamental problems of interferometric FOGs is that the output signal of a basic Sagnac interferometer is a cosine function of the rotation rate Ω to be detected. That means that each signal recovery scheme has to deal with problems of low sensitivity at low rotation rates, distinction of the sign of rotation (CW or CCW), and ambiguity at higher rotation rates.

While the third problem can be solved easily, that is, by tailoring the sensitivity of the fiber coil to the operational range by correct choice of fiber length and coil diameter, thus staying within the first interference fringe, the solution of the first two is more challenging. What must be achieved is the transformation of the measured cosine function into a sine function, thus yielding high sensitivity at zero rotation rate as well as a distinction of the sign of rotation. This can be realized by either differentiation or by introducing a constant phase shift to the signal.

The usual approach is to apply a nonreciprocal phase modulation to the counterpropagating waves inside the fiber coil. Since 1981 a standard solution for that concept has been established[1] which is well known as the reciprocal minimum configuration.

The argument for using a reciprocal configuration was the canceling of disturbing phase variations introduced by the beam splitter. Yet as the quality of present coupling devices has grown, this problem is not as severe for low-to-medium gyro performances.

The main reason for this demand, however, was the use of polarized light and the stringent requirements for the accuracy of the polarizer which must suppress the second polarization mode. The optimum suppression can be achieved by using the same polarizer twice as input and output ports of

the sensor coil, thus giving a reciprocal configuration. In principle the same result is achievable with an exactly oriented polarizer in the nonreciprocal output arm (first hint by Schiffner)[2], but in reality the required accuracy for the alignment of the output of the polarization axis is difficult to obtain.

This disadvantage of nonreciprocal configurations may be one of the reasons for discarding the introduction of a constant phase shift by means of a 3x3 coupler, which was already proposed in 1980 by Sheem[3]. Such a 3x3 directional coupler consists of three fused tapered single mode fibers with a sufficient narrow spacing of the three guiding cores to allow evanescent field overlap. Due to the mechanism of supermode superposition phase shifts between the lightwaves of the output channels are introduced, which are 120 deg. for a symmetric 3x3 coupler with equal splitting ratio of 1/3 for all channels[7]. Due to its inherent phase shift of 120 deg., this device eliminates the requirement for a nonreciprocal phase modulation in the fiber loop. In spite of the use of a passive component, it operates automatically in the vicinity of its maximum sensitivity point. The basic configuration is shown in Fig. 1. Light from light source L is fed into one input port e_1 of a 3x3 coupler.

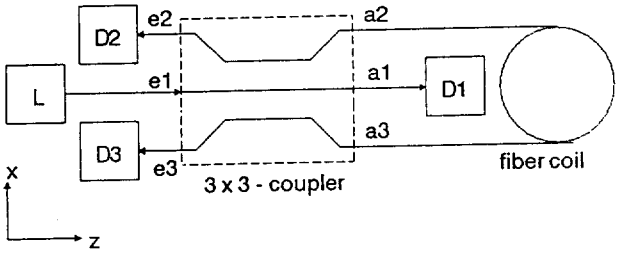


Fig. 1 Principal configuration of a passive FOG

The apparent drawback of such a principally nonreciprocal configuration can be avoided if unpolarized light is used. In this case no polarizer is necesssary, thus eliminating the alignment problem of the optical axis. The use of unpolarized light reduces offset errors of the determined rotation rate just as the combination of polarized light and well-aligned polarizers.[4-6] Consequently, the 3x3 coupler proves to be the ideal device for passive gyro applicaion. Output ports a_2 and a_3 are spliced to the rotation sensing fiber coil, thus launching two lightwaves propagating in clockwise and counterclockwise directions in the fiber coil. The returning lightwaves couple back into the 3x3 coupler interfering along its coupling region. The interference signals are detected by photodiodes D_2 and D_3 at ports e_2 and e_3. Finally, the last port a_1 is coupled to a third photodiode D_1.

The influence of the optical properties of the individual components on the gyro performance will be studied by means of Jones calculus.

8.2.1 Coherence Matrix Representation of Polarization

To characterize the state and degree of polarization, the coherence matrix formalism[8] will be used. Any two orthogonal polarization states may form the basis for the coherence matrix. In the following, the basic states are chosen as linearly polarized along the x and y axes. Let E_x and E_y be the projections of the electric field of a lightwave on this basis and the coherence matrix J in the quasimono-chromatic approach is given by

$$J(\lambda) = \begin{bmatrix} \langle E_x(\lambda)E_x^*(\lambda)\rangle & \langle E_x(\lambda)E_y^*(\lambda)\rangle \\ \langle E_y(\lambda)E_x^*(\lambda)\rangle & \langle E_y(\lambda)E_y^*(\lambda)\rangle \end{bmatrix},$$

where the brackets $< \cdot >$ signify the time average and λ is the wavelength.
The coherence matrix is Hermitian. The degree of polarization P is determined by the eigenvalues (Λ_1,Λ_2), and intensity I is given by the trace

$$J_{ij} = J_{ji}^*,$$

$$P = \frac{|\Lambda_1 - \Lambda_2|}{\Lambda_1 + \Lambda_2},$$

$$I(\lambda) = Tr\{J(\lambda)\}.$$

Equation (1) shows the coherence matrix for a lightwave with degree of polarization P along the x-axis:

$$J(\lambda) = \begin{bmatrix} (1+P)/2 & 0 \\ 0 & (1-P)/2 \end{bmatrix} I_0(\lambda), \qquad (1)$$

where $I_0(\lambda)$ is the spectral intensity distribution of the lightwave.

If a lightwave J_{in} passes through a linear birefringent system, characterized by Jones matrix G, the output coherence matrix J_{out} is given by

$$J_{out} = G \cdot J_{in} \cdot G^+,$$

where + denotes the Hermitian conjugate and $\cdot$ is the scalar product. Consequently, intensity I_{out} of the output lightwave is the trace of J_{out}.

8.2.2 Jones Matrix Representation of the Optical FOG Configuration

The light emitted from source L (see Fig. 1) is assumed to be partially polarized in the x-direction with degree of polarization P and spectral distribution $I_0(\lambda)$. Thus its coherence matrix J_L is given by Eq. (5). The axis of the fiber coil is orientated into the y-direction perpendicular to the z-axis, forming an angle δ with the y-axis.
By using the coherence matrix representation, the detected optical power I_1, I_2, and I_3 at detectors D_1, D_2, and D_3 is

$$I_i(\lambda) = Tr\{G_i(\lambda) \cdot J_L(\lambda) \cdot G_i^+(\lambda)\}, \qquad (2)$$

with

$$G_1 = A_{11}, \tag{3}$$

$$G_2 = A_{23} \cdot S_{CW} \cdot A_{21} + A_{22} \cdot S_{CCW} \cdot A_{31}, \tag{4}$$

$$G_3 = A_{33} \cdot S_{CW} \cdot A_{21} + A_{32} \cdot S_{CCW} \cdot A_{31}, \tag{5}$$

The A_{ij} are Jones matrices of the 3x3 coupler, which is assumed to be polarization independent:

$$A_{ij} = a_{ij}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} = r_{ij}\exp(j\Phi_{ij})\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \tag{6}$$

Here the a_{ij} are the coefficients of the reduced symmetric scattering matrix of the 3x3 coupler, which can be easily evaluated experimentally by measuring the splitting ratio matrix of the coupler.[9,10] An indeal symmetrical 3x3 coupler has equal amplitudes $r_{ij} = \sqrt{1/3}$ for all i and j, and phase shifts $\phi_{ij} = -120°$ for $i \neq j$ and $\phi_{ij} = 0°$ for $i = j$.

The S_{CW} and S_{CCW} describe the clockwise and counterclockwise transmission properties of the fiber coil. These matrices are dependent of the Sagnac phase shift ϕ_s:

$$S_{CW} = \exp(+j\Phi_s/2)S_{0CW}, \tag{7}$$

$$S_{CCW} = \exp(-j\Phi_s/2)S_{0CCW}, \tag{8}$$

where ϕ_s is proportional to rotation rate Ω of the fiber coil

$$\Phi_s = \frac{8\pi AN}{\lambda c}\Omega. \tag{9}$$

Thus, ϕ_s depends on area A of the fiber coil, the number N of windings of the fiber, wavelength λ, and the speed of light in vacuum c.

S_{0CW} is the product of

$$S_{0CW} = \begin{bmatrix}\cos\delta & -\sin\delta\\ \sin\delta & \cos\delta\end{bmatrix}\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}a & b^*\\ -b & a^*\end{bmatrix}\begin{bmatrix}\cos\delta & \sin\delta\\ -\sin\delta & \cos\delta\end{bmatrix} \tag{10}$$

By means of the first and last terms of the matrix product an arbitrary angle δ between the axis of fiber coil y and *y*-axis of the basis coordinate system has been taken into account. The second matrix describes the change of sign of the *x*-component of the electric field vector of the lightwave due to the reflection symmetry of the coordinate systems of incoming to outcoming light of the fiber coil with respect to coil axis y. The third term finally is related to the birefringence properties of the optical fiber inside the fiber coil, made up of individual parts of different birefringence, connected by successive coupling centers.[5,11] Even for nominal standard single mode fibers there exists a certain amount of

birefringence due to lateral internal stress in the case of bendings.[16] Coupling centers are formed, if the circular symmetry of the cross section of the fiber is disturbed e. g. by external forces.In loss-less systems, a and b are related by $aa^* + bb^* = 1$. In this case the whole matrix product in Eq. (10) may be expressed without loss of generality by a single matrix:

$$S_{0CW} = \begin{bmatrix} -\sqrt{1-\alpha}\exp(-j\xi) - \sqrt{\alpha} & \exp(-j\eta) \\ -\sqrt{\alpha} & \exp(+j\eta) + \sqrt{1-\alpha}\exp(+j\xi) \end{bmatrix} \qquad (11)$$

Here α describes the effective coupling strength between the x- and y-polarizations, and ξ and η are effective phase shifts induced by birefringence and coupling centers.

Matrices S_{0CW} and S_{0CCW} of the rotationless fiber coil are related:

8.2.3 Resulting Output Signals at the Photodiodes

Output signals P_1, P_2, and P_3 of photodiodes D_1, D_2, and D_3, which follow from an evaluation of Eqs. (1) - (11) for a not necessarily symmetrical 3x3 coupler, are given by

$$P_1 = v_1 D_1 I_0, \qquad (12)$$

$$P_2 = v_2[A_2 + kB_2\cos(\Phi_s - C_2 + \Phi_F)]d_s I_0, \qquad (13)$$

$$P_3 = v_3[A_3 + kB_3\cos(\Phi_s + C_3 + \Phi_F)]d_s I_0, \qquad (14)$$

where
- v_i = individual sensitivity of the photodiodes i = 1,2,3;
- A_i = constant part of the output signals;
- B_i = amplitudes of the output signals;
- C_i = phase angle introduced by the 3 x 3 coupler (≈120°)
- D_i = splitting ratio of branch 1 of the 3 x 3 coupler (≈ 1/3);
- d_s = attenuation factor of the fiber coil and splices;
- I_0 = intensity of the light source;
- ϕ_F = phase angle introduced by coupling centers;
- k = contrast factor of the interference signal; and
- ϕ_s = Sagnac phase.

Here, the coefficients A_i, B_i, C_i, and D_1 depend exclusively on the optical properties of the 3x3 coupler described by the scattering matrix in Eq. (6):

$$A_i = r_{21}^2 r_{i3}^2 + r_{31}^2 r_{2i}^2, \qquad (15)$$

$$B_i = 2r_{21}r_{i3}r_{31}r_{2i}, \qquad (16)$$

$$C_i = \Phi_{21} + \Phi_{i3} - \Phi_{31} - \Phi_{2i}, \qquad (17)$$

$$D_1 = r_{11}^2. \qquad (18)$$

On the other hand, contrast k and phase offset ϕ_F depend only on the properties of the fiber coil:

$$k = \sqrt{[1 - \alpha + \alpha \cos(2\eta)]^2 + P^2[\alpha \sin(2\eta)]^2}, \qquad (19)$$

$$\Phi_F = \arctan\left[P \frac{\alpha \sin(2\eta)}{1 - \alpha + \alpha \cos(2\eta)}\right]. \qquad (20)$$

To demonstrate qualitatively the polarization sensitivity of the gyroscope, a totally symmetrical 3x3 coupler (equal coupling constants between each channel) has been assumed. The angle between the vector of the residual polarization of the light source and the direction of the fiber coil axis is deliberately set to 45° which means worst case conditions with respect to possible offsets of the output signals. Figure 2 shows the signals at photodiodes D_2 and D_3 as functions of Sagnac phase ϕ_s for an ideal nonbirefringent fiber. The 120° phase shift of the signals caused by the transmission properties of the 3x3 coupler allows low rotation rate detection near the quadrature point. A more realistic fiber coil is linearly birefringent with one or more coupling centers.[10] As an example, Fig. 3 shows the output signals for such a fiber with $\alpha = 0.55$, $\xi = 99.6°$, $\eta = 58.8°$, for degrees of polarization P = 1, P = 0.5, P = 0.2, P = 0.1, and P = 0, respectively. The presence of bending induced birefringent coupling centers lowers the contrast of the interference signal. A significant amount of phase offset is also introduced.

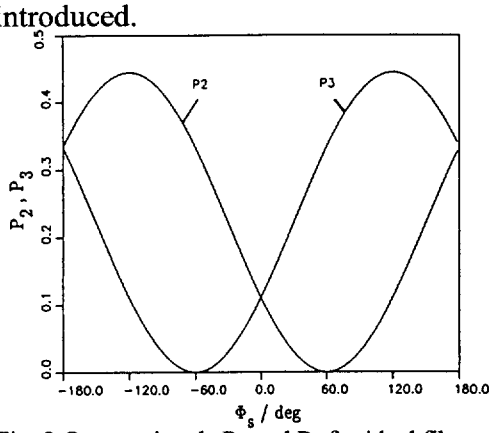


Fig. 2 Output signals P_2 and P_3 for ideal fiber coil

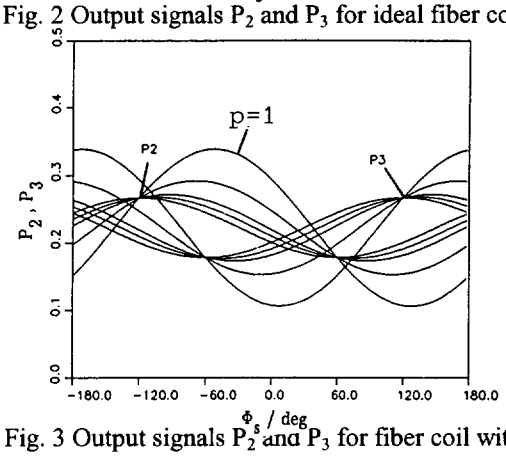


Fig. 3 Output signals P_2 and P_3 for fiber coil with coupling centers

As a consequence, fluctuations of the coupling centers and the birefringent properties of the fiber coil may result in unpredictable fluctuations of phase shift ϕ_F and contrast factor k, thus introducing errors in the derived Sagnac phase and the corresponding rotation rate.

8.3. Contrast Insensitive Signal Recovery

To overcome the problems of bias drift and signal fading caused by unavoidable indeterministic polarization fluctuations inside the fiber coil, a contrast insensitive signal recovery scheme was conceived.

8.3.1 Elimination of Random Bias Variations caused by ϕ_F

According to Eq. (20) the birefringence-induced phase shift ϕ_F is a function of the polarization properties α and η of the fiber coil and of the degree of polarization P. With P = 0 we get

$$\Phi_F(P = 0) = 0$$

independent of the actual values of α and η.

Consequently, to reduce ϕ_F to negligibly small values, an essentially unpolarized light source must be used. In the present experiment, this has been accomplished by means of an edge emitting luminescence diode (ELED) at 1.3 μm. The degree of polarization has been measured to be in the 1 - 5% range, which may be further lowered by means of a Lyot type fiber depolarizer between the ELED and the 3x3 coupler, or at one of the ends of the fiber coil.

8.3.2 Elimination of Random Scale Factor and Bias Variations caused by Contrast Factor *k*

Indeterministic fluctuations of contrast factor *k* of the interference signal affect the scale factor stability of the gyroscope by changing the amplitudes of the cosine functions in Eqs. (13) - (14). Furthermore, in the realistic case of nonsymmetric 3x3 couplers, where $A_2 \neq A_3$, $B_2 \neq B_3$, and $C_2 \neq C_3$, the bias for zero rotation rate is also affected.

This problem can be overcome by taking into account all the information content of the three signals P_1, P_2, and P_3 of the photodiodes. These signals are redundant with respect to rotation rate Ω. Consequently, Eqs. (12) - (14) allow an analytical elimination of the disturbing unknown contrast factor *k* as well as unknown fluctuations of optical power I_0 of the light source. Thus, a single equation can be obtained, relating the rotation rate Ω to be determined to the output signals P_1, P_2,

and P_3 without further dependence of *k* in I_0:

$$\frac{(P_2 - P_3) + f_1 P_1}{(P_2 + P_3) - f_2 P_1} = f_3 \frac{\sin(f_6\Omega + f_4)}{\cos(f_6\Omega + f_5)}, \qquad (21)$$

where the six coefficients, f_1–f_6, are given by

$$f_1 = \frac{v_3A_3 - v_2A_2}{v_1D_1} d_s \simeq 0,$$

$$f_2 = \frac{v_3A_3 + v_2A_2}{v_1D_1} d_s,$$

$$f_3 = \frac{\sqrt{[v_2B_2\cos(C_2) - v_3B_3\cos(C_3)]^2 + [v_2B_2\sin(C_2) + v_3B_3\sin(C_3)]^2}}{\sqrt{[v_2B_2\cos(C_2) + v_3B_3\cos(C_3)]^2 + [v_2B_2\sin(C_2) - v_3B_3\sin(C_3)]^2}},$$

$$f_4 = \arctan\left[\frac{v_2B_2\cos(C_2) - v_3B_3\cos(C_3)}{v_2B_2\sin(C_2) + v_3B_3\sin(C_3)}\right] \simeq 0,$$

$$f_5 = \arctan\left[-\frac{v_2B_2\sin(C_2) - v_3B_3\sin(C_3)}{v_2B_2\cos(C_2) + v_3B_3\cos(C_3)}\right] \simeq 0,$$

$$f_6 = \frac{8\pi AN}{\lambda c}.$$

8.3.3 Calibration

The six coefficients, f_1 - f_6, depending on the performance of the light source of the 3x3 coupler and the photodiodes but not on the polarization dependent birefringence properties of the fiber coil, may be determined by means of the following calibration procedure:

Registration of P_1, P_2, and P_3 at well-known rotation rates over the total range of rotation rate, as shown by Fig. 4

Calculation of f_1 - f_6 by means of a nonlinear equation solver.

Storage of f_1 - f_6 in the memory of the gyroscope microcontroller.

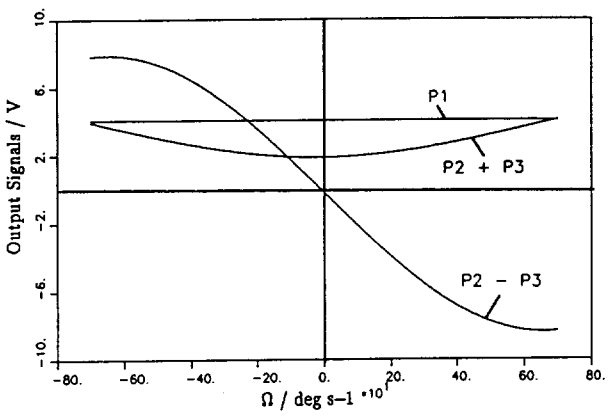


Fig. 4 Measured output signals (P_2 - P_3) and (P_2 + P_3) as functions of rotation rate

Repetition of this procedure at different temperatures and storage of the obtained coefficients together with the temperature data of implemented temperature sensors.

After such an initial calibration procedure the rotation rate may be determined by the gyroscope microcontroller from photodiode signals P_1, P_2, and P_3 according to Eq. (21), allowing variations of the intensity of the light source as well as indeterministic polarization transmission fluctuations within the fiber coil. The associated set of coefficients f_1 - f_6 for the actual internal temperature of the FOG is calculated by interpolation from the
coefficient matrix at the fixed temperature points, generated at the calibration. To avoid the case of zero contrast, where the signal recovery scheme would fail because of lack of information, a crude Lyot depolarizer[12] may be implemented at one end of the fiber coil, yielding a contrast factor with slight fluctuations around the mean value of 0.5.

As a consequence, the requirements normally imposed on the optical quality of the fiber coil can be drastically relaxed. Furthermore, as the splitting ratio properties of the 3x3 coupler
are fully described by the coefficients f_1 - f_6, no strong tolerance requirements on the symmetry an reproducibility of the 3x3 coupler are necessary, allowing the use of cheap, nonselected samples. Thus, the contrast insensitive signal recovery scheme lowers the requirements for the performance of optical components but retains the accuracy of the measured rotation rate.

8.3.4. Results

By using the temperature compensation method described above, a gyro with a coil of 250 m fibre length for a maximum rotation rate of 1000 deg/s has been tested over an extended temperature range.[13]

Fig. 5 shows the bias for temperatures between -54° and +75 deg centigrade, resulting in a 1 sigma value of approximately 100 deg/h.

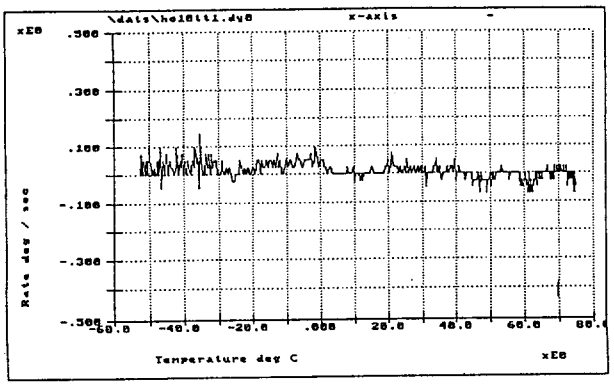


Fig. 5 Bias Error as Function of Temperature

The scale factor accuracy is presented in Fig.6

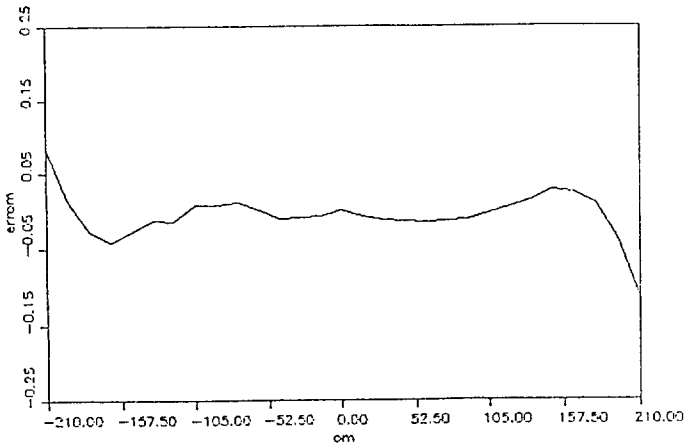


Fig. 6 Scale Factor Error as Function of Rotation Rate

8.4. Application: Passive FOG IMU for Advanced Short Range Air to Air Missiles

This passive FOG has found an application[14] within the Inertial Measurement Unit (IMU) of an advanced short range air to air missile

8.4.1 Basic Configuration

For this purpose, the FOG design described above has been extended to a three axis sensor[15] by positioning three sensing fiber coils orthogonally to each other. The principal configuration is shin Fig. 7. Light, emitted from a temperature stabilized edge emitting lumiscence diode (ELED)is launched into a first 3x3 coupler, which divides the intensity equally into the three sensing blocks, each consisting of an optical fiber coil, a 3x3 coupler and its photodetectors. Only one photodiode P_1 is necessary in order to monitor and compensate variations of the light intensity of the ELED.

By using integrating A/D-converters, the FOG integrates the rotation rate signals during each integration time interval without
loss of any information. Thus the signal output inherently is an angular increment. By dividing this value by the integration time, the mean value of the rotation rate during each integration time interval is calculated.

The acceleration sensing is performed by three individual linear vibrating beam accelero-meters, mounted orthogonally to each other.

This type of accelerometer employs a flexure-suspended pendulous mass in a gas damped, shock caged mechanism. The pendulum is constrained by two resonating quartz beam elements which operate in a push-pull configuration. The pendulum exerts a force on the quartz beams which convert the input acceleration into a varying frequency output that interfaces directly with digital counters, eliminating the need for A/D-conversion. The output signal is obtained by measuring the frequency output of the two quartz crystal beams and the external temperature, and then inputting these values to an algorithm to yield a thermally compensated output in acceleration. This is performed within the common IMU electronics by the same ASIC, which is also used for the control of the integrating A/D-conversion of the current of the photodetectors of the FOG cluster.

8.4.2 Hardware

Theoretical thermal models have been developed in order to avoid hot spots within the IMU, including the internal heating of the IMU by its own power consumption during air carriage as well as the kinetic heating of the missile skin during most of the aircraft flight profiles and during missile free flight. Based on these calculations the final mechanical design of
the IMU structure, which is shown in Fig. 8, had been derived.
A rigid block inside is the basic fixing structure for the three orthogonal fiber coils and the accelerometers. Situated on top is the analog hybrid board for the preamplification and analog to digital conversion of the photodiode signals. Below of that board is the photonics block with the photodetectors, the 3x3 couplers, and the groove for fiber placement. The light source is an ELED launching an optical power of typically 20 μW into the single mode fiber at a wavelength of 1.3 μm. The detectors are InGaAs PINs with 0.8A/W responsitivity. The 3x3 coupler is a fused fiber coupler with a splitting ratio of 0.33±0.03 and excess loss of 0.15 dB. The coil consists of a standard single mode fiber with a depolarizer.

Mounted around the block is the digital electronics circuit. Electrical connections between the several boards are realised by means of connectors on flex layers. Two of the fiber coils can be seen partly at the sides of the hybrids. The IMU fits into a housing of approximately 140 mm diameter and 70 mm height.

8.4.3 Performance

The product performance has been evaluated theoretically by means of a simulation program and experimentally by a series of tests on component and IMU level. The simulation program describes mathematically the properties the IMU. Its purpose is:

- to investigate the influence of the individual optical components on the overall performance
- to identify the most critical components with respect to error terms
- to find criteria for the requirement specification of the components
- to provide a performance simulation subroutine of the IMU for implementation into the overall missile6 DOF program for missile performance evaluation.

These performance studies have been validated by an excessive test program with IMU hardware. With an error budget of the bias in in the range of hundred deg per hour and scalefactor errors of 2000 ppm for a maximum rotation rate of 800 deg/s the gyro provides an accuracy in excess to the requirements of a typical short range missile. The accelerometers are in the accuracy range of 20 mg and 1000 ppm.

8.5.Summary

In summary the above described passive FOG has matured to a qualified product, readily available for the stabilization and short term navigation of short range missiles.

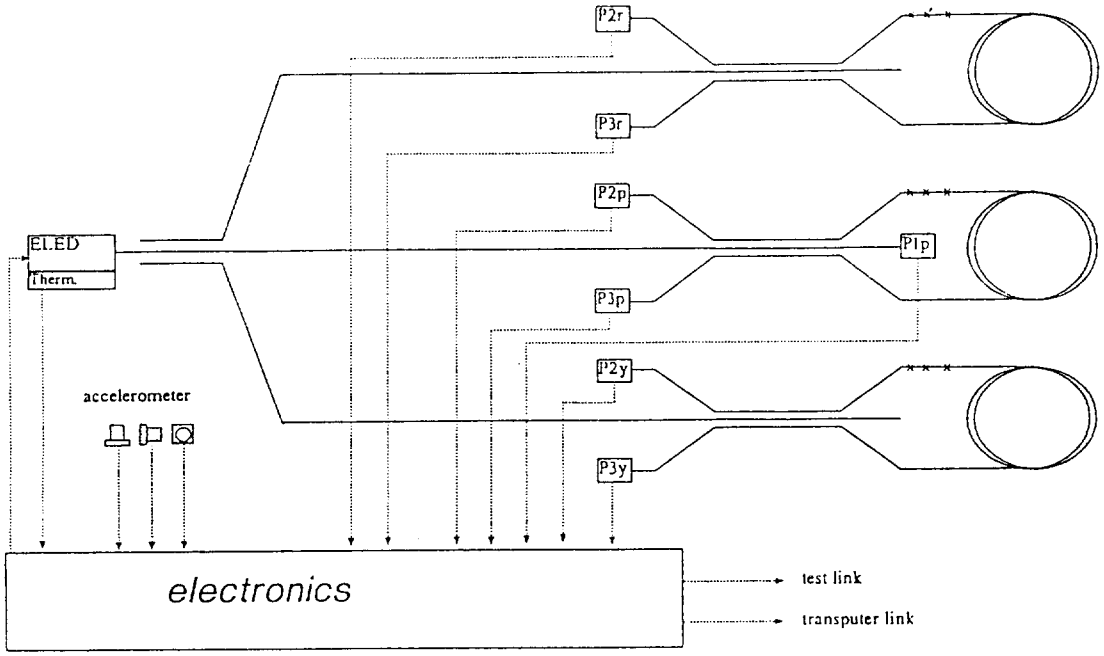


Fig. 7 Principle configuration of the IMU

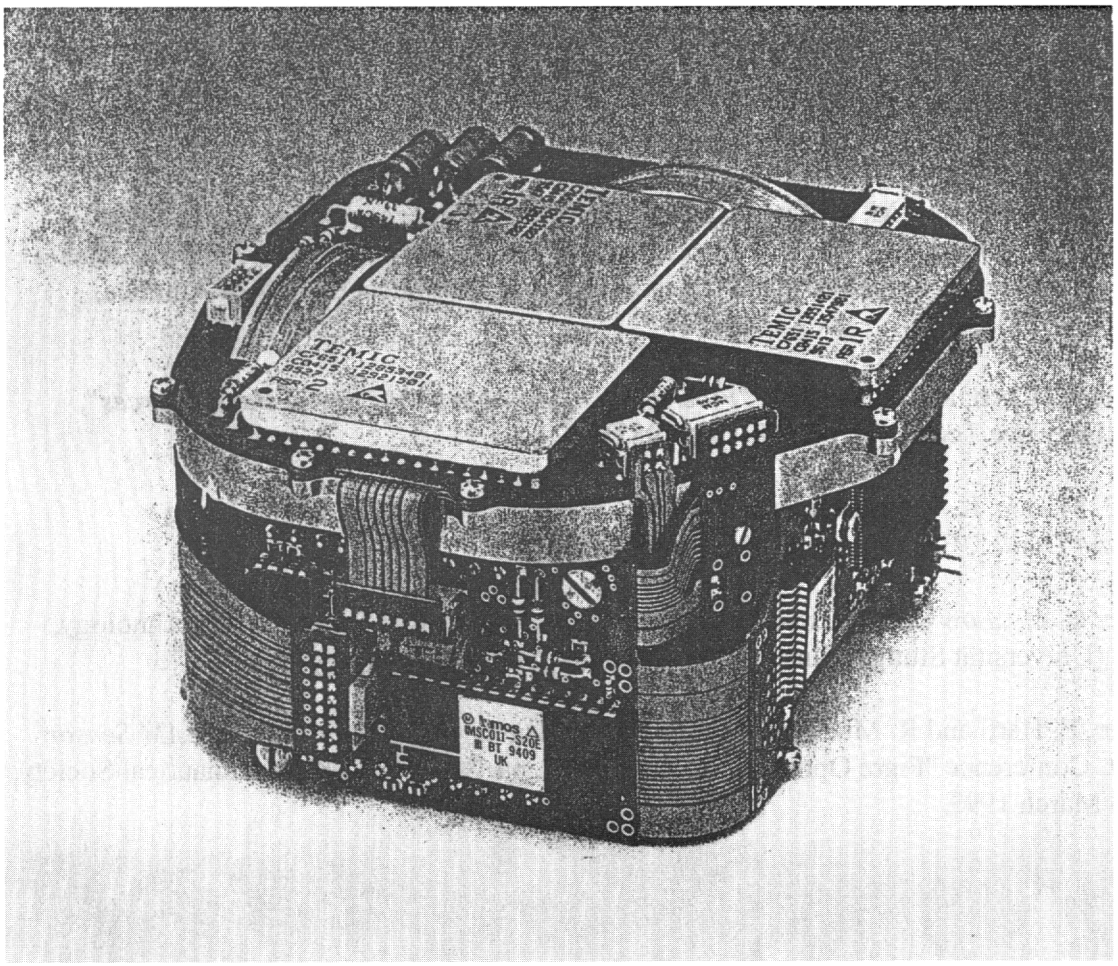


Fig. 8 IMU Hardware

8.6. References

1. S. Ezekiel and H. J. Arditty, *Fiber-Optic Rotation Sensors* (Springer-Verlag, Berlin, 1982)

2. G. Schiffner, W.R. Leeb, H. Krammer, and J. Wittmann, *"Reciprocity of Birefringent Single-Mode Fibers for Optical Gyros"*, (Appl. Opt. 18, 2096-2097 (1979)

3. S. K. Sheem, *"Fiber Optic Gyroscope with (3x3)-Directional Coupler"*, Appl. Phys. Lett. 37, 869-871 (1980)

4. K. Petermann and P. Russer, *"Ringinterferometer"*, West German Patent DE 3,006,580A1 (1980)

5. G. A. Pavlath and H. J. Shaw, *"Birefringence and Polarization Effects in Fiber Gyroscopes"*, Appl. Opt. 21, 1752-1757 (1982)

6. W. K. Burns, R. P. Moeller, and C. A. Villarruel, *"Observation of Low Noise in a Passive Fibre Gyroscope"*, Electron. Lett. 18, 648-650 (1982)

7. S. K. Sheem, *"Optical Fiber Interferometer with (3x3) Directional Couplers: Analysis"*, J. Appl. Phys. 52, 3865-3872 (1981)

8. M. Born and E. Wolf, *Principles of Optics* (Pergamon, London, 1975), pp. 544-553

9. G. Trommer, *"Wavelength Dependence of 3x3 Fiber Coupler for Gyro Application"*, Electron. Lett. 25, 944-945 (1989)

10. M. A. Davis, A. D. Kersey, M. J. Marone, and A. Dandridge, *Characterization of 3x3 Fiber Couplers for Passive Homodyne Systems: Polarization and Temperature Sensitivity"*,Optical Fiber Communication Conference, 1989 Technical Digest Series, Vol. 5 (Optical Society of America, Washington, DC, 1988), p. 103

11. W. K. Burns, C. Chen, and R. Moeller, *"Fiber-Optic Gyroscopes with Broad-Band Sources"*, IEEE/OSA J. Lightwave Technol. LT-1, 98-105 (1983)

12. K. Bohm, P. Marten, K. Petermann, and R. Ulrich, *"Low Drift Fibre Gyroscope Using a Superluminescent Diode"*, Electron. Lett. 17, 352-353 (1981)

13. G. Trommer et al., *„Low Cost Passive Fiber Optic Gyroscope"*, Proc. Symp. Gyro Technology, Stuttgart 1991, (Universität Stuttgart, Institut für Mechanik, Stuttgart, 1991) pp. 17.0-17.13

14. G. Trommer, E. Hartl and R. Mueller, *"Passive Fibre Optic Gyroscope for the ASRAAM Sensor Unit"*, Proc. Int. Conference "Figre Optics in Guided Flight", pp. 3.1-3.12 Royal Aeronautical Society, London, UK, 3 March 1993

15. G. Trommer, *"Passive Fiber Optic Inertial Measurement Unit: From Qualification towards Series Production"*, Proc. SPIE 2837, pp. 124-131 (1996), Int. Symposium on Optical Science, Engineering and Instrumentation Fiber Optcs Gyros: 20th Anniversary conference, Denver, CO, USA

16. R. Ulrich, S. C. Rashleigh, W. Eickhoff, *„Bending-induced birefringence in single-mode fibers"*, Opt. Lett. 5, 273 - 275 (1980)

9. MINIATURE FIBER OPTIC GYRO. FIZOPTIKA IMPLEMENTATION.

V. Listvin
V. Logozinski
V. Solomatin
Fizoptika Co.
109387, PO box 16, Moscow, Russia

SUMMARY

The results of development of the fiber optic rotation sensors (FOG) for industrial applications performed during last several years in FIZOPTIKA Co. are reflected in the report. The major primary FOG components and main steps of the in-line spliceless technology for FOG manufacturing are described. The parameters, main features, and applications of standard FOG models are presented. Specific FOG errors are analyzed and potentials to improve its performance are discussed.

9.1 INTRODUCTION

It has been recognized that open loop fiber optic gyro (FOG) is a cost effective solution for practical use. However the one-fiber manufacturing technique developed by Fizoptika is exceptionally different from other approaches. In this method only one fiber length is used to form all optical components (sensing coil, modulator, couplers, polarizer). The fiber optic components are fabricated step by step along fiber length, so that optical connection between them naturally occurs without splicing. Performance of each components is a subject of in-process control. Special fiber is used for this technique which has main feature - extremely low outer diameter (45μ). It was developed for FOG application. All components fabricated from such fiber have very small size and may be placed in a small volume making possible building of miniature FOG assembly.

The concept of one-fiber manufacturing (in-line technology) demands development of the new methods for fiber optic components fabrication. Special technique based on fusion-tapering and soldering process was created that gives stable over time and temperature components performance (coupler, polarizer, SLD-module).

FOG manufacturing in Fizoptika production line is performed in a following sequence: it begins with winding of sensing coil and PZT modulator (length range 50...200m); next coil coupler, second coupler and polarizer are fabricated. SLD module is made by aligning the SLD crystal to one fiber tail and fix it by soldering. Finally another fiber tail is aligned to photodiode to form receiving module. FOG's optical part is completed by this procedure.

FOG contains both optical part and electronics block which is produced separately. FOG open loop electronics controls PZT modulator, SLD and detects signal of optical block to convert it into rate proportional voltage. Optical block and electronics are integrated into one system which passes acceptance tests to conform performance. Production capacitance of one production line is currently 1000 gyros per year. Modern generation of production equipment is able to produce 3..5 thousand gyros per year.

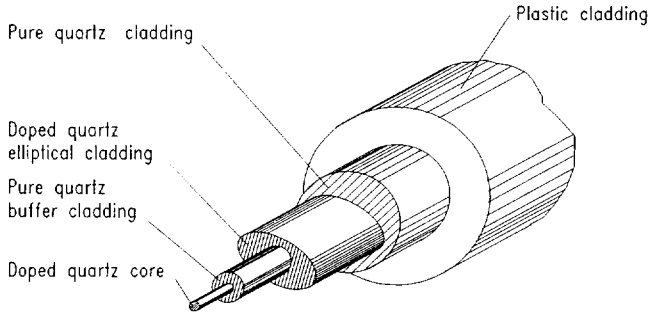
There are several basic models in Fizoptika's production which are delivered to customers:
- VG941 is the most miniature worldwide FOG. Its size ∅ 27 x 58 mm. Recently this model was modernized by implementing gyro electronics inside. VG941 is already widely used in testing equipment for car and aircraft industry.
- VG910 is a FOG which was designed for variety of commercial and industrial applications. Optimal balance between gyro size, cost and performance was realized in this model.Currently it is used in quantity for image stabilization purpose (TV and photo cameras professional, sights) and for river ship motion control.
- VG951 was developed for navigation and accurate rate measurement. Its drift is about 1 deg/hr. Recent development opened opportunity to built gyro of 0.1 deg/hr based on this model which would have great potential for application in marine compassing systems.

In the following sections we briefly describe major FOG components and several FOG models in more details. Analysis of the specific FOG errors is done and potentials to improve FOG performance are discussed.

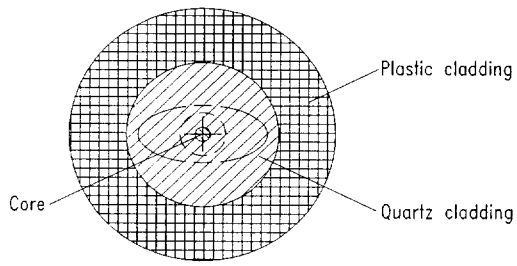
9.2 PRIMARY COMPONENTS

FOG fiber is strongly bent inside the miniature gyro. Its aperture should be high to reduce excess optical loss. Outer diameter is to be very small to reduce mechanical stress which determines reliability and lifetime of quartz components. These were the major reasons for choice of FOG fiber. Its parameters are: cutoff - 780nm; optical loss 10 dB;

numerical aperture - 0.2; polarization beat length - 3 mm; diameter of quartz cladding - 45μ; diameter of plastic (acrylate) coating - 120μ; core diameter - 2μ; H-parameter - 0.001 1/m.



FOG-fiber visible structure.



Intermediate layers are not visible due to refractive index matching.

Fig.9.1.

Fiber design may be described as follows (see Fig.9.1). Optical and mechanical parameters of the fiber are almost unchangeable along its length. Light guiding core is surrounded by buffer layer made from pure quartz to reduce optical loss and to prevent core deformation during fiber manufacturing process. Intermediate elliptical cladding creates at normal temperature cross mechanical tension in the core which is necessary to induce birefringence. The content for elliptical cladding has been selected to match quartz index. This property is extremely important for successful use of tapering technique for main components fabrication.

SLD chip is used for light emitting module manufacturing. It is based on the same structure as a laser diode. Both facets of SLD crystal are made non reflecting to suppress laser mode. Spatial structure of emitted radiation depends on the waveguide formed in the body of SLD crystal. Light amplification which is realized in SLD permits to reach high brightness and low coherence. SLD parameters are: optical power > 3 mW; wavelength - 820 nm; spectrum bandwidth > 15 nm

9.3. FABRICATED COMPONENTS

Fiber optic coupler is made by fusion and tapering technique. Two single-mode fibers are installed together and held by two moveable supports. An arc-flame is used to melt fibers. The supports are moved apart so that a fused tapered region is formed. The speed of separation and heating length control the shape of the taper which has a significant influence on the coupler performance. The movement of the fiber supports is controlled by the sensor providing elongation data. The quality of the single mode fiber used in the fusion process is extremely important. Both fibers should be highly identical. In our case this leads to requirement of very high stability of the fiber parameters along its length. Coupling occurs at about 4 mm taper length and goes without loss. Melting of fibers takes place at very high temperature (>1000°C) so in-process coupling ratio control should take into account coupling temperature dependence. Small fiber diameter result in very small size of the coupler. Its length is only 15mm. Only fiber as row material is utilized for coupler fabrication. That's why fused coupler is intrinsically stable and rigid. Quartz substrate is used to fix the coupler. No additional tension occurs over temperature.

Fiber optic polarizer is a very important FOG component. Its quality influences directly on bias behavior. Crystal-fiber polarizer originally developed in Fizoptika is in-line fiber optic component which takes advantages of tapering technique in fabrication and high quality in performance due to perfect design. To manufacture polarizer fiber taper is initially fabricated using same technique as for coupler fabrication. Taper waist is next being buried in the birefringent media (crystal) by crystal growing process. Crystal axis is aligned to break fiber guiding property for one polarized mode and to keep it for another. An apparent advantage of polarizer is very effective interaction of guided wave with external media because crystal contains fiber inside its body. Optical loss for suppressed mode per unit waveguide's length is proportional to $1/d^3$. Exterior view of polarizer is presented in Fig.9.2.

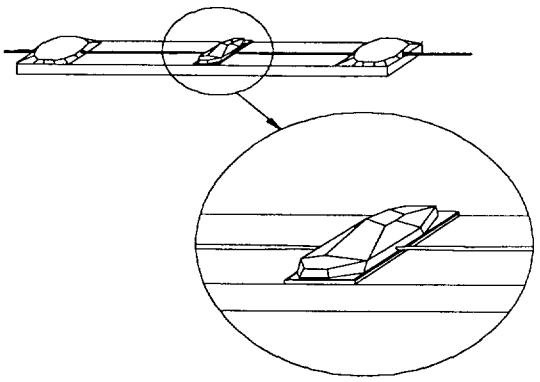
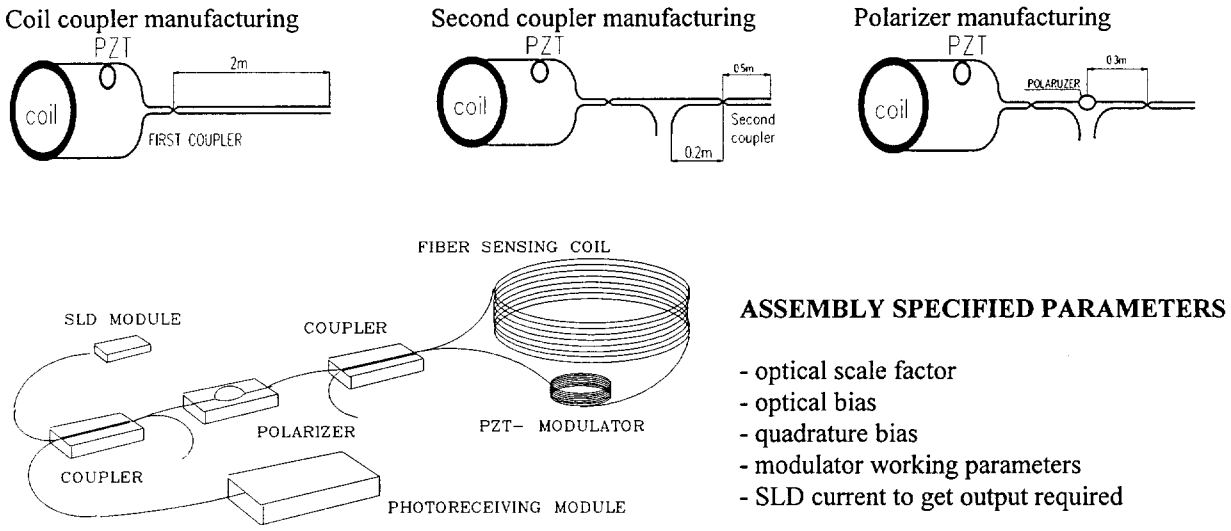


Fig.9.2. Exterior view of fiber optic polarizer

Thus extinction ratio of polarizer rises rapidly with diameter diminishing and may be controlled by taper length. There is also a possibility of polarizer's performance monitoring during the process as in coupler manufacturing. Restrictions of polarizer performance are similar to those of the fused coupler. For successful operation it's necessary to ensure adiabatic transaction of core fundamental mode (HE_{11}). After manufacturing polarizer is fixed on the quartz substrate sized as coupler substrate.

SLD module and photoreceiving module are fabricated by pigtailing technique which includes 3-axis micromotion to align fiber with respect to SLD crystal or photodiode crystal and soldering to fix fiber position. Resultant modules demonstrate very stable behavior over temperature range and over lifetime.

FOG optics fabrication process. The sequence of components fabrication is illustrated in Fig.9.3. During manufacturing the technological light module and receiving module are installed to both coil tails for in-process control of the parameters of all fiber optic components. After entire assembly is finalized its performance is check to conform optics specification.



ASSEMBLY SPECIFIED PARAMETERS

- optical scale factor
- optical bias
- quadrature bias
- modulator working parameters
- SLD current to get output required

Completing optical assembly by pigtailing SLD and photodiode

Fig.9.3. *The main steps of FOG manufacturing.*

9.4 FOG OPERATION

When fiber optic assembly and electronics are connected we obtain one axis gyro system. System is actuated immediately just power is on. In general the output voltage is a function of angular rate Ω and ambient temperature (t°). Within the input range ($-\Omega_{max}$ ...$+\Omega_{max}$) determined by the optical scale factor magnitude a good approximation is given by the expression $U(\Omega, t^\circ) = SF \cdot \Omega \cdot [1 + s(t^\circ)] \cdot [1 - p \cdot (\Omega / \Omega_{max})^2]$.

s(t°) - SF temperature variation; p - linearity factor. Linearity belongs to a very simple algorithm of FOG signal detection realized in our circuit and at $-\Omega_{max}$ could reach 15..20%. From the other side simple circuit permits to achieve very small size of electronics comparable with optical part of the gyro. Moreover linearity is much improved by polynomial approximation to the residual 0.03...0.1% and at low rates it is negligible. After overrating normal operation mode is restored in parts of second. Scale factor variation s(t°) is a reproducible function of temperature with a scale 2..4% over temperature range. It belongs to thermal behavior of electronic components and SLD wavelength temperature dependence. Initial bias and its temperature slope are caused by electronic detecting circuit. The irregular drift depends on optical components quality and determines the limit of gyro accuracy. Output noise is caused by natural quantum properties of light and thermal noise of preamplifier. It results in data variance (RMS) dependent on measurement period.

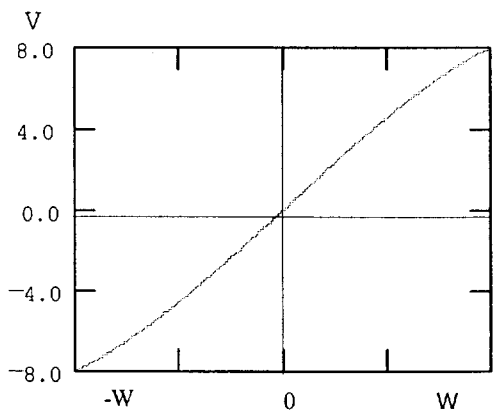


Fig.9.4. *FOG system output characteristic*

9.5. FIZOPTIKA GYRO PRODUCTS

Common features. Open loop FOG offers:
in MEASUREMENT: instant readiness; instant response; high sensitivity and wide dynamic range; very low weight; opportunity to apply at high accelerations and over temperature range;
in POWERING: standard low voltages roughly conditioned, low power consumption- typically <1 Watt;
in INTERFACING: natural and universal analog output.

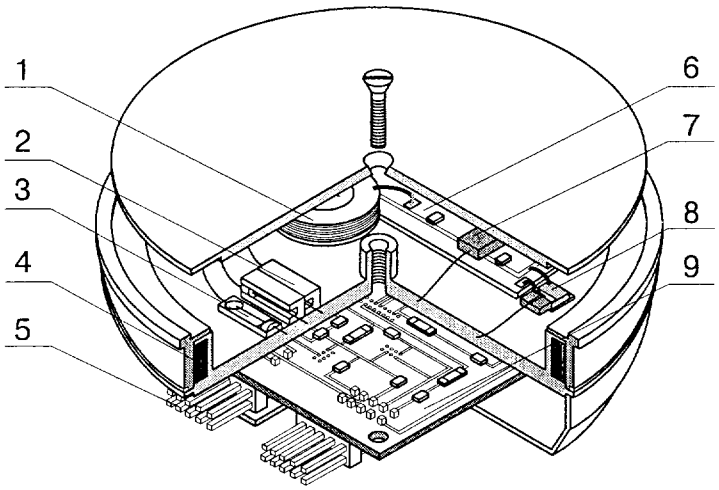
VG910.
VG910 is gyro system incorporating FOG and processing PCB in one housing (see Fig.9.5).

PERFORMANCE OUTLINE

Range	100 to 300 °/s
Scale factor repeatability	0.2 to 0.4%
Bias repeatability	20 to 60 °/h
Bias stability	5 to 15 °/h
Output noise	3 to 6 °/h / √Hz
Dimensions	∅ 80 mm x 18 mm
Weight	110 gram

VG910 FEATURES
Reasonable compromise between size, accuracy and price. Potential to grow in higher performance.

APPLICATIONS
Land navigation, TV stabilization, Inspection systems, GPS integrated systems



1 - PZT phase modulator
2 - Fiber optic coupler
3 - Fiber optic polarizer
4 - Fiber coil
5 - Connectors
6 - Preamplifier board
7 - Photoreceiving module
8 - SLD module
9 - Printed board

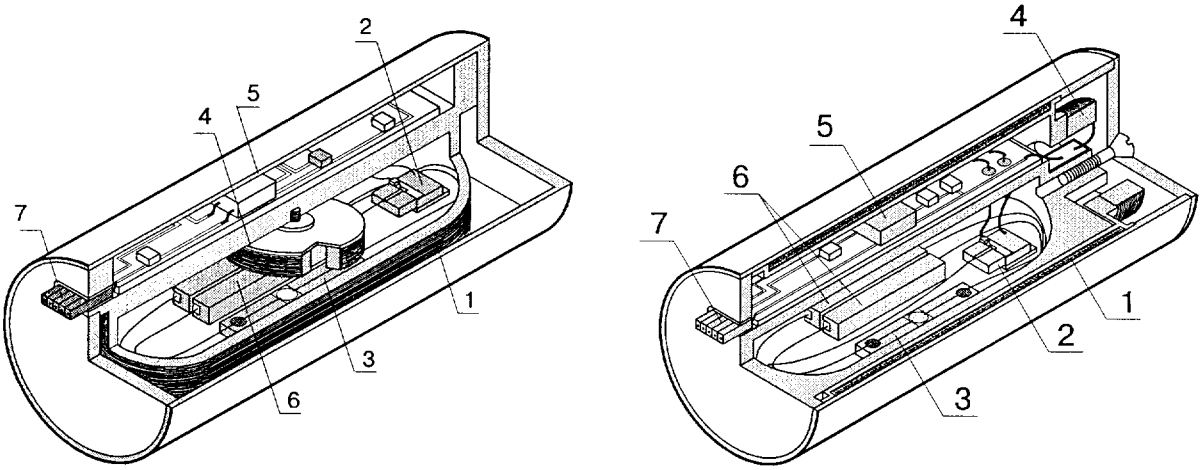
Fig.9.5. *VG910 Internal view*

VG941
VG941 is gyro system incorporating FOG and processing external electronics PCB. VG941 is the smallest FOG (see Fig. 9.6.) sized as ≈ ∅1" x 2".

PERFORMANCE OUTLINE

Range	500 °/s (expandable)
Scale factor repeatability	0.2 to 0.4%
Bias repeatability	0.1 °/s
Bias stability	0.01 to 0.03 °/s
Output noise	0.003 to 0.01 °/s / √Hz
Frequency range	0...1000Hz (expandable)
Dimensions	∅ 25.4 mm x 58 mm
Weight	40 gram

FEATURES Very rigid design. Available in **both** axis to body orientations (VERSION **A** and **B**). Convenient assembling of multi axis systems for a variety of applications. Frequency response may be extremely high (to 3kHz). Recently electronics in SMT implementation has been integrated inside and lower power consumption achieved (typ. 0.7W at just +5V powering).



VG941A VG941B

1 - fiber coil 2 - SLD module 3 - fiber polarizer 4 - PZT modulator
5 - receiving module 6 - fiber coupler 7 - external cable

Fig.9.6. *VG910 Internal view*

APPLICATIONS Inspection systems, dynamic test equipment, motion control

VG951
VG951 is gyro system incorporating FOG with enhanced scale factor and PCB type OE002. This is most accurate model within Fizoptika FOG family. VG951 system does not require data compensation since its electronics block contains circuits for periodical built-in calibration. In the near future due to modification of production technique this model will have accuracy 0.1 deg/hr.

PERFORMANCE OUTLINE

Range	40 °/s
Scale factor repeatability	0.2 %
Bias repeatability	1 to 3°/h
Bias stability	0.3 to 1°/h
Output noise	1°/h / √Hz
Frequency range	0...60 Hz
Dimensions	∅ 150 mm x 30 mm
Weight	650 gram

APPLICATION Precise testing equipment for motion control, navigation equipment.

9.6. FOG MAIN ERRORS. POTENTIAL TO IMPROVE PERFORMANCE.

Miniature FOG manufacturing reveals specific fundamental problems of main optical components quality. These problems are not so essential if FOG has middle or big size. Natural degradation of fiber's H-parameter due to its small diameter and strong bending inside gyro case is unavoidable in a miniature FOG. Moreover, during manufacturing of couplers and polarizer the orientation of birefringence axes of the fibers is not controlled because of twisting flexibility of the fiber. That results in the increasing of FOG bias components dependent on parameters of polarization anisotropy of the FOG.

The sign of this component depends on a radiation frequency and is partially averaged due to SLD spectrum bandwidth. The residual value of this component is the more the bigger coupling between polarization modes of the fiber. There are two components of FOG bias: optical bias and quadrature bias. Optical bias may be in general decreased by enhancing of the extinction ratio of the polarizer. The origin of the quadrature signal is a birefringence modulation of

the fiber wound on the PZT cylinder. This signal is shifted at 90° in phase with respect to useful (gyroscopic) signal. After lock-in detection the quadrature signal is suppressed but not completely that gives at the output the quadrature bias comparable (or even exceeding) optical bias. Modulation of birefringence results in modulation of light intensity after passing polarizer. Depth of intensity modulation depends on polarizer orientation with respect to birefringent axes.

Two main mechanisms of arising of the quadrature signal are explained by Fig.9.7, 9.8a, 9.8b.

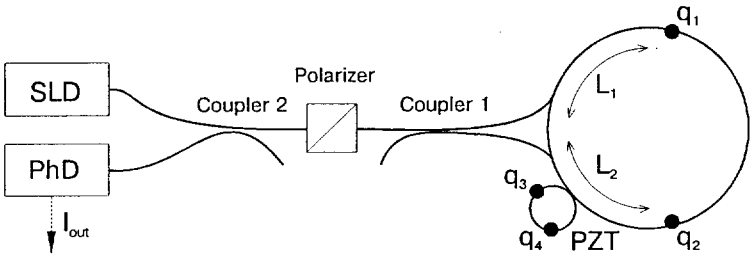
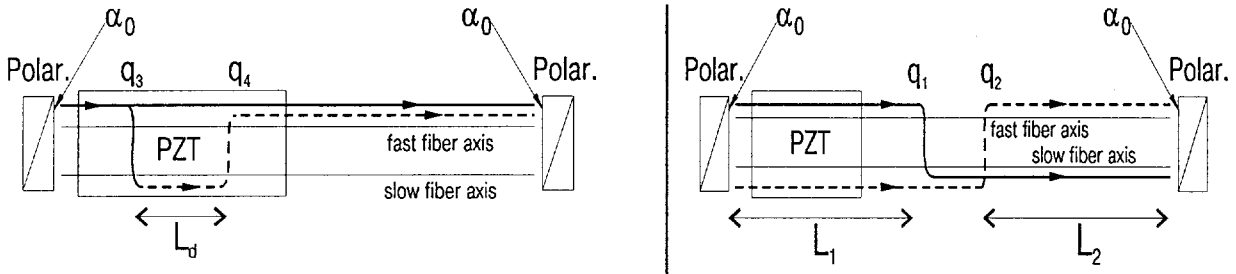


Fig.9.7. Location of the coupling centers in the FOG.

Fig.9.8a. The wave at the input of modulator propagates in the "fast" axis of the fiber. At the coupling center q_3 (fiber inhomogeneity) part of the wave power is branched off to the "slow" fiber axis. At the center q_4 a part of branched power returns back to the "fast" axis. The phase difference between waves propagating in "fast" axis and "slow" axis of the fiber (in the different polarization modes) is modulated during periodical elongation of the fiber. If the distance between centers q_3 and q_4 is less than depolarization length ($L_d = l_c \cdot l_b / \lambda$, l_c - coherence length of radiation, l_b - polarization beat length of the fiber, λ - radiation wave length) the waves are coherent and the power of wave propagating in "fast" axis is modulated.



a). *b).*

Fig.9.8. *Two main mechanisms of arising of the quadrature signal (amplitude modulation of the optical signal in the FOG).*

Fig.9.8b. The wave at the input of modulator propagates both in the "fast" axis of the fiber and in "slow" one. The interfering waves pass the fiber section between centers q_1 and q_2 in the same axis but the sections L_1, L_2 in both axes. In case of $|L_1-L_2| < Ld$ the waves are coherent at the polarizer and after passing polarizer light intensity becomes modulated. Dependence of quadrature signal (q) on polarizer orientation (α_0) with respect to birefringent axes may be estimated as

$$q \sim \delta\varphi\, (L_d/l)\, q_3\, q_4 \cos^2 2\alpha_0 + \delta\varphi\, q_1\, q_2 \sin^2 2\alpha_0$$

where q_1, q_2, q_3, q_4 - the coupling coefficients between polarization modes, l - length of fiber section wound on PZT cylinder, $\delta\varphi$ - modulation amplitude of phase difference between fiber polarization modes. The magnitude of $\delta\varphi$ (~0.1rad) is proportional to fiber elongation and to fiber birefringence. Taking into account that $L_d/l \sim 0.1$ we conclude from the above formula that quadrature signal can be suppressed both by decreasing of $\delta\varphi$ and by polarizer alignment with respect to birefringent axes of fiber.

For the new generation of our FOGs we will align the axis of polarizer with respect to birefringent axes of the fiber and will use PZT modulator with special winding of the fiber providing possibility of $\delta\varphi$ compensation. Such kind of changes in FOG design allows to reach the bias drift below 0.1 deg/hour.

10. Progress in Fiber-Optic Gyro Development and Applications

Wilfried Auch

SFIM Industries Deutschland GmbH
Gottlieb-Daimler-Str. 60
D-71711 Murr, GERMANY

ABSTRACT

The SFIM Industries single-axis fiber-optic gyros and inertial measurement units are described. Application areas are identified and rationale is given for the use of the fiber-optic gyro in different market segments. The applications reach from retrofit of mechanical rate gyros to integrated hybrid navigation systems. The ongoing development and engineering effort is now concentrating on miniaturized multiple-axes solutions. The most promising subsystem is a tactical grade inertial measurement unit with a fiber-optic gyro triad as rotation rate sensors.

1. INTRODUCTION

Single-axis fiber-optic gyro (FOG) basic design can be considered as being completed. Areas in which engineering effort is still necessary are improvements of performance limiting elements and analysis of error mechanisms in practical designs and under industrial production conditions. The main driver for ongoing development activities is the need for cost reduction. The known designs have a large number of common characteristics but have also different solutions to FOG specific problems with respect to the application area and market segment.

Device/System	Configuration	Gyro Performance	Application
RGP Rate Gyro Package (1 axis)	1 FOG	0.1 ... 0.3 deg/s 0.5 ... 1 % 100 deg/s	Single-axis stabilisation; Turn indicator; Heading reference for ground vehicle navigation (improved performance required)
RGA Rate Gyro Assembly (3 axes)	3 FOG or 1 FOG triad	0.1 ... 0.3 deg/s 0.5 ... 1 % 100 ... 800 deg/s	Multiple-axes stabilisation; Flight control systems
IMU Inertial Measurement Unit	3 FOG or 1 FOG triad 3 Accelerometers IMU processor	1 ... 10 deg/h 300 ... 500 ppm 400 ... 800 deg/s	Subsystem for guidance and control, and for integrated navigation systems
AHRS Attitude/Heading Reference System	3 FOG or 1 FOG triad 3 Accelerometers 1 Magnetometer System processor	1 ... 3 deg/h 300 ... 500 ppm 400 ... 800 deg/s	Attitude and heading of a vehicle; Magnetic heading reference
GPS/INS Integrated GPS-aided Inertial Navigation System	3 FOG or 1 FOG triad 3 Accelerometers 1 Magnetometer 1 GPS receiver System processor	0.1 ... 1 deg/h 100 ... 300 ppm 400 ... 800 deg/s	Complete position and navigation information; Attitude and heading; Velocity

Table 1.1 FOG devices, subsystems, and systems; the FOG performance requirements are given as bias stability, scale factor error, and measurement range

Marketing opportunities for the single-axis FOG are a straight forward replacement of single-axis mechanical gyro packages; the use in a very special two-dimensional navigation task, which is as heading gyro in navigation systems for unmanned vehicles; as sensor within a classical inertial sensor assembly consisting of three individual gyros and three accelerometers; and very specific measurement tasks, where package flexibility and environmental conditions are not

achieved with classical gyros. FOG devices and systems are listed in Table 1.1 in increasing order of complexity and performance, reaching from a simple rate gyro to inertial grade triads. Drivers for new designs of multiple-axis FOGs are the still high power consumption of the FOG, its still considerable size compared to a pure mechanical gyro, and again cost considerations. They are all together the reason that FOG multiple axes designs are moving in the direction of triad solutions with multiplexed signal processing and common use of optical components.

2. SINGLE-AXIS FIBER-OPTIC GYROS

2.1 Design and Basic Configuration

The configuration of the Alcatel SEL FOGs has been described in Ref./1,2/. The optical part is very similar to the well known minimum configuration of a practical FOG. The readout and signal processing electronics are highly integrated and are implemented as a combined analog and digital ASIC. Block diagrams of the gyro optics and signal processing electronics of the 1991 rate gyro production units are shown in Fig. 2.1.

For completion the main features of the design are briefly summarized and repeated here as follows:

- As light source a low-cost multimode laserdiode is used with an operational wavelength of 820 nm and up to 1.5 mW of optical power launched into the single-mode fiber pigtail. It is mounted in a half butterfly package with thermo-electric cooler, temperature sensor and monitor diode.

- The multifunctional integrated-optics module contains a pigtailed IOC with polarizer, Y-branch and a pair of phase modulators on a z-cut Lithiumniobate substrate with Ti-indiffused waveguides
- Polarization-preserving fiber is used for the pigtails and the fiber coil, which has a length of 100 m and a Sagnac phase shift of approximately 1 μrad for 1 deg/h input rate.

- The detector module consists of a Si-PIN photodiode with a hybrid preamplifier circuitry.

- Square-wave phase modulation at the proper frequency of 1 MHz with a digital demodulation scheme provides the required sensitivity improvement. The demodulation technique with the 8-bit converter and without any analog RF mixing and filtering has already been implemented in 1986 within a first in-house designed ASIC and is described in Ref./1/.

- A new ASIC with analog and digital functions is now used to generate the modulation signal, to demodulate the detector output, and to give improved scale factor performance. Open-loop and a closed-loop operation can be selected. Scale factor calibration as well as temperature error modelling of bias and scale factor is performed by the single-chip micro controller.

- The gyro has an extended built-in test capability, including power-on BIT, initiated BIT, and periodic BIT for monitoring critical parameters during operation. A minimum number of components is used in the mechanization of the optics and the electronics thus improving the reliability of the device.

The decision to operate the present FOGs in the 820 nm wavelength region is based on the fact that the cheapest light sources with the highest reliability out of a mass production are available in this wavelength region. Polarization preserving fiber was chosen for a number of reasons: It gives the cleanest optics concept with the smaller number of splices compared to standard single mode fiber with two depolarizers in the loop. It relieves the requirements for the source incoherence considerably, and it makes at least in first order the gyro insensitive to magnetic fields. The cost considerations are tightly coupled to the length of fiber used. The higher cost of the fiber must be compared to the amount of labour cost for e.g. 4 or 6 additional splices if standard single mode fiber were used.

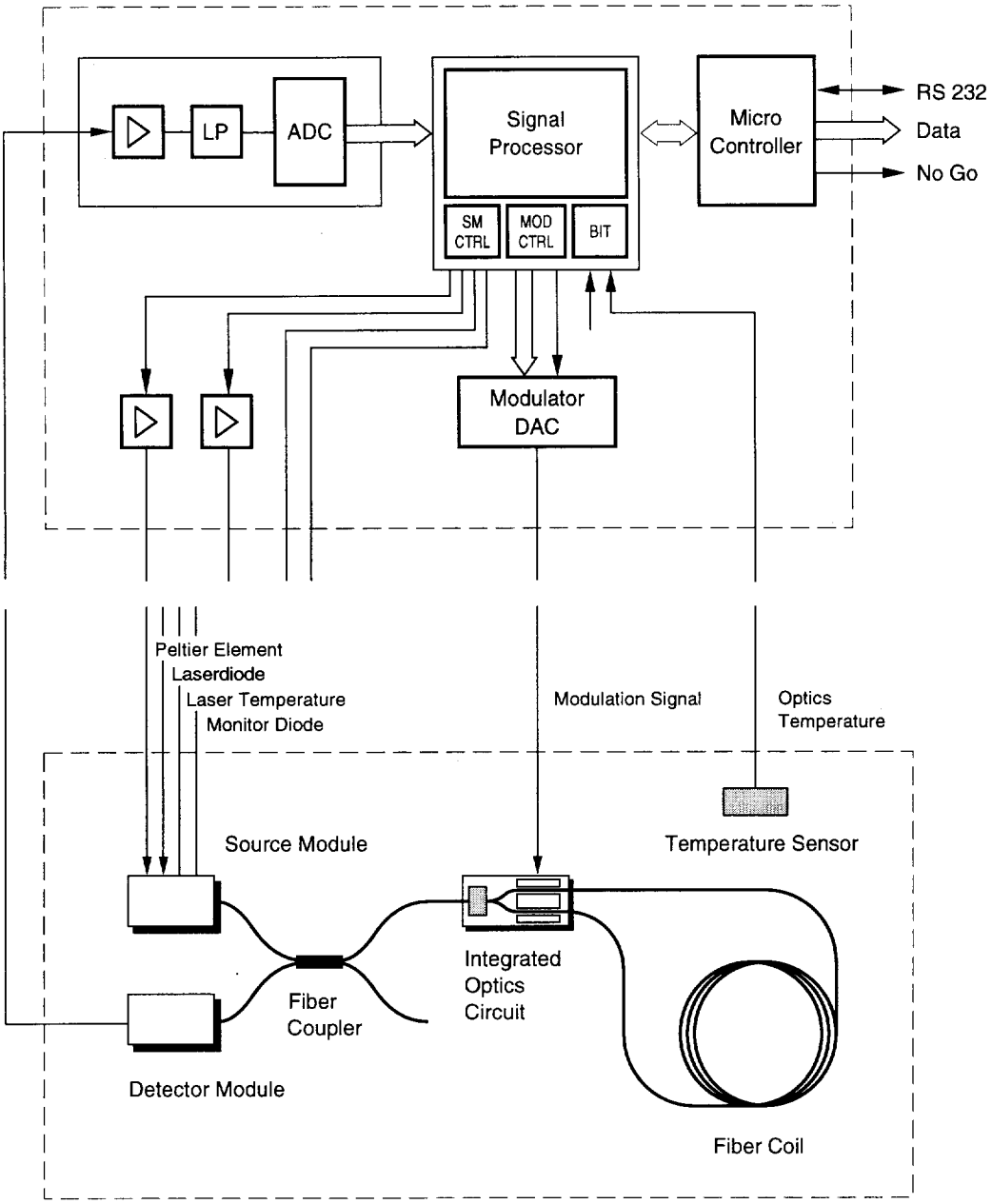


Fig. 2.1 Block diagram of the FOG-P1-A optics (bottom) and control and signal processing electronics (top)

2.2 Packaging and Production Configurations

Today two single-axis FOG product families are available. Both are based on the same concept, are build with the same optical modules and with the same ASIC, hybrid, and micro controller. They have different packages and have been developed in accordance with different customer requirements.

The FOG-1-B is a gyro with separated sensor and signal processing electronics. Its mechanical design is for applications where the sensors are mounted on a sensor block and the electronics boards are sitting on a motherboard in the PCB compartment of an avionics equipment box. The sensor has a cylindrical shape and is connected to the electronics

board with one flexible printed cable. The electronics board is of Standard European size. The performance and the main parameters of this FOG are summarized in Fig 2.2.

The FOG-P1-A is a complete package with sensor, electronics and power conditioning. It has been designed as a form, fit, and function replacement for mechanical rate gyro packages. Therefore it can be delivered with an analog output instead of the regular serial digital rate signal output. The performance of the FOG-P1-A is in its most accurate version similar to the performance of the FOG-1-B. In its standard version it is a 10 ... 50 deg/h gyro with a scale factor accuracy in the range of 500 ... 2000 ppm. The performance levels depend on the degree of calibration and error modelling applied to the units. The less accurate values are achieved without individual calibration of the gyros. This gyro is a fully qualified product. A photograph of the FOG-P1-A is shown in Fig. 2.3.

Performance	
Signal	Rate output
Range	800 deg/s
Scale factor	100 deg/h/LSB
Scale factor repeatability	500 ppm (1σ)
Scale factor nonlinearity	500 ppm (1σ)
Scale factor temperature sensitivity	500 ppm (1σ)
Bias repeatability	3 deg/h (1σ)
Bias temperature sensitivity	3 deg/h (1σ)
Random walk	0,17 deg/$\sqrt{h}$
Signal repetition rate	1 kHz
Dimensions and parameters	
Sensor diameter	80 mm
Sensor height	28 mm
Fiber length	100 m
Coil radius	35 mm
Sensor weight	260 g
Electronics size	4 x 6 inches
Electronics weight	320 g
Operating temperature range	-54 ... + 85 °C

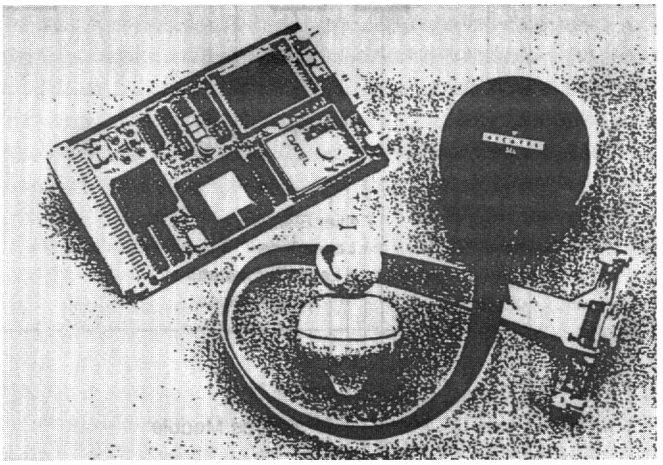


Fig. 2.2 Performance characteristics, parameters, and photograph of the FOG-1-B

2.3 Single-Axis FOG Applications

2.3.1 Retrofit for Mechanical Gyros

Generally this market opportunity is the replacement of single axis mechanical gyros, in most cases of rate gyros. The mechanical gyros can be replaced on two different levels due to the fact that those existing gyros are used either in a gyro package, which could be a single, dual or three axes package, or as a pure sensor mechanically integrated into a subsystem or system, like a gimballed mounting of a seeker camera. Therefore the retrofit can be on package level or on sensor level.

The first approach is easily feasible with today's components and FOG size and their level of miniaturization. Two typical examples are a single axis rate gyro package for avionics applications, widely used in any kind of aircraft motion control, damping, and stabilization in one axis. The second example is a single axis package for gun stabilization in tanks. The size of these existing candidates is approximately 65 x 50 x 115 mm without a sometimes quite large mounting flange. The FOG-P1-A as described above has been designed and developed for this market segment. It has been successfully tested and integrated into tank weapon stabilization systems. Performance requirements have easily been met, performance margins and improvements have been demonstrated. The FOG package has a superior bias and scale factor characteristic compared to its mechanical counterparts. The large FOG bandwidth is beneficial for the stabilization loop design. Since the package has to replace older designs an analog output can be added to the normally digital signal processing and output of the FOG electronics. A breakthrough for the FOG in this market area depends on a further cost reduction on the sensor side and on a change in the users cost evaluation and philosophy: The FOG is a competitive candidate if life cycle cost is the selection criteria and not a case by case cost comparison between the mechanical rate gyro and the FOG on a sensor and spare part procurement level.

The second possibility is to replace the sensor itself. The typical size of the mechanical gyros inside of these packages is a 50 mm long and 25 mm diameter cylinder, weighing less than 140 grams. Due to its implementation and specific mechanical gyro characteristics these gyros need only some drive signal and pick-off signal inputs from the host system and provide an output which can be directly used by the system without further signal processing. Normally it is an AC output which is demodulated within the systems electronics. To approach this retrofit opportunity the FOG must be integrated into this cylinder volume and it has to emulate exactly the input and output signals of the old mechanical gyros despite of its completely different sensing mechanism and signal processing. Studies have been done, and the feasibility to place a FOG optics inside of this volume has been demonstrated together with a volume allocation for the electronics. On the electronics side a significant size reduction and integration would be necessary. This can only be done if a large enough number of identical or at least very similar models of gyros can be identified as replacement candidates.

2.3.2 Ground Vehicle Navigation

A very attractive application for the FOG is in the field of land vehicle or car navigation systems, especially for private cars. The commercial market segment is addressed for example in Ref./3/. A very special type of land vehicle or car navigation is the navigation and positioning of unmanned ground vehicles in a limited short distance two-dimensional environment like a factory floor or a container terminal. Vehicles to be guided are fork lifts, tool carriers, van carriers, trucks etc. Sensors used are a heading gyro, odometers, distance measurement sensors, position markers and means for alignment of the vehicle. The need for gyro compassing to get the north reference is circumvented by other solutions. In some cases where the distances are larger and the vehicles are used outside of a building, GPS is the preferred main solution and the gyro in addition to the odometer is installed only for aiding and to guarantee a hundred percent time and position coverage.

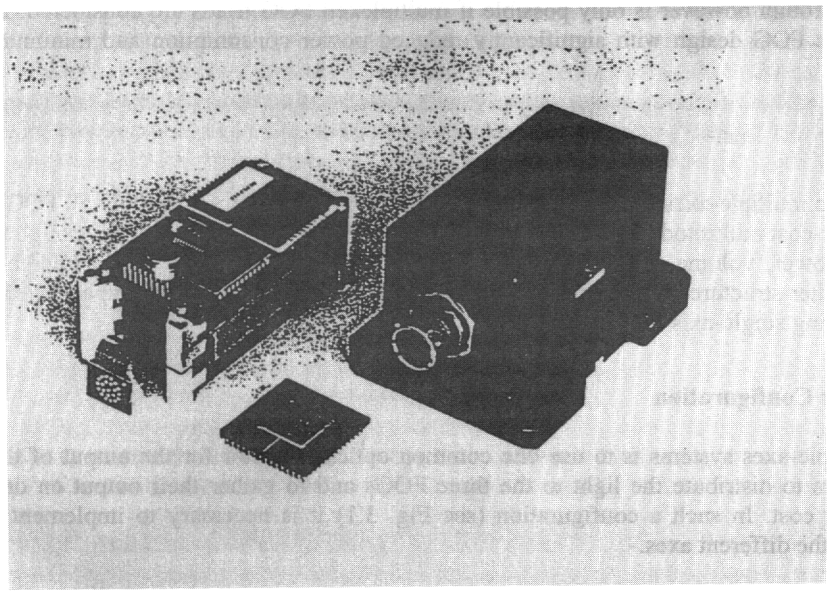


Fig. 2.3 Photograph of the single-axis gyro package FOG-P1-A

In contrast to the rate gyro retrofits the performance requirements for these gyro application are much more demanding. Whereas bias performance requirements are relieved by position updates, zero-velocity updates etc., the scale factor accuracy of the gyro has to be in the order of 300 to 500 ppm due to 90 deg or 180 deg turns performed by the vehicles before updates take place. On the other hand these systems are new and modern and therefore the digital interface of the FOG, its BIT capabilities and its simple error mechanisms and error models are well acknowledged by the system designers, meaning that today's FOG cost/performance numbers are very competitive.

2.3.3 Inertial Sensor Assemblies for AHRS, INS, and Derivatives

The classical approach in an AHRS or INS is an inertial sensor assembly (ISA) with individually replaceable sensors. Since the FOG as a stand-alone sensor still has a volume of 8 to 20 cubic inches, in addition requiring a signal processing electronics of a similar volume, its use as replaceable unit in an ISA is today limited to those applications where size and weight are not the main driving factors, and where on the other side FOG specific characteristics are beneficial for the overall systems design. These could be its fast response time, its bandwidth, or unique environmental scenarios like high shock or acceleration loads to the instruments. Another reason is simply the availability of a tactical grade 1 ... 3 deg/h, 300 ... 500 ppm, 1000 deg/s gyro at a competitive cost level. Systems available and based on SFIM Industries FOGs are a vehicle motion measurement system, an AHRS, and an integrated hybrid navigation system.

3. MULTIPLE-AXIS FIBER-OPTIC GYROS

3.1 Design Considerations

Complete Inertial Measurement Units (IMU) to determine motion of a vehicle in three-dimensional space have as a minimum without redundancy six sensors, normally three linear accelerometers and three single axis gyros. Driving design parameters in the field of sensor packages for avionics, flight control, and missile guidance are volume, power consumption, and also weight. Especially small size is often more important than an individual performance parameter. Due to the modular structure of the FOG and especially as a result of the flexibility in design, shape, and size of the sensing coil, the size of the optical part of the FOG meets almost all requirements known for today's rate gyro packages and IMUs. An example was given in Section 2.3.1.

In contrast to the small FOG optics is the electronics for a single axis FOG. Depending on performance, ASIC or hybrid implementation, available supply voltages, and interface features, it may have a size up to 100 x 150 mm printed circuit board. Power consumption of a complete FOG with a light source containing a thermo-electric cooler is temperature dependent and is in the range of 5 ... 12 W with the highest power consumption at the temperature extremes. Therefore new IMU designs are aimed at a reduction of size and power dissipation of the electronics rather than smaller optical modules. A size reduction can be obtained without changing the proven readout and signal processing schemes by higher integration of the electronics, e.g. digital and analog ASICs, three channel hybrids, external power conditioning, and tailoring for specific applications. A breakthrough however is only possible if multiplexed FOG triads are considered. In order to achieve an optimized multiple-axes FOG design with significantly reduced power consumption and minimum number of parts several steps are possible:

3.2 Single-Source Configuration

The first and most obvious step to optimize a multiple-axes system is to use one common light source to supply all FOGs. Light sources available today (especially low-cost multimode laserdiodes with 820 nm wavelength) are powerful enough to supply three FOGs in parallel. Electrical power, volume, and cost of two light sources can be saved with a tolerable reduction of the optical signal level. If all other structures of the axes are left unchanged, there is nearly no impact on the already existing configuration of the underlying single-axis FOG.

3.3 Single-Source and Single-Detector Configuration

The next step for an improvement of multiple-axes systems is to use one common optical detector for the output of the interferometers. The optical (3 x 3)-couplers to distribute the light to the three FOGs and to gather their output on one detector are readily available at reasonable cost. In such a configuration (see Fig. 3.1) it is necessary to implement a method to separate the signals coming from the different axes.

3.3.1 Frequency Multiplexing

By applying different modulation frequencies to the different axes the output signals are also at different frequencies and can be separated by bandpath filters. An advantage of this method is the in principle continuous operation of all axes in parallel. Its disadvantage is the high suppression required to avoid crosstalk between the continuously active modulation and detection circuits. At least three independent modulator circuits are required in the electronics; the length of the fiber coils must be different in order to match the loop transit times to the modulation frequencies.

3.3.2 Time Multiplexing

A straight forward approach is the time multiplexed operation of the different FOG triad axes. The signal processing electronic circuits of a single-axis FOG can be used for a multiple-axes system if the individual axes can be switched on and off in a controlled manner. Only simple digital logic is then required to direct the result of the signal processing to an individual output channel. Due to the very fast transfer characteristics of the FOG the switching can be done at a sufficiently high repetition rate without loss of information, e.g. in a vibrating environment. In general two types of switching mechanisms are possible:

- Optical Switching
uses active optical switches in the light path to direct light into only one axis of the triad. Complex integrated optics circuits are necessary to achieve high performance switching characteristics.

• Modulation Switching
uses the fact that the single-axis FOG described above provides rate output only then when an electrical modulation signal is applied to the phase modulator in the light path of the interferometer. Therefore simple switching of an electrical signal can be used to activate only one axis at a time. In a consistent design the same electrical functions are used for all axes in a time sequence, which on the other hand avoids any crosstalk between the axes as only the functions for one axis are active at the same time. With the exception of the electrical switches, which are readily available, and a small overhead of digital logic to control the switching, the same optical configuration and the same proven signal processing scheme can be used as in the single-axis FOG. No special optical circuit and no new signal processing scheme have to be developed. For these reasons the modulation-switched time multiplex technique has been chosen by SFIM Industries to be the preferred solution for a compact multiple-axes system.

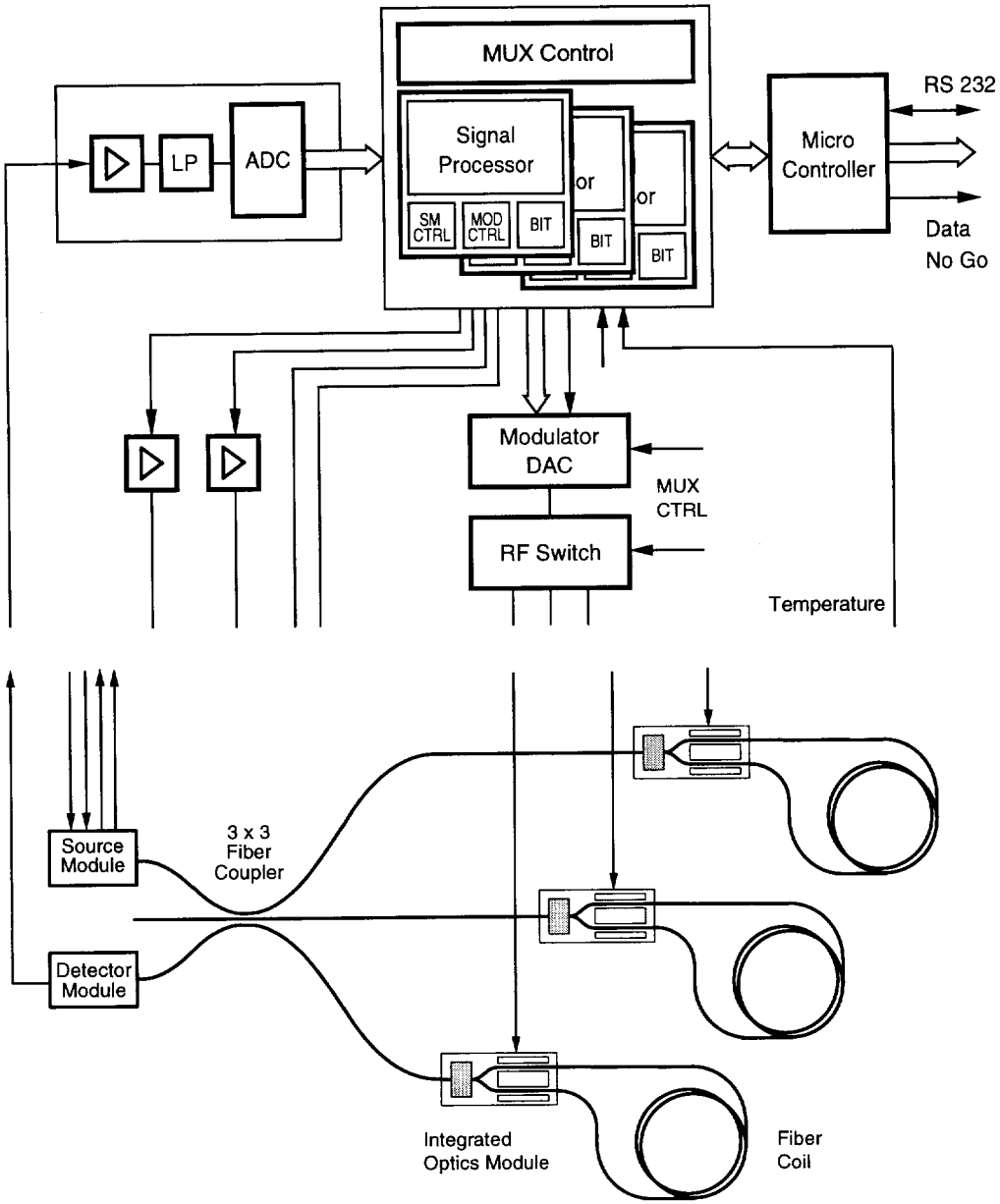


Fig. 3.1 Time multiplex FOG triad block diagrams of the optics (bottom) and electronics (top)

4. FOG BASED INERTIAL MEASUREMENT UNIT

4.1 Design

The core of the Inertial Measurement Unit FOG-I-3A/B is the time-multiplexed FOG triad. In order to make it an IMU it is completed by three single-axis linear accelerometers, an ASIC for the readout of the accelerometers, and an additional micro-controller to perform IMU related functions like temperature modelling and axes misalignment compensation. The Inertial Sensor Assembly (ISA) and two purely digital electronics boards are mounted into a single box. The box has mounting provisions for an optional interface board. The IMU can be operated with different supply voltages and provides compensated delta-angle and delta-velocity outputs through a fast serial or parallel interface. A top-level block diagram of the IMU func-tions is shown in Fig. 4.1.

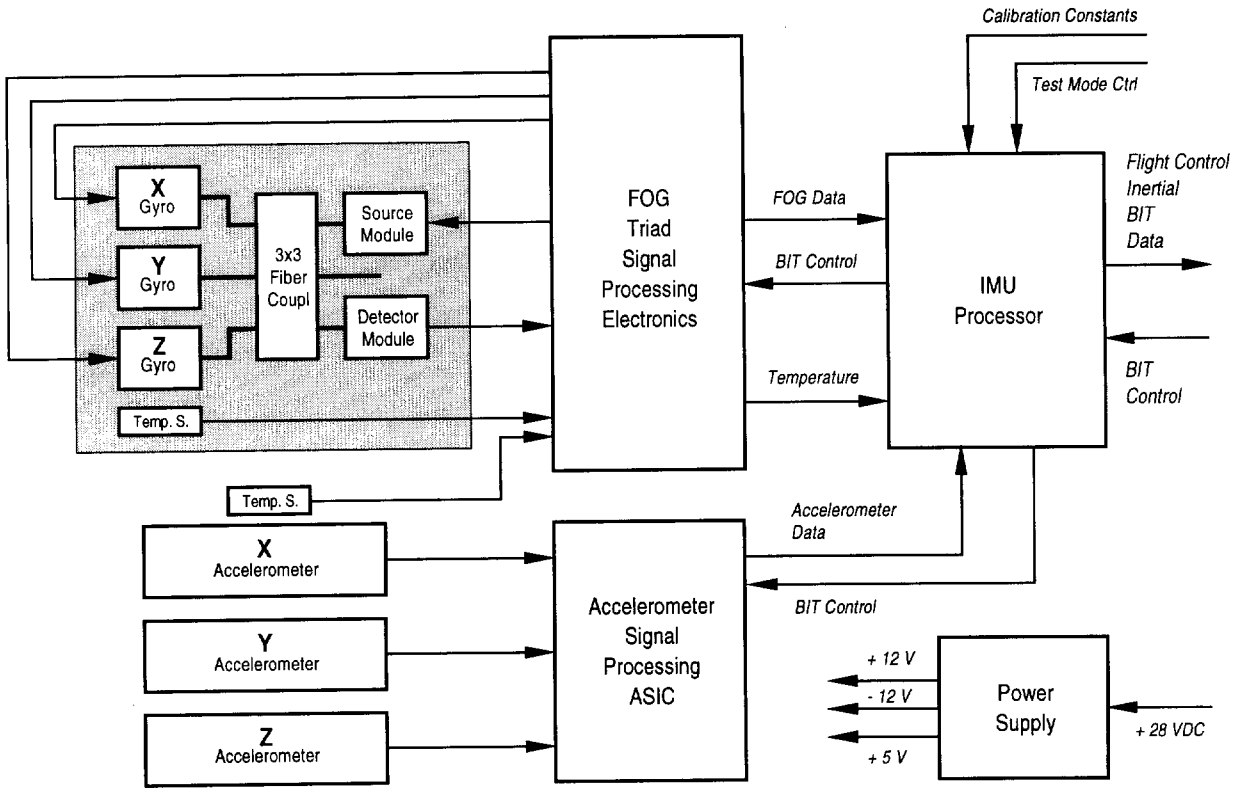


Fig. 4.1 Block diagram of a FOG based Inertial Measurement Unit (IMU)

4.2 Hardware description

4.2.1 Optics

The modules used in the FOG triad optics are identical to the modules in the single-axis FOGs. These are a source module, a detector module, a fused 3 x 3 fiber coupler, three IOMs and three fiber coils. The shape of the coils is adapted to the overall mechanical design of the sensor block. Due to the very low failure rate of FOG fiber coils they are potted into the block structure. The design is such that the block can accommodate either 100 m or 300 m coils, thus giving some flexibility for the cost/performance trade-off. Two electronics modules are attached to a frame holding the block in order to make the electrical interface between the sensor assembly and the processing electronics less critical with respect to shielding, grounding, crosstalk etc. These are the two analog subassemblies containing the functions of the AD conversion and the DA conversion with modulation signal switching. In addition to these FOG modules the block carries three single

axis accelerometers and temperature sensors. It has mounting provisions for vibration isolators. A photograph of these IMU is shown in Fig. 4.2. The size of the box is 155 x 118.5 x 115 mm^3.

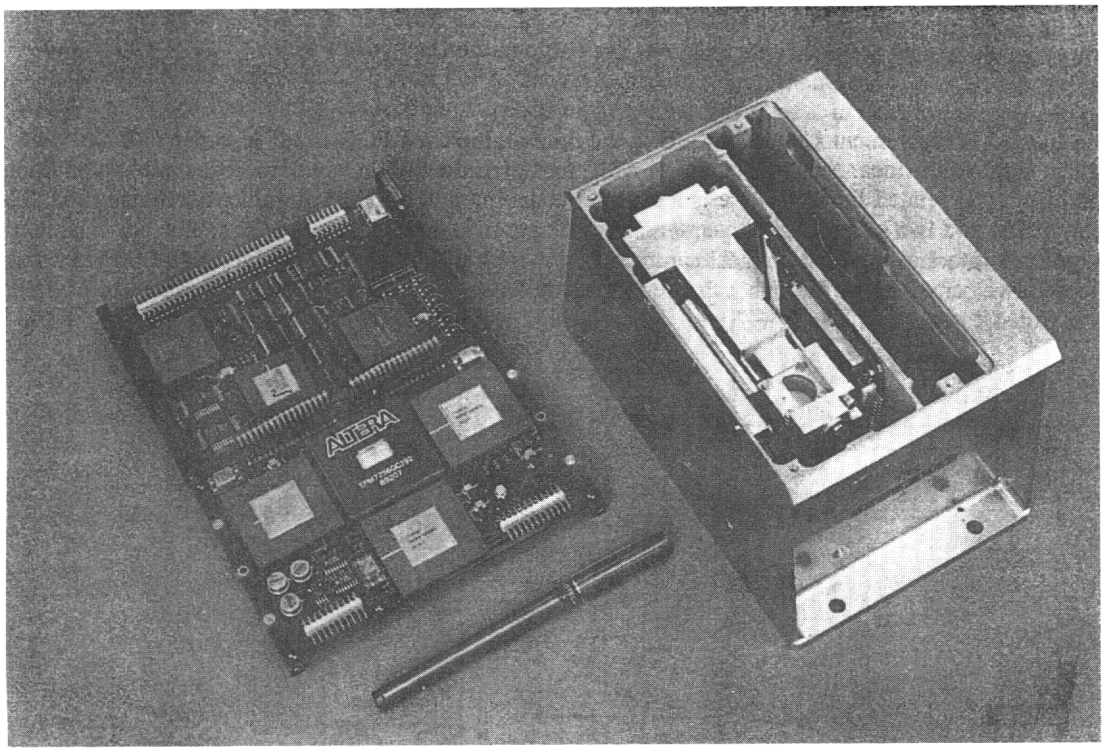


Fig. 4.2 Photograph of the FOG-I3-B Inertial Measurement Unit. The two digital electronics boards are shown outside of the box.

4.2.2 Electronics

The electronics is split into analog pre-processing which is the AD conversion, DA conversion plus switching, and two purely digital boards. The first implementation of the multiplexed FOG triad is based on the existing single-axis FOG ASIC. Therefore three ASICs are used which are completed by a programmable ASIC for the multiplex overhead. The latter distributes the AD flash converter output synchronously to the switching of the modulation signal into the three ASICs. These four ASIC parts are located on the first PCB. A second PCB carries two micro controllers together with the accelerometer readout ASIC and some peripheral circuitry. The tasks of the FOG micro controller are hardware related, mainly to initialise and control the FOG ASICs. The so-called IMU processor is mainly for accelerometer readout and sensor error compensation. The IMU can be delivered with a multifunctional interface board including DC/DC converters and serial or parallel output.

CONCLUSION

Today SFIM Industries Deutschland GmbH is providing a familiy of mature single axis and multiple axes FOG products, as well as FOG based Inertial Measurement Units. They operate in the drift performance range of 1 ... 10 deg/h. Devices with higher performance are described in the contribution by H.C. Lefevre within this volume.

REFERENCES

1. W. Auch, M. Oswald, D. Ruppert "Product development of a fiber optic rate gyro"
   Proceedings ***Symposium Gyro Technology***, Ed.: H. Sorg, Deutsche Gesellschaft für Ortung und Navigation, Stuttgart, 1987,

2. W. Auch, M. Oswald, R. Regener "Fiber Optic Gyro Productization at Alcatel SEL"
   Proceedings ***Fiber Optic Gyros: 15th Anniversary Conference***, Ed.: S. Ezekiel, E. Udd; SPIE Vol. 1585, Boston,MA, 1991,

3. Y. Nishi, T. Iwashita, Y. Nishiura, K. Washimi, K. Okamoto, and A. Oaka "Single Mode Fiber Based Fiber Optic Gyroscope for Automobile Navigation System"
   Proceedings ***9th Optical Fiber Sensors Conference***, Firenze, 1993

11. Progress in Fiber-Optic Gyroscope Applications II with Emphasis on the Theory of Depolarized Gyros

G. A. Sanders and B. Szafraniec*

September 9, 1998

ABSTRACT

Fiber-optic gyroscopes (FOGs) are under development as the next generation inertial sensor and are showing considerable potential toward meeting even the most stringent of applications. We report promising results of efforts aimed at inertial-navigation quality (0.01°/hr) using a low cost depolarized FOG design. This design using a single-mode (SM) fiber sensing coil presents unique challenges which are the focus of a detailed analysis. Both a Jones matrix formulation and a Poincaré sphere description are used. We also present the progress of efforts on precision (0.001°/hr) FOG designs as well as update the status of a medium grade open-loop product for attitude heading reference system (AHRS) applications.

1 Introduction

Fiber-optic gyros[1] are undergoing product development throughout the world as the next significant inertial technology. These all-solid-state devices offer a combination of long life and high reliability along with the potential of low mass-production cost. Most importantly, their performance can be readily scaled to fulfill a broad range of applications.

The primary focus of this chapter is the medium to ultra-high performance range of FOG designs and applications. In particular, we start with a brief update of the progress of the Honeywell medium performance open-loop FOG. This development represented a significant step for FOG technology, as it was one of the earliest FOG products. This is covered in section 2.

* Authors are with the Honeywell Technology Center, Honeywell Inc., Phoenix, AZ USA.

In addition, higher performance product developments are showing excellent progress toward meeting their applications. One area of great interest is that of inertial navigation ($< 0.01°/\text{hr}$ bias stability) for many applications including aircraft, land vehicles and spacecraft. Because this market is cost-sensitive, we have[2] undertaken the challenge of trying to improve the performance of the previously lower performance (but low cost) depolarized designs to that of navigation quality. These designs use a single mode fiber sensing coil. Performance results for our depolarized FOGs are reported in section 3. They establish navigation-grade performance feasibility, and show encouraging progress toward an affordable inertial grade FOG product.

There are several interesting problems related to the design and the mitigation of error mechanisms in the depolarized design. Because depolarized gyros are significantly more complex than FOGs using the more expensive polarization maintaining (PM) fiber, special attention needs to be paid to a number of issues. These include the issues of signal fading, wavelength stability, time-varying non-reciprocity, magnetic sensitivity, and polarization errors. All of the issues are in some way related to the complex behavior of polarization in depolarized gyros. While brute force numerical treatments of the depolarized gyro may be adequate for some purposes, such an approach tends to overcome the insight of the behavior of light in this gyro design. Because of that issue, we have described and contrasted this optical behavior from the point of view of the Poincaré sphere, as well as the Jones matrix representation. The blend of the two techniques is intended to provide both a mathematical model and an intuitive interpretation of the model. The discussion of the above topics is in section 5, and offers theoretical support for the performance results presented in section 3.

We have also been pursuing the PM FOG technology for precision applications such as submarine navigation and spacecraft control. Section 4 contains a brief discussion of our progress toward satisfying requirements in the realm of $0.001°/\text{hr}$ bias stability or better.

2 Medium Performance FOG Product Status

Several developers have been pursuing FOGs for the medium performance range having stabilities in the $0.1°/\text{hr}$ to $5°/\text{hr}$ range. These applications have included tactical missiles,[3–5] spacecraft, trains, autonomously guided vehicles,[7] and marine gyrocompasses.[8] One of the most mature applications today is that of commercial aircraft attitude heading reference systems. In this application the sensor is only responsible for providing vehicle direction information, and hence, the scale factor requirements are modest (approx. 1000 ppm). This has enabled the use of an open loop design in the first commercially available FOG product, shown in Figure 1, developed by Honeywell. The design[9] uses a standard all-fiber minimum reciprocal configuration. The required bias modulation is effected by stretching a length of fiber wrapped around a piezoelectric cylinder. The output after demodulation is a signal representative of the rotation rate. The first application for this device was for the Dornier 328, 33–passenger commuter aircraft.

Another application of this design was for the Boeing 777 airplane in the Standby Attitude

and Air Data System (SAARU). The SAARU provides attitude and heading information about the 777 in the event that the primary navigation system fails. The SAARU system was certified in the first half of 1995 and has been in revenue service since mid-1995. By early 1998 over 3000 units have been produced for all applications combined, and the design successfully meets the performance requirements of its intended applications.

To get one perspective of the rate of progress of the FOG technology, it is instructive to look at the success and improvements over the first six years of this product. To do this, we collected a set of statistics comparing the performance of 18 pre-production units from 1991 with a sample of 100 production units from 1997. The bias and scale factor repeatability over time due to all causes, including in-run stability, turn-on to turn-on repeatability, run-to-run repeatability and any variations over time caused by environmental stress testing are shown in Figure 2. The mean bias repeatability, illustrated in Figure 2a, of the pre-production sample was 17.4°/hr, which was beyond the 3-σ specification of 4°/hr for this application. By careful attention to design processes, angular control of splices, and signal processing designs, the average bias stability has improved to 0.15°/hr. The standard deviation of the distribution has improved by five-fold, indicating significant progress in process control. The scale factor performance history, illustrated in Figure 2b, shows similar improvement. The scale factor repeatability is now 420 ppm. This is particularly notable in an open-loop system since the scale factor stability is influenced by electronic and optical gains. This performance result is surprisingly comparable to moderate-grade closed-loop FOG performance. This example clearly illustrates the maturation of the technology and that with careful attention to process controls, FOG designs can be successfully produced.

3 Navigation Grade FOG Status

Inertial navigation systems such as for aircraft navigation typically operate by determining the vehicle position and initial heading of the vehicle relative to the earth's axis, and then integrating linear acceleration and rotation throughout the flight. The rotation information from the gyro must typically be stable to better than 0.006°/hr. Requirements are comparable for many land navigation and missile applications. In many of these applications, especially those in highly maneuverable vehicles like missiles, the scale factor requirements can be as stringent as 10 ppm over a rotation range of up to 1000°/s.

The navigation problem becomes more challenging since this performance level must be met during exposure to harsh physical environments; namely, 6-9 g rms random vibration (typical), −54°C to +71°C vehicle temperature range, 10 Gauss magnetic fields, and 0.5°C/min to 1.0°C/min temperature transients (at sensor location). Furthermore, competitive issues with other technologies and requirements for smaller, lighter systems are forcing the sensor size requirements down, diminishing inherent rotation sensitivity.

Despite these challenges the FOG technology is making excellent progress toward meeting

the intended requirements. Since the scale factor requirements are tight, a closed loop system is required. The most common high performance FOG architecture[10] is shown in Figure 3, with most developers using a polarization maintaining PM fiber sensing coil type.

In this design a multi-function integrated optics chip (IOC) is used for splitting the light into clockwise and counterclockwise waves, and for electro-optically imparting a phase modulation to the light waves in the loop. The signal processing design is based on conversion of the photodetector signal to a digital representation of the detected light intensity, followed by digital demodulation and integration. The loop is closed by driving the integrated optical phase modulator with a voltage ramp whose slope is proportional to the sensed rotation rate. This ramp imparts an optical phase shift between light waves to restore the sensor to its non-rotating condition. Three primary phase ramp loop closure methods have been developed throughout the industry. These are the digital phase step technique,[11] the serrodyne phase ramp technique[12] and the dual ramp technique.[13]

Several developers have reported navigation-grade, or near navigation grade bias stability in ambient room temperature environments using a PM fiber sensing coil. In 1996, Cordova *et al.* have in fact reported navigation-quality performance[14] (using a PM fiber coil of less than 3 inches outside diameter) at 0.0081°/hr bias stability over a temperature profile comprising a −54°C to +71°C temperature range and ramp rates of 0.5°C/min to 1.0°C/min. More recently, we have reported, for the first time, similar results in a FOG using a single mode fiber coil. These results are described later in this section.

The PM fiber technology has clearly received the greatest attention, up to this point, for high performance applications. While considerable progress has been demonstrated using the PM fiber technology, attainment of navigation grade performance at a sensor cost that is sufficiently attractive to displace existing sensors is a significant issue. While we believe that automation technologies are quite important, it is also essential to reduce the inherent cost of the FOG design. One of the most critical areas where cost reduction is necessary to realize an attractive product cost is that of the sensing coil itself. Honeywell[15] has therefore been pursuing the use of standard single-mode coil fiber in conjunction with polarization randomization in the loop, i.e., the depolarized FOG technology.[16,17] Figure 3 shows the contrast between the navigation-grade closed-loop PM and SM designs. In the PM design, the fiber maintains the light polarization along one of its axes throughout the coil, whereas depolarizers on either end of the SM loop (see section 5.4) randomize the light polarization in the SM coil case. The depolarized technology is particularly attractive for the higher performance FOGs, where large amounts of PM fiber are used. For instance, in a navigation grade gyro of 0.005°/hr. which typically uses 1000 m of fiber, the cost of the coil is expected to be the most dominant cost component of the gyro. In order to reach navigation-grade performance in a depolarized FOG, however, several performance barriers have had to be overcome. Some of these, such as signal processing issues and the coil sensitivity to thermal transient gradients, are common to both FOG types. Others are either unique to the depolarized FOG or more extensive in the depolarized FOG.

Perhaps the most important performance issue which is more pronounced in a depolarized

gyro is that of polarization-induced errors[18] which degrade the bias stability. These errors are treated in section 5.6. This well known polarization-related errors in the FOG are caused by light that is cross-coupled to the wrong polarization state prior to the IOC polarizer, travels in the wrong state in part of the interferometer loop, and then cross-couples back to interfere with the gyro signal light. It erroneously appears as a rotate rate. Since a depolarized gyro uses one or more depolarizers purposely containing 45° splices, this effect is 10–30 times larger in a PM FOG. Recently, however, this issue and some others (see section 5) have been overcome. Navigation-grade performance levels comparable to PM gyros have been obtained.[18] An illustration of navigation grade performance is shown in Figure 4, where the rate uncertainty of an open loop depolarized gyro is plotted versus integration time to τ for ambient temperature conditions. The coil length was 1 km and the coil diameter was nominally 3 inches. As is shown in the plot, the rate uncertainty is proportional to $\tau^{-1/2}$ for short integration times indicating that the gyro is limited by random noise. It does not depart from this dependence until it gets to the 0.0035°/hr level, which is representative of its bias stability.

More recently we have built depolarized FOGs with a sensing coil diameters smaller than 2.2 inches and tested them over rugged temperature environments. The applied thermal profile of the temperature chamber (approx. −50°C to +70°C), the resultant coil-surface temperature profile (approx. −40°C to +70°C, with a range of ramp rates up to 0.75°C/min), and the compensated bias stability are plotted in Figure 5a. The bias residuals after compensation, which include terms for temperature and temperature rate, were deduced to be around 0.018°/hr using the Allan variance (cluster) analysis technique similar to that in Figure 4. Figure 5b shows the compensated output response of another depolarized FOG prototype to a comparable thermal profile. As shown in Figure 5b, very little residual thermal sensitivity is observed. The resulting bias stability after compensation was better than 0.011°/hr. These results show excellent progress toward navigation grade performance in a very small coil size and are consistent with the best reported results for PM designs.

Another obvious concern in depolarized FOGs is the sensitivity to magnetic fields via the Faraday effect. This is treated analytically in sections 5.2.2 and 5.4. In PM FOGs the state of polarization (SOP) is very nearly linear, and Faraday sensitivity is only a consequence of residual ellipticity[19] in the SOP. In depolarized FOGs, light can propagate in all SOPs, including elliptical and circular SOPs, which are sensitive to the magneto-optic Faraday effect. As is shown in later sections, this magnetic sensitivity can be overcome by sufficiently randomizing the light polarization entering the sensing loop.[20]

The experimental data shows that it is possible to reach navigation grade bias stability in a two-depolarizer depolarized FOG architecture with the aid of standard μ-metal shielding. The requirements are stringent, since typically the gyro bias must vary only a fraction of 0.005°/hr in a fluctuating magnetic field of up to 10 Gauss along any axis. Since the evolution of the SOP in a depolarized FOG is quasi-random over time, the magnetic sensitivity even in a constant field, will fluctuate from some minimum to maximum value. Hence, it is important that the maximum sensitivity be less than 0.0005°/hr/Gauss.

Since the gyro random noise and bias errors are quite appreciable compared to the required magnetic field sensitivity, a synchronous AC measurement technique was used to measure the gyro response. In this technique the applied magnetic field was modulated at 0.05 Hz to differentiate magnetic sensitivities from other error mechanisms. Since the response of the gyro was small, a 10 Gauss field amplitude was used and the output of the gyro was demodulated at 0.05 Hz. The gyro sensitivity is plotted vs. time in Figure 6. The gyro employed the basic depolarized architecture of Figure 3, and used a 1000 m SM fiber coil of less than 2.2 inch outside diameter. The results show varying sensitivity over the 8 hr time span as expected, with fields along the x, y, and z (rotation) axes giving bias sensitivities of -0.00005 ± 0.00016, -0.00007 ± 0.00017, and 0.00016 ± 0.00013 measured in °/hr/Gauss. These clearly meet navigation grade performance requirements and compare favorably with comparable PM FOG results.[14]

The results shown in Figure 4, 5, and 6 mark a significant milestone in FOG development since they illustrate sufficient understanding of FOG error mechanisms to achieve navigation grade performance using single-mode fiber. This improves the prospects for low cost while still meeting performance objectives. Much of this progress is common to that of the PM gyro such as the development of the basic reciprocal design, winding techniques and thermal transient mitigation, and signal processing techniques. However, as discussed earlier, there are several issues that are unique to depolarized gyros. These issues affect the random noise, bias stability, magnetic sensitivity and scale factor performance of the instrument. Since the SOP in a single mode fiber coil varies freely due to internally-and externally-induced stresses, several questions arise in the study of this design. What is the best way to model the behavior of the SOP in the gyro, giving both quantitatively correct results without sparing intuitive insight? How do optical imperfections affect performance? How "depolarized" does the light have to be? These are questions that we attempt to answer in section 5. In the end, we believe that if one approaches the design correctly, the random noise will be slightly higher (but adequate for this application) than in PM gyros, but bias stability and magnetic sensitivities may be equal or even smaller. The latter statement seems ironic, but stems from the fact that there are several error-averaging processes that are available to the designer. The benefits of these processes are analyzed and illustrated in section 5.

4 Precision FOG Development Status

FOGs are also undergoing development for very high precision applications such as for precision spacecraft and submarine usage.[2,22] These applications typically have the most stringent of all accuracy requirements (better than 0.001°/hr) but reside in relatively benign physical environments. In both cases, the FOG technology offers the advantages of very low angle random walk (ARW), very long life, high reliability, and no mechanical dither. In the spacecraft application, it also offers the ability to achieve almost arbitrarily high resolution, represented by the smallest angle for which an instantaneous indication of rotation is provided by the gyro.

A schematic of a typical precision FOG is shown in Figure 7. A superfluorescent erbium-

doped fiber light source (FLS) is used to produce broadband light in the 1.53μm to 1.56μm wavelength region. The fiber light source is well suited for this application. First, because the source wavelength is primarily dependent on the erbium ion transition characteristics and it has one to two orders of magnitude lower sensitivity to temperature that of semiconductor sources.[23,24] Hence, the gyro scale factor stability can be in sub-ppm range. Secondly, the fiber light source offers high optical power (greater than10mW) which is needed to reduce the ARW.

While high optical power is necessary for reducing the shot noise component of the random noise, it is not sufficient for achieving ultra-low ARW. As pointed out by Morkel,[25] beating between independent oscillators in the broadband optical spectrum contributes excess noise which increasingly dominates the gyro signal to noise ratio for higher detected power levels. This is especially true for the use of erbium-based FLS designs since their spectral widths are typically narrower that those of most SLDs and LEDs. One method of compensating the gyro output for excess noise, typically employed in a precision FOG, is direct detection of the noise on the light source intensity with a second photodetector. The output of the second photodetector is subsequently subtracted from the noise on the gyro signal.[26]

Efforts are ongoing at Honeywell to develop[2,22] precision FOGs for both space pointing and submarine navigation applications. In efforts by Sanders *et al.*, a gyro employing a 4 km, 5.5 inch diameter coil of PM fiber and an Er-doped fiber source was used. A closed-loop bias stability test taken under temperature-stabilized conditions (permissible for this application) was conducted. The gyro output data illustrated in Figure 8, shows performance consistent with the intended application. The Allen variance analysis similar to that in Figure 4, shows that the angle random walk (ARW) was around $0.00019°/\sqrt{\text{hr}}$ with a bias stability of better than 0.0003°/hr. This bias performance, amongst the best closed-loop performance published to date, is equivalent to the rotation rate of less than one revolution per century and a half!

5 Analysis of Depolarized Gyro

There are many commonalities between PM gyros and depolarized gyros. In this section we will focus on issues which are unique to depolarized technology. All of these issues are in some way linked to the complex behavior of polarization. In PM gyros light travels through the entire interferometer in one polarization state, aligned to a principal axis of the fiber. Some light does couple to the wrong polarization state, by means of polarization cross-coupling at the splices or over a long distance of PM fiber in the sensing coil, but it is removed by the polarizer. This makes the description of polarization behavior in PM gyroscopes relatively simple. On the other hand, polarization analysis of the depolarized gyro is far more complex as the polarization state is not controlled and it may change anywhere within the single mode sensing coil depending on its birefringence. Birefringence is present in a SM coil because of inherent imperfections in the SM fiber and because of internally or externally induced stresses, e.g.: fiber bending or twisting. The character and magnitude of birefringence in SM coils varies with environment; for example, fiber within the coil structure may develop larger linear birefringence as the temperature changes and

stresses within the coil increase. A fact which is very real but often overlooked, is that changes in the birefringence are fast enough to introduce phase errors between counter propagating waves (i.e. "time varying non-reciprocity"). The problem of time-varying birefringence is similar to the Faraday effect since it is a non-reciprocal effect. The difference is that unlike the Faraday effect it may involve any polarization state, rather than only circular states. Of course, the Faraday effect itself is a potential issue as well, since among the polarization states which exist within the sensing coil, there are circular polarization states which make the gyro sensitive to the non-reciprocal circular birefringence induced by a magnetic field. Issues of non-reciprocal birefringence are described in sections 5.2.2 and 5.4.

In depolarized FOGs, unlike in PM FOGs, individual wavelengths comprising the broadband spectrum experience different polarization evolution as light propagates through the interferometer. This polarization evolution results in a different attenuation for different wavelengths when gyro polarizer is encountered and it may even lead to a total extinction of some wavelengths. As a consequence, the original spectrum of the light source is altered. This brings up several issues: gyro light loss or signal fading, wavelength instabilities and increased relative intensity noise of the new spectrum. These issues are treated in sections 5.3 and 5.5. Lastly, in section 5.6 special attention is paid to the classical polarization errors which are considerably more complex in depolarized gyros than in PM gyros.

Because of the importance of polarization and birefringence we start (section 5.1) by reviewing and noting useful aspects of the Poincaré sphere and Jones matrices for analyzing depolarized FOGs. Jones matrices provide a mathematically rigorous method of tracking polarization which is essential, however, they often lead to a less intuitive description. Despite the mathematical complexity of some phenomena observed in depolarized gyros, the Poincaré sphere provides elegant and simple geometrical interpretations. The two techniques are combined in sections 5.2.2 and 5.4 and offer both a mathematical representation and intuitive interpretation of non-reciprocal polarization effects in a depolarized gyro.

5.1 Description of Polarization and Birefringence

There are several methods of describing the polarization state of light in an optical system. Jones vectors and matrices are mathematically very simple; however, the pair of complex numbers in Jones vector is sometimes rather difficult to connect with a specific polarization state. On the other hand, the location of the point on the Poincaré sphere can be connected with a specific polarization state immediately. These two methods of describing polarization behavior are of course equivalent and any operation performed using one method can be represented by the other method. A brief review of both techniques for purposes of analyzing depolarized gyros is presented in this section.

The Poincaré sphere is shown in Figure9. Points located on the equator represent linear polarization states. The horizontal state, H, and vertical state, V, are on the opposite sides of

the sphere, as orthogonal polarization states. The same applies to $+45°$ and $-45°$ linear states of polarizations (SOP). Right circular, R, and left circular, L, polarization states are located on the poles. Points located between equator and the poles describe elliptical SOPs, with the ellipticity increasing with closer proximity to the poles. All points of the same longitude correspond to polarization states with the same azimuth, all points of the same latitude correspond to SOPs with the same ellipticity. To convert for example, the polarization state H into an arbitrary polarization state P_1, we recall two important properties of Poincaré sphere:

1. Any birefringent medium has two orthogonal eigenmodes, i.e. principal polarization states, that are preserved while propagating through the medium.[27–29] These orthogonal states are located on the opposite sides of the sphere and form an axis; for example, a linearly birefringent medium like an ideal PM fiber has eigenmodes modes H and V which form the axis H–V. Light propagating through this medium (in SOPs other than H and V) changes its polarization state and these changes map on a circle on the Poincaré sphere whose center is located on the axis established by the eigenmodes, in this case, axis H–V. This operation is represented as circle A in Figure 9. and describes the property of optical retardation.

2. If light propagates from one medium to another whose principal axes of birefringence are misaligned from the first medium by an angle γ, then we rotate the polarization state on the sphere east or west by twice the angle of misalignment 2γ before propagating light through the second medium as described in (1). For light propagating in polarization state H in the first medium, this operation corresponds to the rotation from state H to state P_0 on the equator in Figure 9. This operation of rotation is used to describe splices in gyro circuit.

By using retardation and rotation we can convert any polarization state into a new polarization state. Consider light of linear polarization H, propagating from the polarizing waveguide on integrated optics chip into a PM fiber whose axes of birefringence are misaligned by an angle γ_1 from those of the IOC. First, we rotate H on the sphere along the equator by the angle $2\gamma_1$ to the point P_0. This represents H with respect to the new principal axes of PM fiber. Then, to describe the propagation of light through linearly birefringent fiber whose retardation angle is equal to η_1, we rotate the state of polarization from point P_0 to point P_1 along circle A centered on axis H–V. These two operations are illustrated in Figure 9. Every polarization on the Poincaré sphere can be conveniently defined in terms of angles γ and η which correspond directly to the construction of the interferometer circuit described in the next section[1*]. The normalized Jones

[1*] The angles which are typically used for the description of polarization on the Poincaré sphere are angles α and θ which define a measure of azimuth and ellipticity, respectively. The angles γ and η, and α and θ are related by the following equations:

$$\tan 2\alpha = \tan 2\gamma \cos \eta$$
$$\sin 2\theta = \sin 2\gamma \sin \eta$$

vector described in terms of angles γ and η takes a very simple form:

$$M = \begin{bmatrix} \cos\gamma \\ \sin\gamma\, e^{j\eta} \end{bmatrix}.$$

The polarization state can also be described in Cartesian system of coordinates as defined in Figure 10a. Such a vectorial description of polarization represents three out of four elements of Stokes vector (degree of polarization is not included):

$$\overrightarrow{P_1} = (\cos 2\gamma_1,\ \sin 2\gamma_1 \cos\eta_1,\ \sin 2\gamma_1 \sin\eta_1)\,. \tag{1}$$

An important property of the Poincaré sphere, which will be used in the next section, is that if P_1 represents the state of polarization and P_2 is the transmission state of a polarization analyzer, then the fraction of the intensity, F, transmitted through the analyzer depends on the angle σ between vectors $\overrightarrow{P_1}$ and $\overrightarrow{P_2}$ corresponding to those states:

$$F = \frac{1}{2} + \frac{1}{2}\cos\sigma. \tag{2}$$

If I represents intensity of light impinging on the analyzer the transmitted light is equal to FI. Since $\cos\sigma$ can be expressed as a dot product we can rewrite the above equation in terms of angles γ and η as

$$F = \frac{1}{2} + \frac{1}{2}\cos 2\gamma_1 \cos 2\gamma_2 + \frac{1}{2}\cos(\eta_1 - \eta_2)\sin 2\gamma_1 \sin 2\gamma_2. \tag{3}$$

While the above discussion sufficiently addresses the description of an SOP for our purposes it is necessary to also describe the birefringence of the optical system in which the light is propagating. The description of birefringence can also be approached from both the Jones matrices and Poincaré sphere viewpoints. In determining the Jones matrices for a medium of arbitrary birefringence, the elementary matrices[27] shown in Table 1 are virtually essential. The elementary matrices characterize infinitesimally thin layers of linear and circular birefringence, and their advantageous property is that they can be easily combined to describe a medium of general birefringence properties. The Jones matrices in turn, are exponential functions of elementary matrices. Therefore, to describe an anisotropic medium the desired amounts of "elementary media" are added together and the exponential function is formed; i.e. $\exp(N_0 z + N_{45} z + N_c z)$, where z denotes the thickness of the medium and N's are defined in Table 1. In all three elementary matrices β is equal to one-half of the difference between two principal propagation constants.

The three simplest special cases are Jones matrices which describe 0° linear, 45° linear, and circular birefringence. They are associated with the elementary matrices as follows. For 0° linear birefringence,

$$D(\psi_0) = e^{N_0 z} = \begin{bmatrix} e^{j\psi_0} & 0 \\ 0 & e^{-j\psi_0} \end{bmatrix},$$

for linear birefringence inclined by 45° to the chosen coordinate system,

$$D_{45}(\psi_{45}) = e^{N_{45} z} = \begin{bmatrix} \cos\psi_{45} & j\sin\psi_{45} \\ j\sin\psi_{45} & \cos\psi_{45} \end{bmatrix},$$

Symbol	Elementary matrix	Description
N_0	$\beta_0 \begin{bmatrix} j & 0 \\ 0 & -j \end{bmatrix}$	linear birefringence at 0°
N_{45}	$\beta_{45} \begin{bmatrix} 0 & j \\ j & 0 \end{bmatrix}$	linear birefringence at 45°
N_c	$\beta_c \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$	circular birefringence

Table 1: Elementary Matrices

and for circular birefringence,

$$R(\psi_c) = e^{N_c z} = \begin{bmatrix} \cos\psi_c & \sin\psi_c \\ -\sin\psi_c & \cos\psi_c \end{bmatrix}.$$

The angle ψ_i, which appears in the Jones matrices, denotes $\beta_i z$. Of course, the circular birefringence matrix is identical to the commonly used rotation matrix, since both matrices result in a rotation of the azimuth. The role of the 45° linear component of arbitrary birefringence is illustrated in Figure 10b. Just like polarization, we can view the birefringence of a medium as vector pointing in any direction. The 45° component is essential in defining an arbitrary type of birefringence since it can not be represented by merely combining the other two orthogonal components. The length of the vector defines the total birefringence of the media and it is expressed as $\beta = \sqrt{\beta_0^2 + \beta_{45}^2 + \beta_c^2}$. The direction of the vector is uniquely defined by its components $(\beta_0, \beta_{45}, \beta_c)$ or by the angles γ and η.

5.2 Sagnac Interferometer with Arbitrary SOPs

To study the polarization evolution and effects of non-reciprocal birefringence we define a Sagnac interferometer, illustrated in Figure 11, in which arbitrary polarization states P_1 and P_2 can be produced at the end of the PM fiber leads of an IOC. At this point it is assumed that the IOC itself is an ideal polarizer. The two PM fibers are connected to both legs of the IOC at angles γ_1 and γ_2. Each fiber has linear birefringence defined by retardation angles η_1 and η_2, respectively. Therefore, as described in the last section, the SOP at the end of each fiber will evolve into arbitrary polarization states P_1 and P_2 like those shown on the Poincaré sphere in Figure 9. Between the PM fibers we assume a birefringent element, T, whose thickness is small and whose affect on the polarization state is negligible. It is assumed that the element T can represent a medium of arbitrary birefringence which is different for clockwise (cw) and counterclockwise (ccw) waves, i.e. non-reciprocal birefringence. One special case is that of non-reciprocal circular birefringence, where the optical element is simply a length of SM fiber placed in magnetic field. Two other special cases are 0° and 45° linear non-reciprocal birefringence that may by thought as a time-varying linear birefringence. For modeling purposes, we do not need to consider the position of the element T with respect to the center of the loop, we simply assume that counter-propagating waves observe different birefringence. All the derivations in this section

will be performed assuming the use of monochromatic light. The analysis will be generalized to include a broadband light source in sections 5.3 and 5.4.

One of the goals in this section is to point out simple interpretations on Poincaré sphere. We will examine the polarization states of light emerging from PM IOC leads, as illustrated in Figure 12a, and predict the gyro response to the non-reciprocal birefringence which is assumed to be placed between the leads joined into the loop. As illustrated in Figure 12, polarization states are described differently before and after the leads are joined into the loop. This is due to a change in the direction of propagation which must be accounted for. The difference can be easily corrected for by changing the signs of γ_1 and η_1 which flips polarization state P_1 on another side of the plane established by points $HVRL$ on Poincaré sphere. This step is essential when moving from the Jones matrix notation describing the loop to the Poincaré sphere.

The interferometer from Figure 11 can be described in terms of Jones matrices as follows:

$$\begin{aligned} E_{cw} &= \tfrac{1}{2}\, P \cdot R_2 \cdot D_2 \cdot T_{cw} \cdot D_1 \cdot R_1 \cdot P \cdot \mathrm{v}\, e^{j\varphi(t)}, \\ E_{ccw} &= \tfrac{1}{2}\, P \cdot R_1 \cdot D_1 \cdot T_{ccw} \cdot D_2 \cdot R_2 \cdot P \cdot \mathrm{v}\, e^{-j\varphi(t)}, \end{aligned} \tag{4}$$

where,

E_{cw}, E_{ccw}	are clockwise and counter-clockwise waves,
$\varphi(t)$	is the bias modulation at the proper frequency,
v	is the Jones vector describing horizontal light,
P	is an ideal polarizer,
D_1, D_2	describe linear birefringence η_1 and η_2 of the PM leads,
R_1, R_2	describe rotations γ_1 and γ_2 between PM fiber axes and IOC axes,
T_{cw}, T_{ccw}	describe the optical element between PM fiber axes for *cw* and *ccw* waves.

In the model we assume that input intensity is equal to one. Therefore, the intensity, $|E_{cw} + E_{ccw}|^2$, describes a fraction of light, F, transmitted through a gyro circuit in the absence of bias modulation, $\varphi(t)$. For a monochromatic light source of intensity I, the power at the detector can be calculated as $F \cdot I$. For a broadband light source, F becomes dependent on the optical frequency ν and it can be treated as a gyro transfer function $F(\nu)$, as is done in sections 5.3 and 5.5.

In the presence of bias modulation, the fraction of transmitted intensity varies with time and can be written as $F'(t)$ where,

$$F'(t) = A \sin\varphi(t) + B \cos\varphi(t) + C, \tag{5}$$

and A, B and C are constants. The signal which is detected in the gyro is contained in the term $A \sin\varphi(t)$ and comprises only odd harmonics of the bias modulation signal. The strength of the signal can be determined by finding the value of the coefficient A. It can be easily shown that for the Sagnac phase shift, φ_S, the coefficient A can be expressed as

$$A = \frac{1}{2} F \sin(\varphi_S). \tag{6}$$

For the approximate case where $\varphi_S \ll 1$, A can be expressed as

$$A = \phi_S = \frac{1}{2} F\, \varphi_S. \tag{7}$$

This notation is convenient because for the common case of $\varphi_S \ll 1$, the coefficient ϕ_S is proportional to the phase shift φ_S and it is measured in radians. Note that ϕ_S is essentially the non-reciprocal phase shift due to rotation, φ_S, multiplied by the term F, representative of the amount of detected light that experiences this non-reciprocal phase shift. As such, ϕ_S can be treated as a phase shift. A similar nomenclature is used in later sections when various gyro errors are analyzed. It will be shown in the section 5.2.2 how to normalize gyro errors with respect to the signal $A = \phi_S$ measured for the Sagnac phase shift φ_S, and described by the equation (7). It is worth noting that the above treatment is quite general and did not assume any specific form of bias modulation or demodulation. The only assumption is that the bias modulation is at the proper frequency which is very typical for gyros with IOCs.

5.2.1 Fraction of Transmitted Intensity

We start by calculating the fraction of transmitted intensity $|E_{cw} + E_{ccw}|^2$ for T_{cw} and T_{ccw} being identity matrices and no bias modulation:

$$F = \frac{1}{2} + \frac{1}{2}\cos 2\gamma_1 \cos 2\gamma_2 - \frac{1}{2}\cos(\eta_1 + \eta_2)\sin 2\gamma_1 \sin 2\gamma_2. \tag{8}$$

The above equation with exception for two signs is identical to equation (3) which was derived from the Poincaré sphere. As mentioned previously, the sign discrepancy can be eliminated by changing the sign of γ_1 and η_1. The equivalence of equations (3) and (8) provides essential insight into the behavior of light in the depolarized FOG. Equation (8) determines the fraction of light in polarization P_1 being effectively transmitted through an analyzer P_2. In the gyro loop, the *ccw* wave associated with P_2 actually defines an analyzer for the *cw* wave (P_1) and vice versa. This is true because it is the state P_2 which could be "time-reversed" through the PM lead and pass unattenuated through the polarizer. Thus, assuming that T_{cw} and T_{ccw} do not alter the SOP, the transmitted light is a function of a cosine of the angle between P_1 and P_2, as was illustrated on the Poincaré sphere in Figure 9 and in equation (2). If P_1 and P_2 are orthogonal (opposite sides of the sphere) no light gets out of the loop, whereas if P_1 and P_2 are identical output is maximized. In general, the transmitted portion of polarization state P_1 is its P_2 component, the transmitted portion of P_2 is its P_1 component. This rule extends to all the points along the gyro loop. Polarization states P_1 and P_2 change along the fiber in the loop, however, the angle between them as measured on Poincaré sphere remains the same everywhere. (Thus the computation of the intensity is not dependent on which point in the loop is considered.) It is important to realize that reciprocity does not force the polarization state to be the same at any given point for the two counter-propagating waves. When the two waves meet they are in different polarization states and they define for each other the surviving polarization components which are different again.

5.2.2 Non-Reciprocal Birefringence

Most of the non-reciprocal effects such as the Sagnac effect itself or the Shupe effect are common between PM gyros and depolarized gyros. The unique effect which can be much more pronounced in a depolarized gyro has to do with the birefringence of a SM coil which can vary over time and with environmental changes. In PM gyros the birefringence is nominally linear throughout the gyro and so is the polarization of light. Some perturbations are introduced by imperfect PM fiber whose eigenmodes are not entirely linear, but to a small degree elliptical, and by stresses within the sensing coil which distorts intentionally induced internal stresses. These imperfections make PM gyros sensitive to some degree, to the non-reciprocal circular birefringence caused by the magneto-optic Faraday effect.[19,1] The SM fiber which is used in a depolarized gyro is not intended to have any birefringence. In practice, however, geometrical deviations from circular construction and stresses induced by the fiber jacket make the SM fiber birefringent. This became a concern in telecommunications where birefringence causes polarization mode dispersion (PMD), and spreading of signal pulses in SM fiber systems.[21] When SM fiber is employed in gyro coils, additional winding stresses are introduced due to bending, twisting and complex interactions between coil layers and the coil mount. Thus the birefringence of a SM coil is actually defined by both the coil structure and fiber imperfections. In addition, since stresses within the coil structure vary with the environment so does the birefringence. This time-varying birefringence results in a non-reciprocal effect since the two counter-propagating waves experience different birefringence due to the coil delay time. The analysis of non-reciprocal birefringence and its effects on the fiber optic gyro is the subject of this section.

We now analyze the effect of an observed difference in birefringence between the two counter-propagating waves, i.e., T_{cw} and T_{ccw} represent birefringent elements and $T_{cw} \neq T_{ccw}$. Such non-reciprocal birefringence due possibly to the Faraday effect or time-varying birefringence, affects various polarization states differently. The interaction between an element with non-reciprocal birefringence and lightwaves in arbitrary polarization states P_1 and P_2 is now calculated for various cases. The intuitively simplest case is that of non-reciprocal circular birefringence induced by the Faraday effect. This can be approached by inserting an element which represents non-reciprocal circular birefringence into the gyro model. Matrices $T_{cw} = R(x_c)$ and $T_{ccw} = R(-x_c)$ are selected so that the counter propagating waves observe rotation by an opposite angle. The phase modulation term $\varphi(t)$ is also included in equation (4). In examining the calculated fraction of transmitted intensity signified by equation (5), the coefficient A is of interest since it represents the component of the detected signal due to the non-reciprocal phase shift. Note that x_c is defined such that, for circularly polarized *cw* and *ccw* waves, the non-reciprocal phase shift is equal to $2x_c$. Then, the coefficient A takes a form analogous to that of equation (6) and it is expressed as $\frac{1}{2}\sin(2x_c)$. For $x_c \ll 1$, as in equation (7), A is denoted by ϕ_c and is simply equal to x_c.

For the case of arbitrary polarization states P_1 and P_2, ϕ_c is given by,

$$\phi_c = \phi_{cP1} + \phi_{cP2} = -\frac{1}{2}x_c \sin 2\gamma_1 \sin \eta_1 - \frac{1}{2}x_c \sin 2\gamma_2 \sin \eta_2. \tag{9}$$

Note that ϕ_c has two terms, each a contribution from P_1 and P_2 respectively. Physically the

two terms in ϕ_c can be understood as being the non-reciprocal phase shift x_c multiplied by, or weighted in proportion to the component of light in P_1 or P_2 that is sensitive to the non-reciprocal circular birefringence. This is in sharp contrast to the case of rotation where all the transmitted light experiences the Sagnac phase shift, φ_S [equation (7)]. The coefficient ϕ_c from equation (9) can be viewed as a "weighted" non-reciprocal phase shift. It should be noted that ϕ_c gives a non-reciprocal phase shift that is equivalent to and indistinguishable from the Sagnac phase shift. This perceived Sagnac phase shift, φ'_S, can be determined by comparing equations (7) and (9), and expressed as:

$$\varphi'_S = \frac{2}{F}\phi_c,$$

where F is the fraction of transmitted intensity described by equation (8). The fact that equation (9) is a sum implies that the polarization states of both counter-propagating waves are equally important and essential in computation of ϕ_c. The sensitivity to the Faraday effect is maximized when $\gamma_1 = \gamma_2 = \pi/4$ and $\eta_1 = \eta_2 = \pi/2$. This implies that the PM leads form quarter wave plates and that polarization states P_1 and P_2 are circular. This is not surprising since this is a typical construction of a Sagnac interferometer-based current sensor. (However, it is somewhat surprising that sensitivity of the current sensor represented by ϕ_c will drop only by 50% when one of the quarter wave plates is missing.) Expression (9) has a very simple interpretation on the Poincaré sphere. Notice that $\sin 2\gamma_1 \sin \eta_1$ represents a cosine of an angle ζ between the polarization state P_1 and the L–R axis (see Figure 9). Similarly, $\sin 2\gamma_2 \sin \eta_2$ represents a cosine of an angle ξ between polarization state P_2 and the L–R axis. Therefore, the sensitivity to the circular non-reciprocity is proportional to the sum of L–R axis direction cosines for polarization states P_1 and P_2.

Similar results can be obtained for the cases of 0° linear and 45° linear non-reciprocal birefringent elements, the only difference is that the angles ζ and ξ are measured from axes H–V or 45°–45°, respectively. In terms of angles γ and η the expressions for the "weighted" non-reciprocal phase shifts are:

$$\phi_0 = \phi_{0_{P1}} + \phi_{0_{P2}} = -\frac{1}{2}x_0 \cos 2\gamma_1 - \frac{1}{2}x_0 \cos 2\gamma_2 \tag{10}$$

for 0° linear non-reciprocal birefringence, and

$$\phi_{45} = \phi_{45_{P1}} + \phi_{45_{P2}} = \frac{1}{2}x_{45} \sin 2\gamma_1 \cos \eta_1 - \frac{1}{2}x_{45} \sin 2\gamma_2 \cos \eta_2 \tag{11}$$

for 45° linear non-reciprocal birefringence. In equation (10) x_0 represents half of the non-reciprocal phase shift observed by counter-propagating waves that are linearly polarized and aligned with the axes of birefringence. The birefringent element used in this derivation was $T_{cw} = D(x_0)$ and $T_{ccw} = D(-x_0)$. Similar assumptions apply to equation (11). Equation (10) indicates that the maximum sensitivity to variations of 0° linear birefringence takes place for $\gamma_1 = \gamma_2 = 0$ and is independent of η_1 and η_2. This scenario corresponds to a PM gyro in which the sensitivity to 0° linear non-reciprocal birefringence is actually maximized. Since only one polarization mode is used in a PM gyro, this sensitivity can be viewed as the change of propagation constant in that polarization mode. In equation (11) $\phi_{45_{P1}}$ and $\phi_{45_{P2}}$ approach their extreme values for $\gamma_1 = \gamma_2 = \pi/2$ and $\eta_1 = \eta_2 = \pi$, however, they subtract resulting in $\phi_{45} = 0$. This

special case is illustrated in Figure 12 and corresponds to a complete signal fading as polarization states P_1 and P_2 are orthogonal. Both equations (10) and (11) can be converted to rotation equivalent non-reciprocal phase shift by multiplying by $2/F$.

In summary, light in a 0° linear SOP is most sensitive to a change of 0° linear birefringence, light in a 45° linear SOP is most sensitive to a change of 45° linear birefringence and circularly polarized light best senses a change of circular birefringence. Since these three polarization states are orthogonal on the Poincaré sphere they cannot sense a change of birefringence corresponding to a orthogonal state, e.g., light in a 0° linear SOP cannot sense a change in magnitude of either circular or 45° linear birefringence.

The three components of birefringence can be used to define the birefringence vector $\overrightarrow{x}$ on the Poincaré sphere as:

$$\overrightarrow{x} = (x_0, x_{45}, x_c).$$

Then, the most general case of non-reciprocal birefringence requires that we consider the birefringence to be changing in value, azimuth and ellipticity from the initial value $\overrightarrow{x}$ (observed by *cw* wave) to the final value $\overrightarrow{x'}$ (observed by the *ccw* wave). This case is considerably more complex, however, the final result can be written in a very simple form. As before, the "weighted" non-reciprocal phase error, ϕ, can be simply described in term of angles ζ and ξ which are measured from the axis established by the birefringence change vector, $\overrightarrow{\Delta x} = \overrightarrow{x'} - \overrightarrow{x}$. The error is proportional to the length of the vector $\overrightarrow{\Delta x}$:

$$\phi = \frac{1}{4}\left|\overrightarrow{\Delta x}\right|(\cos\zeta + \cos\xi).$$

The angles ζ and ξ are illustrated in Figure 13. Alternatively, the above equation can be rewritten in the form of a dot product if vector notation is used:

$$\phi = \phi_{P1} + \phi_{P2} = \frac{1}{4}\overrightarrow{P_1}\cdot\overrightarrow{\Delta x} + \frac{1}{4}\overrightarrow{P_2}\cdot\overrightarrow{\Delta x}.$$

Again, the interpretation using the Poincaré sphere offers simplicity and elegance. It is worth noting that the simplicity of mathematical expressions in this section resulted from the initial assumption that the birefringent element is "thin," i.e., it does not affect the polarization state. The assumption of an infinitesimally thin wave plate does not limit the generality of the derived equations as even an element of large birefringence can be modeled by slicing it into thin wave plates. The cumulative effect of the element can then be predicted by integrating over the length of the element. Similarly, the integration can be performed over the source wavelength spectrum which will be illustrated in section 5.4. The important conclusion from this section is that non-reciprocal birefringence, which is certainly present in depolarized gyros, gives rise to a non-reciprocal phase shift that depends on the polarization state. In the later sections we identify the design which maximizes immunity from these effects.

5.3 Signal Fading and Wavelength Stability

The optical circuit of Figure 11 was chosen to study polarization behavior and effects of non-reciprocal birefringence on depolarized gyros. Despite its simplicity, this circuit describes a fairly general configuration of a depolarized gyroscope which is of interest in this section. By making angles γ_1 and γ_2 equal to 45° a configuration with two depolarizers is obtained. The lengths of the depolarizers are controlled by angles η_1 and η_2, which for the case of the broadband light source, become wavelength dependent. To obtain a configuration with one depolarizer γ_1 can be set to 45° and $\eta_2 = 0$. The matrix element T mathematically represents the SM sensing coil including its birefringence.

In our analysis thus far, we have assumed that the light is monochromatic light. However, a broadband light source is essential for FOG performance.[16] Thus in this section and subsequent sections the analysis is extended for the case of broadband light source. We assume that the broadband light source is described by the normalized power spectral density function, $S(\nu)$, such that $\int_{-\infty}^{\infty} S(\nu)d\nu = 1$. An important special case is that of a Gaussian power spectrum, described as

$$S(\nu) = \frac{1}{\Delta\nu\sqrt{\pi}} \exp\left(-\frac{(\nu - \nu_0)^2}{\Delta\nu^2}\right), \tag{12}$$

where ν is the optical frequency, ν_0 is the mean optical frequency and $\Delta\nu$ is $1/e$ half-bandwidth. Some of the results derived for Gaussian spectrum can be easily generalized for other spectra; e.g., a superfluorescent fiber light source spectrum can be described as a sum of Gaussian functions. At first let us assume a gyro configuration with an ideal single depolarizer and negligible coil birefringence. The assumption of an ideal depolarizer is not necessary; however, it gives the simplest mathematical description. In equation (8), which describes the fraction of light transmitted to the detector, we select $\gamma_1 = 45°$ (ideal first depolarizer) and $\eta_2 = 0$ (no second depolarizer). The retardation angle η_1 is wavelength dependent and it can be expressed as $\eta_1 = 2\pi\nu\tau_1$, where τ_1 denotes difference in propagation time between the fast and slow polarization mode in the PM fiber (birefringence induced delay). We assume that the dispersion is negligible and that the delay, τ_1, is wavelength independent. Under the assumptions listed above the equation (8) can be rewritten as

$$F(\nu) = \frac{1}{2} - \frac{1}{2}\cos(2\pi\nu\tau_1)\sin 2\gamma_2.$$

It is important to realize that $F(\nu)$ represents a fraction of transmitted light at the optical frequency ν and can be treated as the gyro transfer function which undulates the original source spectrum $S(\nu)$. Therefore, the light spectrum observed at the detector is $F(\nu)\,S(\nu)$. The angle γ_2, in the above equation, can be interpreted as a rotation of a linearly polarized input state within the SM coil. Depending on the angle γ_2 the source spectrum is undulated to a different degree; as the angle γ_2 changes from 0° to 45°, the undulation increases and becomes maximal at 45°. At this point, some wavelengths are transmitted through the gyro unattenuated and some are extinguished. The detected spectrum then resembles a "rounded comb function" within a Gaussian envelope. The frequency of spectral modulation increases with the birefringence induced delay τ_1 (length of depolarizer) and, for a depolarizer long enough, the undulation is so

dense that both the power and mean wavelength observed at the detector remain constant. The question which is often asked is how long does the depolarizer need to be in order to ensure this stability?[2] To answer this question we can represent the above single depolarizer gyro transfer function by the form

$$F(\nu) = \frac{1}{2} + a\cos(2\pi\nu\tau_1).$$

The total power is

$$P = \int_{-\infty}^{\infty} F(\nu)\, S(\nu)\, d\nu. \tag{13}$$

The mean optical frequency is then given by

$$\bar{\nu} = \frac{1}{P}\int_{-\infty}^{\infty} \nu F(\nu)\, S(\nu)\, d\nu. \tag{14}$$

Using $F(\nu)$ in equation (13) and assuming the total source power is normalized to unity,

$$P = \frac{1}{2} + a\int_{-\infty}^{\infty} S(\nu)\cos(2\pi\nu\tau_1)\, d\nu = \frac{1}{2} + a\,\Gamma(\tau_1), \tag{15}$$

where $\Gamma(\tau_1)$ is the degree of coherence.[30] Thus, the detected power fluctuates with the delay τ_1. Equation (15) represents what is commonly called "signal fading". If τ_1 is small, then $\Gamma(\tau_1)$ approaches one and it is possible to have no light at the detector. For a longer, more ideal depolarizer, small power fluctuations occur around the average power, $P_0 = \frac{1}{2}$. The amplitude of fluctuations, ΔP, when normalized to the average power, is given by

$$\frac{\Delta P}{P_0} = 2a\,\Gamma(\tau_1). \tag{16}$$

The obtained results are not unexpected; that is, the interference of two waves which decorrelate while propagating in orthogonal polarization modes of the depolarizer is governed by the degree of coherence $\Gamma(\tau_1)$, where τ_1 represents the birefringence induced delay. For the Gaussian spectrum (12) the power fluctuations are given by,

$$\frac{\Delta P}{P_0} = 2a\cos(2\pi\nu_0\tau_1)\exp\left(-\pi^2\Delta\nu^2\tau_1^2\right). \tag{17}$$

The effect on the mean detected optical frequency which is directly related to the gyro scale factor can be calculated. By solving integral (14) it can be shown that the fractional mean frequency shift is given by,

$$\frac{\Delta\bar{\nu}}{\bar{\nu}} \approx 2\pi a\tau_1\frac{\Delta\nu^2}{\nu_0}\sin(2\pi\nu_0\tau_1)\exp\left(-\pi^2\Delta\nu^2\tau_1^2\right). \tag{18}$$

Based on the equations above, the expected signal fading and mean wavelength instability were calculated for $a = 1/2$, $\nu_0 = 1555nm$ and $\Delta\nu = 15nm$, and illustrated in Figure 14. For convenience, the length of the depolarizer was described in beat lengths, instead of corresponding delay

time, τ. One beat length is the fiber length over which a phase difference of 2π is acquired between light propagating in the fast and slow propagation mode. Equation (18) gives a surprising result; namely, a very rapid decay of variations in power and wavelength with increasing length of the depolarizer. A depolarizer 100 beat lengths long, which gives only four visible peaks within the envelope of the considered Gaussian spectrum, gives wavelength instabilities which are only $3ppm$. This remarkable stability is a consequence of the powerful dependence on the coherence function, which in this case is an ideal Gaussian function. In practice, for longer depolarizers, the stability is determined by the "floor" of the coherence function but sub-ppm stability is easily achievable. Typically, wavelength instability is the largest contributor to the scale factor errors and, therefore, a primary concern. Small intensity fluctuations are of secondary importance; however, it is worth noting that they always appear together and both are caused by partially polarized light existing within the gyro loop.

It is very apparent from Figure 14b and equation (18) that there exists a depolarizer delay, τ_0, which maximizes fluctuations in the mean optical frequency. It can be easily shown from (18) that this delay is related to the optical bandwidth $\Delta\nu$ and takes a value close to the coherence time, τ_c:

$$\tau_0 = \frac{1}{\pi\sqrt{2}\Delta\nu} = \frac{\tau_c}{\sqrt{\pi}}.$$

The maximum fractional shift of the mean optical frequency corresponding to τ_0 is given by,

$$\frac{\Delta\bar{\nu}_{\max}}{\bar{\nu}} \approx a\sqrt{\frac{2}{e}}\frac{\Delta\nu}{\nu_0}. \tag{19}$$

Therefore, for $a = \frac{1}{2}$ (full undulation of the spectrum) the shift of the mean optical frequency approaches half of the optical bandwidth $\Delta\nu$! Of course, a frequency shift that large never takes place for an effective depolarizer since its birefringence must induce a delay substantially larger than the coherence time to accomplish depolarization. However, as is treated next, small delays on the order of τ_0 can result from the SM coil birefringence and have a significant impact on sub-optimum depolarized designs. Equation (19) becomes relevant when analyzing cases of non-ideal depolarized gyros involving imperfect depolarizers combined with a birefringent SM coil.

We now consider the case of an imperfect depolarized gyro. We can easily represent imperfect depolarizers by setting splice angles γ_1 and γ_2 unequal to 45°; however, the representation of the coil birefringence has to be chosen very judiciously to properly capture the impact of the SM coil properties on wavelength stability. As mentioned before, SM fiber within the sensing coil has elliptical birefringence which is introduced by geometrical imperfections of glass, stresses in the jacket, and stresses that are introduced by winding. In the latter case, every turn within the coil adds (to the initial elliptical birefringence of the SM fiber) linear birefringence induced by bending and stresses within the coil structure which is substantially smaller than a full-wave of retardance. These small segments of birefringence are perturbed at turn-to-turn and layer-to-layer transitions, where the geometry of bending and stresses is different, forming a very complicated pseudo-periodic structure. The cumulative effect of thousands of turns (like thin wave plates) is very wavelength dependent resulting in a spread in the birefringence induced delays measured with a broadband light source. The distribution of delays is centered about a zero mean. The delay

τ_0, which yields the maximum wavelength shift (19), is always within the described distribution of delays. Furthermore, the distribution of delays, although continuous, tends to form discrete peaks which change their location with slight perturbations of the environment.[31] Therefore, the worst case assumption for this analysis is that the coil birefringence may take a value which will induce the maximum shift of optical frequency described by equation (19). The coil is thus represented as a linearly birefringent element, placed between the depolarizers at an arbitrary angle ψ and having retardance angle denoted as $\eta_0 = 2\pi\nu\tau_0$.

To find the intensity at the detector we use the gyro description as given by equations (4) and follow the procedure introduced in the last section. As before, the resulting intensity depends on optical frequency, ν, and represents the fraction of transmitted light $F(\nu)$ which is treated as the gyro transfer function. The transfer function has many terms which reflect delays encountered in the individual depolarizers, in the SM coil, and in various combinations of them. Below we list only a few terms of $F(\nu)$. Among them, the important term which corresponds to the delay, τ_0, encountered within the SM coil is present:

$$\begin{aligned} F(\nu) &= 1/2+ \\ & 1/4\cos 2\pi\nu\tau_1 \sin 2\gamma_1 \cos 2\gamma_2 \sin 4\psi- \\ & 1/4\cos 2\pi\nu(\tau_1+\tau_2)\sin 2\gamma_1 \sin 2\gamma_2 \sin^2 2\psi+ \\ & 1/2\cos 2\pi\nu\tau_0 \cos 2\gamma_1 \cos 2\gamma_2 \sin^2 2\psi+ \\ & \vdots \end{aligned}$$

Various terms which appear in the transfer function modulate the original power spectrum; however, as long as this undulation is "dense" (large τ) the impact on the mean wavelength is negligible. Among all of the terms of transfer function $F(\nu)$ only the birefringence of the SM coil is low enough to induce modulation of the source that is of concern. We now assume that the angles $\Delta\gamma_1$ and $\Delta\gamma_2$ are small and describe angular deviation of the splices from 45°, and that $\psi = 45°$. Thus, the maximum shift of the mean optical frequency as defined by equation (19) is given by,

$$\frac{\Delta\bar{\nu}_{\max}}{\bar{\nu}} \approx 2\sqrt{\frac{2}{e}}\Delta\gamma_1\Delta\gamma_2\frac{\Delta\nu}{\nu_0}.$$

For 0.5° splice misalignment, $\nu_0 = 1555$nm and $\Delta\nu = 15$nm, the maximum shift is equal to about 1.3ppm which is well within a typical specification for scale factor stability of a navigation grade gyro (20ppm). However, if only one depolarizer were used, the maximum shift would increase to about 144ppm ruling out the use of this configuration for inertial navigation purposes. The above example gives the answer to the question posed in section 3 on how depolarized does the light have to be. The misalignment of 0.5° results in light being 96.6% depolarized.

The two cases presented here are very strongly connected. As a matter of fact, just from the first derivation one can predict the significance of coil birefringence and the importance of the second depolarizer. From the first derivation it was observed that the depolarizer (or any birefringent element) needs to be sufficiently long. If the depolarizer is sufficiently long and its 45° splice is ideal, light gets completely depolarized and its propagation through any additional birefringent elements, e.g. SM coil, is of no consequence. However, if the splice angle deviates

from 45°, regardless of the depolarizer length, a portion of light will remain polarized. When this residual polarized component enters the SM coil we have to again ask the same question: Is the coil birefringence sufficiently large to maintain a stable mean wavelength? Of course, the coil birefringence is not large and it can result in wavelength shifts comparable to the maximum wavelength shift discussed earlier. Therefore, the only solution to maintaining the excellent mean wavelength stability is to sufficiently depolarize the light which is injected into the SM coil. This can be accomplished by using two depolarizers and making suitably accurate 45° splices.

5.4 Optimal Configuration

The optimal depolarized configuration is of course application dependent. Since our goal was the attainment of navigation grade performance we have focused on a closed loop system with an integrated optics chip. The integrated optics technology enables fabrication of a phase modulator with sufficient bandwidth to implement various loop closure signal processing schemes. Once this is established, the next question is how many depolarizers should be used, and where should they be located in the optical circuit? For example, it has been established that two depolarizers are necessary for maintaining satisfactory mean wavelength stability, but can one achieve the best bias stability by placing both of them on one side of the sensing coil? It has been shown that two depolarizers are necessary to reduce magnetic sensitivity[20,34] and that for the configuration with an IOC, this forces the placement of the two depolarizers on the opposite sides on the gyro loop. Two depolarizers were also found to be essential in suppression of the sensitivity to thermal transients.[33] Our gyro test results presented in section 3 indicate navigation-grade bias stability in the twin depolarizer case; whereas, in gyro experiments conducted at Honeywell using a single depolarizer, we observed bias stability degradation of up to 100 times. But how does this compare to the errors stemming from non-reciprocal birefringence discussed in section 5.2.2? The configurational requirements to achieve high bias stability in the case of the general non-reciprocal birefringence is the subject of discussion in this section.

To begin the discussion, we have to generalize the results in section 5.2.2 to characterize the propagation of broadband light in a depolarized FOG. One can envision the representation of the two polarization states P_1 and P_2 (see Figure 11) on the Poincaré sphere for the two cases which were analyzed in the last section. The first case of a single, ideal depolarizer is represented by letting $\eta_2 = 0$, $\gamma_1 = 45°$ and $\eta_1 = 2\pi\nu\tau_1$. The angle γ_1 rotates the linearly polarized light H in the IOC to the +45° point once light enters the fiber depolarizer (see Figure 15). Then, the retardance angle of the depolarizer η_1 rotates the +45° SOP around the axis H–V. Because the retardance angle η_1 changes with optical frequency, ν, as the frequency increases the SOP revolves around the great circle which contains points R, L and +45°. Consequently, a broadband light wave is represented on the Poincaré sphere as a great circle. Light is considered to be depolarized when the curve formed by all the polarization states has the center of gravity in the center of the sphere. The points along the curve, in this case a great circle, are not weighted equally – they are weighted by the power spectral density function, $S(\nu)$. In the case treated in the last section, for a Gaussian spectrum, about four revolutions around the sphere were sufficient to

make the wavelength stable to within $3ppm$. This simply implies that for a Gaussian spectrum the center of gravity is nearly in the center of the sphere and it does not have a tendency of moving towards any specific wavelengths which have their SOPs preferentially distributed along the great circle. From this pictorial representation it becomes intuitively clear that we can always depolarize broadband light by simply making enough revolutions around the Poincaré sphere.[2*] On the other hand, if the rotation angle γ_1 is not equal to 45°, the SOP revolves around the circle whose center of gravity is not aligned with the center of the sphere and light remains partially polarized regardless of the depolarizer length (see Figure 15). The above describes the behavior of polarization state P_1 for light propagating through the first depolarizer. The behavior of polarization state P_2 is analogous if a second depolarizer is used; however, in a circuit without the second depolarizer , the polarization state P_2 remains stationary on the sphere.

With this new picture in mind we can reexamine the sensitivity to non-reciprocal birefringence which can cause bias instabilities in the gyro. Consider the case of polarization state P_1 under the influence of non-reciprocal circular birefringence. This may be due to the presence of magnetic field or a time varying torsion of the fiber in a vibration environment. As shown previously, the sensitivity to non-reciprocal circular birefringence is proportional to the cosine of the angle, ζ, between the polarization state and the axis L–R (see Figure 9). As the wavelength changes the polarization state P_1 revolves around the great circle, and the angle ζ varies between 0° and 180°. Thus the cosine of ζ varies between -1 and 1. As a consequence, the non-reciprocal phase shift (bias) experienced by each wavelength varies in value and sign over the wavelengths comprising the light source spectrum. This leads to an averaging process when the full spectral width is considered. The mathematical description of the averaging process is virtually identical to the description of signal fading. In both cases we integrate a sinusoidally varying function weighted by the power spectrum density function, $S(\nu)$, over the optical frequency. For signal fading these sinusoidally varying functions are contained within the transfer function $F(\nu)$. For non-reciprocal circular birefringence and polarization state P_1, sinusoidal variations are introduced by $\phi_{c_{P1}}$ after substituting $2\pi\nu\tau_1$ for η_1 in equation (9). In both cases sinusoidal variations reflect rotation of the polarization state P_1 around the Poincaré sphere. To determine the form of equation (9) for a broadband light source the behavior of both polarization states P_1 and P_2 must be considered. For simplicity, we only consider $\phi_{c_{P1}}$ since the behavior of $\phi_{c_{P2}}$ is analogous. For a broadband light source the P_1 component of the "weighted" non-reciprocal phase shift, $\phi_{c_{P1}}$, is multiplied by $S(\nu)$ and integrated over the optical frequency, ν. This yields a new quantity $\Phi_{c_{P1}}$ which describes the P_1 component of the "weighted" non-reciprocal phase shift for the broadband light source under the earlier defined assumption that $\int_{-\infty}^{\infty} S(\nu)d\nu = 1$.

$$\Phi_{c_{P1}} = \int_{-\infty}^{\infty} \phi_{c_{P1}}(\nu)\, S(\nu)\, d\nu.$$

After substituting for $\phi_{c_{P1}}(\nu)$ we obtain:

$$\Phi_{c_{P1}} = \frac{1}{2} x_c \sin 2\gamma_1 \int_{-\infty}^{\infty} S(\nu)\, \sin 2\pi\nu\tau_1\, d\nu,$$

[2*] This argument is equivalent to the argument of the depolarizer being sufficiently longer than the light source coherence length.

where x_c represents circular non-reciprocal birefringence described by Jones matrices $R(x_c)$ and $R(-x_c)$ for counter propagating waves. The above integral defines the sine Fourier transform of $S(\nu)$ and yields a function whose envelope is identical to the degree of coherence, $\Gamma(\tau)$; however, its oscillatory behavior is shifted in phase by 90° relative to $\Gamma(\tau)$. Since we are primarily interested in the decay of oscillations and not in their phase we denote this function also as $\Gamma(\tau)$. Assuming that the angle γ_1 is nearly equal to 45° the above equation can be rewritten as:

$$\Phi_{c_{P1}} = \frac{1}{2} x_c \, \Gamma(\tau_1)$$

Similar to the results obtained for the treatment of signal fading (16), we observe that $\Phi_{c_{P1}}$ decays rapidly with increasing length of the depolarizer, with the decay defined by the degree of coherence, $\Gamma(\tau_1)$. By the same token, if the depolarizer is very short (or not used) $\Gamma(\tau_1)$ approaches unity and $\Phi_{c_{P1}}$ approaches half of the non-reciprocal phase shift x_c. By including both polarization states P_1 and P_2, we can capture the equivalent of the "weighted" non-reciprocal phase shift Φ_c for broadband light:

$$\Phi_c = \Phi_{c_{P1}} + \Phi_{c_{P2}} = \begin{cases} 1/2\ x_c\ \Gamma(\tau_1) + 1/2\ x_c & \text{for single depolarizer} \\ 1/2\ x_c\ \Gamma(\tau_1) + 1/2\ x_c\ \Gamma(\tau_2) & \text{for two depolarizers} \end{cases} \tag{20}$$

It becomes obvious from equation (20) that only the configuration with two depolarizers on opposite sides of the loop is immune to the presence of the non-reciprocal circular birefringence. The suppression of effects from the non-reciprocal circular birefringence x_c is proportional to the degree of coherence, $\Gamma(\tau)$. Just as for the case of signal fading, the degree of suppression is eventually limited by the "floor" of the coherence function. For many light sources this is below 10^{-5}. The single depolarizer configuration does not provide inherent attenuation of this effect since the second-term of (20) contains x_c alone. It is worth noting that the net non-reciprocal phase shift or the perceived Sagnac phase shift, φ'_S, takes a form very similar to (20). This is because the fraction of transmitted intensity, F, is constant for depolarized light and equal to $P_0 = \frac{1}{2}$. Therefore,

$$\varphi'_S = \frac{2}{F}\Phi_c = 4\Phi_c.$$

Similarly, to determine the sensitivities for other types of non-reciprocal birefringence, one can examine the behavior of polarization state P_1 only. By inspection of Figure 15 or equation (11) one can tell that the suppression of sensitivity to 45° linear non-reciprocal birefringence is the same as for circular birefringence. The only difference in representation on the Poincaré sphere is that the angle ζ is measured from the 45°–45° axis instead of the L–R axis; however, the averaging process is identical since both axes intersect the great circle traced by revolutions of the polarization state P_1. A different mathematical description [equation (10)] is obtained for the case of 0° linear non-reciprocal birefringence effects. In this case the angle ζ is measured from H–V axis. For an ideal 45° splice the angle ζ is orthogonal to the plane of great circle traced by the polarization state P_1 and, therefore, the sensitivity is always equal to zero. In practice, splices always deviate from 45° which makes ζ different than 90° and the sensitivity finite. Since $\phi_{0_{P1}}$ in equation (10) is independent of optical frequency, ν, we can omit integration over optical

frequency when determining quantity $\Phi_{0_{P1}}$ for the broadband light source. Simply,

$$\Phi_{0_{P1}} = \phi_{0_{P1}} = -\frac{1}{2} x_0 \cos 2\gamma_1,$$

where x_0 represents linear non-reciprocal birefringence described by Jones matrices $D(x_0)$ and $D(-x_0)$ for counter propagating waves. Such birefringence can be realized by applying alternately tensile or compressive forces to the element represented by T in the direction aligned with the principal axes of birefringence of the depolarizer. As the angle γ_1 approaches 45°, $\Phi_{0_{P1}}$ becomes negligibly small. However, if γ_1 approaches zero (no depolarizer) $\Phi_{0_{P1}}$ approaches half of the non-reciprocal phase shift x_0. By including both polarization states P_1 and P_2, and denoting the small deviation of γ from 45° by $\Delta\gamma$ we can express the "weighted" non-reciprocal phase shift, Φ_0, for broadband light source as,

$$\Phi_0 = \Phi_{0_{P1}} + \Phi_{0_{P2}} = \begin{cases} 1/2\ x_0\ \Delta\gamma_1 + 1/2\ x_0 & \text{for single depolarizer} \\ 1/2\ x_0\ \Delta\gamma_1 + 1/2\ x_0\ \Delta\gamma_2 & \text{for two depolarizers} \end{cases} \tag{21}$$

where it was assumed that $\gamma_2 = 0$ for the case of one depolarizer. Again, it is obvious from equation (21) that only the configuration with two depolarizers on opposite sides of the loop is immune to non-reciprocal linear birefringence at zero degrees of inclination. For the two depolarizer configuration, the suppression of non-reciprocal phase shift x_0 is determined by both $\Delta\gamma_1$ and $\Delta\gamma_2$, and it is essential that both 45° splices are equally good. Intuitively this requirement is not obvious since it takes only one good 45° splice to depolarize light or to eliminate signal fading. Typical values of $\Delta\gamma$ are in the neighborhood of 0.01; therefore, the suppression of sensitivity to linear non-reciprocal birefringence is not as effective as it was for circular non-reciprocal birefringence. In the latter case, the degree of suppression was established by the "floor" of the coherence function.

In the most general case when the birefringence changes in value, azimuth and ellipticity all three of the orthogonal components, described above, have to be considered. To suppress sensitivity to arbitrary non-reciprocal birefringence we need a depolarizer which is long enough and spliced precisely at 45°. These two elements are thus defining ideal depolarization. Making the depolarizer longer is quite simple, however, the precision of a 45° splice is typically limited to around 0.5°. This practical consideration limits the suppression of arbitrary non-reciprocal birefringence errors to about 2 orders of magnitude. Cascading two 45° splices improves the suppression by another 2 orders of magnitudes, however, it tends to complicate the optical circuit.

It is important to realize that the above mathematical description was based on a gyro configuration with two depolarizers and an infinitesimally thin element with non-reciprocal birefringence properties. The element was assumed to be placed between the depolarizers as shown in Figure 11. For this case, the gyro sensitivity to circular and 45° linear non-reciprocal birefringence is associated with depolarizer length and the sensitivity to 0° linear non-reciprocal birefringence is determined by the precision of 45° splice. In general, when a birefringent coil is placed in the gyro circuit and when polarization states evolve along the gyro loop, the sensitivity to non-reciprocal

birefringence depends on both: the depolarizer length and its alignment. This effectively makes the error suppression limited by the alignment accuracy of the 45° splices.

In summary, two depolarizers are required to provide maximum immunity to non-reciprocal birefringence, however, they need to be placed on the two opposite sides of the loop to ensure that depolarized light enters the gyro loop from both its ends (as is illustrated in Figure 16). The experimental data supports these conclusions.[20,15,32,33] For depolarizer splice misalignment of 0.5°, which corresponds to light being 96.6% depolarized, we demonstrated the navigation grade gyro bias stability of 0.005°/hr (section 3).

5.5 Random Noise

The model of random noise in a depolarized gyro is the same as in a PM gyro with one exception: the relative intensity noise (RIN) in a depolarized gyro is higher.[35] The purpose of this section is to explain and quantify the increase in RIN in the optimal depolarized configuration shown in Figure 16 under the assumption that light source exhibits thermal properties. This assumption is fully satisfied for superfluorescent fiber light sources (FLS) which are often used in navigation grade gyros, and sufficiently satisfied for LEDs and SLDs. FLS designs based on erbium doped fibers typically have higher intensity noise than most LEDs and SLDs because of their narrower spectrum. Most often, the intensity noise is described as a consequence of beating between individual oscillators comprising the broadband spectrum. The narrower the spectrum is, the fewer the number of oscillators are, and stronger their interference is relative to the average optical power. In statistical optics this effect is described in terms of the coherence time,[36] τ_c, and is captured by a very simple formula,

$$RIN = \sqrt{2\tau_c}, \tag{22}$$

where RIN is expressed in terms of the fractional intensity noise per $\sqrt{Hz}$ bandwidth. The above formula is applicable to thermal light that is passed through a polarizer (the FOG case) and therefore is generated by only half as many independent oscillators as in the original, unpolarized source. This gives rise to the $\sqrt{2}$ in equation (22). The coherence time can be calculated directly from the complex degree of temporal coherence,[36] $\mathbf{\Gamma}(\tau)$, as

$$\tau_c = \int_{-\infty}^{\infty} |\mathbf{\Gamma}(\tau)|^2 \, d\tau. \tag{23}$$

Using equations (22) and (23) we can calculate the intensity noise in a depolarized gyro from the coherence function of the light incident on the detector. The analysis can be started by calculating the gyro transfer function for the architecture shown in Figure 16 assuming that the 45° splices and the polarizer are ideal. As usual, some considerations have to be given to the model of the SM coil. Since we do not expect the coil birefringence properties to affect random noise it can be assumed, for simplicity, that the coil only rotates the SOP by an arbitrary angle ψ. The description of the gyro is then given by equations (4) with matrices T_{cw} and T_{ccw} being

equal to $R(\psi)$. The obtained expression for gyro transfer function is:

$$F(\nu) = \tfrac{1}{2} - \tfrac{1}{2}\cos^2\psi\cos 2\pi\nu(\tau_1+\tau_2) + \tfrac{1}{2}\sin^2\psi\cos 2\pi\nu(\tau_1-\tau_2).$$

The optical circuit transfer function $F(\nu)$ will thus modulate the original source spectrum predominantly at two frequencies which correspond to sum and difference of depolarizer delays τ_1 and τ_2. This undulation changes the original source coherence function, $\mathbf{\Gamma}(\tau)$, into the new coherence function, $\mathbf{\Gamma}'(\tau)$ of the light at the detector. We find $\mathbf{\Gamma}'(\tau)$ by taking the Fourier transform of $F(\nu)S(\nu)$ and normalization. The new coherence function of the detected light can be expressed in terms of the original coherence function as

$$\begin{aligned}\mathbf{\Gamma}'(\tau) &= \mathbf{\Gamma}(\tau) + \\ &\quad 1/2\sin^2\psi\,\mathbf{\Gamma}(\tau+(\tau_1-\tau_2)) + 1/2\sin^2\psi\,\mathbf{\Gamma}(\tau-(\tau_1-\tau_2)) - \\ &\quad 1/2\cos^2\psi\,\mathbf{\Gamma}(\tau+(\tau_1+\tau_2)) - 1/2\cos^2\psi\,\mathbf{\Gamma}(\tau-(\tau_1+\tau_2)).\end{aligned}$$

Four additional peaks appear on $\mathbf{\Gamma}'(\tau)$ at locations $\pm(\tau_1-\tau_2)$ and $\pm(\tau_1+\tau_2)$. The shape of the additional peaks is identical to the shape of the $\tau = 0$ peak for $\mathbf{\Gamma}(\tau)$. By using equation (23) and assuming that there is no overlap between the peaks (a necessary condition for depolarization) we can express the coherence time of light seen at the detector, τ_c', as function of source coherence time τ_c:

$$\tau_c' = \tau_c\left(1 + \frac{1}{2}\sin^4\psi + \frac{1}{2}\cos^4\psi\right).$$

The coherence time τ_c' is always larger than τ_c. The minimum increase of the coherence time corresponds to the rotation angle $\psi = \pi/4$ and a coherence function with four additional peaks of the same height. The maximum increase of τ_c' corresponds to the rotation angles of $\psi = 0$ or $\psi = \pi/2$ and a coherence function with only two additional peaks. The values of τ_c' are thus bounded by inequality

$$1.25\tau_c \leqslant \tau_c' \leqslant 1.5\tau_c$$

The average increase of the coherence time is equal to 37.5%. In Figure 17 we show one side of the coherence function measured at the output of a gyro circuit with a SM birefringent coil (the other side is the same by definition). The peaks at $\tau_1 - \tau_2$ and $\tau_1 + \tau_2$ appear as expected; however, they are broader than the $\tau = 0$ peak and smaller in amplitude. This can be explained by the existence of coil birefringence, which was assumed to be absent in the above model. The computed increase of the coherence time is 27.7% which is within the limits defined above.

The increase of intensity noise, as apparent from the derivation, is a result of depolarization. It can be explained by assuming that the light source emits wavepackets which change randomly in phase. These wavepackets propagate in both polarization modes of the depolarizer and evolve into some final polarization state at its end. Changes of phase of the emitted light, therefore, result in random variations of the polarization state at the end of the depolarizer. In the gyro circuit using a polarizer this translates into a change in the instantaneous intensity. This process can be described as conversion of phase noise into intensity noise. The model which was chosen for this analysis did not require a detailed description of this process.

The RIN increases in depolarized gyro on average by 17.2% over that of its PM counterpart. It should be noted that this is smaller than the 41% increase of shot noise due to additional

$3dB$ loss in the loop. The first question which comes to mind is why is this 17.2% significant, especially that we can overcome intensity noise by increasing the modulation depth?[1] Indeed, this is not a problem for a navigation-grade FOG. The concern is related to the high precision applications where the noise subtraction techniques are often employed to decrease the random walk coefficient (RWC) below $0.0003°/\sqrt{hr}$.[37] Since the additional intensity noise in depolarized gyroscopes stems from source phase variations it can not be subtracted; therefore, this may preclude use of depolarized gyroscopes in some very high performance applications.

For completeness we include a formula for the RWC which in addition to the intensity noise, includes shot noise and Johnson noise. The RWC is expressed as a function of the modulation depth, φ_m:

$$RWC\left(\varphi_m\right) = \frac{\sqrt{2}\lambda c}{2\pi LD} \frac{\sqrt{\frac{4kT}{R_f} + 2e\eta I\left(\frac{1+\cos\varphi_m}{2}\right) + \frac{11}{4}\tau_c\eta^2 I^2\left(\frac{1+\cos\varphi_m}{2}\right)^2}}{\eta I \sin\varphi_m},$$

where,

λ	is the wavelength,
c	is the speed of light in vacuum,
L	is the coil length,
D	is the coil diameter,
k	is the Boltzman constant,
T	is the temperature,
R_f	is the detector preamp transimpedance resistor,
e	is the electron charge,
η	is the detector responsivity and
I	is the light intensity at the detector.

5.6 Classical Polarization Errors

Polarization errors which are due to the imperfect control and filtering of the polarization state of light in the FOG have been described by many authors.[1,16,38,15,40] Unlike the truly non-reciprocal effects described in sections 5.2.2 and 5.4 on non-reciprocal birefringence (which exist even for a perfect gyro polarizer) these effects stem from component and splice imperfections. They always involve a parasitic light wave which cross-couples to the "wrong" polarization mode either at the splices or within the coil and propagates through the polarizer in its reject axis. Therefore, polarization errors depend on the quality of the polarizer and on the amount of the polarization cross-coupling. The errors or bias instabilities that these parasitic waves generate depend on the degree of coherence between them and the main signal light wave (primary wave) or other parasitic waves they interfere with. The suppression of these "classical" polarization errors in depolarized gyroscopes is substantially more difficult than in PM gyroscopes because of two reasons:

1. The error terms are significantly more abundant, and
2. They are one to three orders of magnitude larger.

Before analyzing these errors for the optimal configuration of Figure 16, it is useful to first outline the design difficulties inherent in the depolarized gyro. Typically, two groups of polarization errors are identified: amplitude-type errors and intensity-type errors.[1,39] Intensity-type polarization errors result from the interference of two cross-coupled waves which propagate through the reject-axis of the polarizer. Because of the excellent polarizers which are available nowadays, intensity errors are of lesser concern. Of primary concern are the amplitude-type polarization errors, especially in gyroscopes using semiconductor light sources which often have coherence functions with multiple coherence peaks.

Amplitude-type polarization errors, denoted by the apparent optical phase difference observed in the gyro, Φ_A, result from the interference of the main signal wave and a cross-coupled wave. The cross-coupled wave couples to the "wrong" polarization in front of the polarizer, propagates through the reject axis of the polarizer and cross-couples back to the "right" polarization within the sensing coil. A typical PM gyroscope with an IOC is constructed to guide linearly polarized light in the fast axis, therefore, the cross-coupled wave spends some time propagating in the slow axis and encounters a birefringence induced delay with respect to the signal wave. The adjustment of this delay is a very powerful technique for suppression of polarization errors. The problem is thus identical to interfering an optical wave $E(t)$ with its own delayed image $E(t-\tau)$, well known as a temporal coherence problem. The above discussion can be summarized in a very simple equation

$$\Phi_A \propto k_1 k_2 \varepsilon \Gamma(\tau),$$

where

Φ_A	is the effective optical phase error observed in the gyro,
k_1, k_2	are amplitude cross-coupling coefficients (measure of splice misalignment),
ε	is the amplitude extinction of the polarizer and
$\Gamma(\tau)$	is the degree of coherence of the interfering waves.

The following design practices make polarization error control in PM gyroscopes especially simple:

1. Use of the same optical axis throughout the gyroscope accumulates birefringence induced delays as cross-coupled waves propagate through the gyroscope. This means that the delay τ is very large and, for most practical light sources, $\Gamma(\tau)$ is very small.

2. Use of large birefringence in the optical circuit introduces a large delay and again minimizes the coherence and interference from parasitic waves ($\Gamma(\tau)$ is small). The PM sensing coil itself is highly birefringent and therefore eliminates many polarization errors.

3. Use of readily available light sources that have coherence functions with very small and very few peaks; e.g., a broadband fiber light source has no peaks other than that located at $\tau = 0$.

We now compare the above advantages of the PM gyro with the depolarized configuration of Figure 16 to point out the challenges of designing a low-drift (high bias stability) depolarized FOG. The two 45° depolarizer splices are essential in providing wavelength stability and immunity to non-reciprocal birefringence (as shown in the previous sections), however, they maximize polarization cross-coupling. Assuming that, in a PM gyroscope, the splices in the loop can be aligned to within 1° ($k = 0.017rad$) the increase of polarization errors caused by the 45° splices is about 40 times ($1/k\sqrt{2}$) for amplitude errors and 1600 times ($1/k^2\sqrt{2}$) for intensity errors! Not only are the errors substantially larger but the 45° splices unfortunately bar the use of the very powerful technique mentioned earlier, i.e., the use of one principal axis of birefringence throughout the gyroscope. Delays introduced by the depolarizers can add to or subtract from the delays encountered in the front end of the gyroscope and IOC. The SM coil itself is birefringent and its birefringence can also add to or subtract from these delays. Also, delays encountered on one side of the loop can combine with delays seen the other side of the loop. Consequently, there are numerous parasitic waves and, therefore, several hundred errors which have to be accounted for in a design suitable for navigation grade performance. Despite these difficulties, careful analysis shows that the control of polarization errors is possible.[18]

To analyze polarization errors in the gyro circuit shown in Figure 16 we have to modify the gyro description given by equations (4). A cross-coupling point of amplitude k_0 is added in front of the IOC (represented by a small angle rotation matrix), matrix P is modified to describe a polarizer with a finite extinction ratio (ε), two additional matrices D_{01} and D_{02} are added to capture the delays encountered by the cross-coupled waves traveling through the IOC in its reject axis (relative to the waves on the transmission axis), and the matrices $T_{cw} = T_{ccw}$ are defined as arbitrary rotation matrices $R(\psi)$ to capture all the possible ways is which delays from both sides of the gyro can combine. The modified set of equations is:

$$\begin{aligned} E_{cw} &= \tfrac{1}{2}\, K_0 \cdot P \cdot D_{02} \cdot R_2 \cdot D_2 \cdot T_{cw} \cdot D_1 \cdot R_1 \cdot D_{01} \cdot P \cdot K_0 \cdot \mathrm{v}\, e^{j\varphi(t)}, \\ E_{ccw} &= \tfrac{1}{2}\, K_0 \cdot P \cdot D_{01} \cdot R_1 \cdot D_1 \cdot T_{ccw} \cdot D_2 \cdot R_2 \cdot D_{02} \cdot P \cdot K_0 \cdot \mathrm{v}\, e^{-j\varphi(t)}, \end{aligned}$$

The rest of the procedure is the same as described in the section on non-reciprocal birefringence. The outcome from the modified model is the collection of all the polarization errors whose full listing would serve no purpose. Instead we will examine a few key error terms and provide a brief description of the rest. It is worth noting that the outcome from the Jones matrix model, as illustrated in previous sections, describes the monochromatic-light case. The extra steps which are taken to make the results applicable for a broadband light source are, weighting by the power spectral density function, $S(\nu)$, and integration over optical frequency, ν. This yields the dependence on the degree of coherence $\Gamma(\tau_i)$ provided that τ_i is not a function of optical frequency for the spectral width of interest.

Perhaps the most surprising result of all is that the 45° splices themselves do not cause the largest errors. From the knowledge of PM gyroscopes one would expect that the largest errors are associated with waves which travel in "wrong" polarization mode between the cross-coupling point in front of the IOC (k_0) and the 45° splices. If $\Delta\gamma_1$ an $\Delta\gamma_2$ denote the deviation from 45°, the apparent optical phase differences between *cw* and *ccw* waves (i.e. errors) can be described

as

$$\Phi_{\tau_{01}} \propto 2\,k_0\,\varepsilon\,\Delta\gamma_2\,\Gamma(\tau_{01})\ ,$$
$$\Phi_{\tau_{02}} \propto 2\,k_0\,\varepsilon\,\Delta\gamma_1\Gamma(\tau_{02})\ .$$

The error $\Phi_{\tau_{01}}$ is due to the light propagating in the wrong polarization mode between a cross-coupling point in front of the IOC and the first 45° splice. It is suppressed as $\Delta\gamma_2$ becomes small for the second splice which is located on the opposite side of the loop. The same applies to $\Phi_{\tau_{02}}$ which is suppressed by diminishing $\Delta\gamma_1$. More careful inspection indicates that each nearly 45° splice induces two polarization errors which cancel if the splice on the opposite side of the loop is perfectly equal to 45°.

Another unexpected result is that the largest polarization errors which exist in the depolarized gyroscope involve the combinations of delays τ_{01} or τ_{02} with the delays induced in the depolarizers, τ_1 and τ_2. In this case, the equations are

$$\Phi_{\tau_{01},\tau_1,\tau_2} \propto k_0\,\varepsilon\ \Gamma(\tau_{01} \pm i\tau_1 \pm j\tau_2)\ ,$$
$$\Phi_{\tau_{02},\tau_1,\tau_2} \propto\ k_0\,\varepsilon\ \Gamma(\tau_{02} \pm i\tau_1 \pm j\tau_2)\ ,$$

where $i = 0$ or 1, $j = 0$ or 1 and $i + j \neq 0$. Since the depolarizer delays can add to or subtract from each other, depending on the polarization evolution in the SM loop, the number of errors is minimized when one depolarizer is twice the length of the other (the difference in length is equal to the length of the shorter depolarizer). It should be noted that the errors listed above are of low order; i.e., they involve only one or two small angles representing a misalignment from 0° or 45°. Lower order polarization errors are few in number and, therefore, relatively easy to control by judicious choice of delays in the optical circuit. In order to achieve navigation grade performance all polarization errors have to be smaller than $\approx 10^{-8} rad$ of optical phase shift and higher order error terms like $k_1 k_2 k_3 \varepsilon \Gamma(\tau)$, where the k's represent misalignment from 0° or 45°, also have to be considered ($k_i \approx 10^{-1.5}, \varepsilon \approx 10^{-3}$). These higher order error terms are smaller than the lower order terms, but there are many more of them. Thus controlling them becomes more difficult. These challenges have nevertheless been overcome in our depolarized design, providing for the navigation grade bias stability illustrated in section 3.

6 Summary and Conclusion

The FOG technology continues to make rapid progress toward satisfying higher accuracy requirements, ranging from 1°/hr AHRS systems to the most stringent of precision applications. The open loop FOG is in production for commercial aircraft. Results show excellent performance that is progressing with time. On the other end of the spectrum, the precision FOG is showing promise for meeting even the most demanding applications, with bias stability better than 0.0003°/hr and an angle random walk of around $0.00019°/\sqrt{\text{hr}}$. FOGs are also showing that they can meet inertial navigation requirements over severe environments. The depolarized FOG is looking remarkably attractive for this marketplace, showing rapidly-improving performance over environments in a very cost-competitive design. Because of this progress and the growing

interest in navigation grade depolarized designs, we have presented a relatively comprehensive treatment aimed at understanding the unique behavior of polarization phenomena in depolarized FOGs. The model addressed the issues of signal fading, wavelength stability, classical polarization errors and random noise. In addition, a mathematical treatment was presented for the first time (to our knowledge) on the issues of non-reciprocal birefringence effects. The suppression of these effects is essential for realizing a high performance depolarized FOG. While the analysis points out the challenges of this technology, its final result is consistent with the experimental results: a navigation grade depolarized FOG is realizable.

7 Acknowledgments

This chapter presents accomplishments of Fiber Optic Sensor Group at Honeywell Technology Center of Phoenix, AZ, USA, and Professor James Blake of Texas A&M University. This work was supported, in part, by the United States Defense Advanced Research Projects Agency under the GPS Guidance Package Program.

8 Acronyms

AHRS	Attitude Heading Reference System
ARW	angle random walk
FLS	fiber light source
FOG	fiber optic gyroscope
GPS	Global Positioning System
IOC	integrated optics chip
LED	laser-emitting diode
PM	polarization maintaining
PMD	polarization mode dispersion
RIN	relative intensity noise
RWC	random walk coefficient
SAARU	Standby Attitude and Air Data System
SLD	superluminescence diode
SM	single mode
SOP	state of polarization

9 REFERENCES

[1] H. Lefèvre, *The fiber-optic gyroscope*, Boston–London, Artech House, 1993.

[2] G. A. Sanders, B. Szafraniec, R.-Y. Liu, C. Laskoskie, L. Strandjord and G. Weed, "Fiber optic gyros for space, marine and aviation applications," *SPIE Proceedings*, Vol. 2837, Fiber Optic Gyros: 20th Anniversary Conference, pp. 61-71, 1996.

[3] H. Kajioka, T. Kumagui, H. Nakai, T. Dohsho, H. Soekawa, T. Yuhare, "Commercial applications of mass-produced fiber optic gyros," *SPIE Proceedings*, Vol. 2837, 18-23, 1996.

[4] H. Lefevre, P. Martin, J. Morisse, P. Simonpietri, P. Vivenot, H.J. Arditty, "High dynamic range fiber gyro with all-digital processing," *SPIE Proceedings*, Vol. 1367, pp. 70-80, 1990.

[5] R. -Y. Liu, T.F. El-Wailly, M. Olson, G.W. Adams, J.E. Carroll, "Miniaturized, low- cost FOG IMU for tactical weapon applications," *SPIE Proceedings*, Vol. 2837, pp. 82-92, 1996.

[6] G. Pavlath, "The LN200 fiber gyro based tactical grade IMU," *Proc. of Guidance, Nav. and Control*, AIAA, pp. 898-904, 1993.

[7] K. Sakuma, "Application of fiber optic gyros at JAE," *SPIE Proceedings*, Vol. 2837, pp. 72-79, 1996.

[8] M. Ribes, G. Spahlinger, M.W. Kemmler, "0.1 deg/hr DSP-controlled fiber optic gyroscope," *SPIE Proceedings*, Vol. 2837, pp. 199-207, 1996.

[9] J. N. Blake, J. R. Feth, J. A. Cox, R. Goettsche, "Design and test of a production open loop all-fiber gyroscope," *SPIE Proceedings*, Vol. 1169, 1989.

[10] H. C. Lefevre, J. P. Bettini , S. Vatoux and M. Papuchon, "Progress in optical fiber gyroscopes using integrated optics," *NATO/AGARD Conference Proceedings*, Vol. 383, 9A-1, 9A-13, 1985.

[11] H. Lefevre, Ph.D. Graindorge, H. Arditty, S. Vatoux, M. Papuchon, "Double closed- loop hybrid fiber gyroscope using digital phase ramp," *Proc. of OFS 3*, SOA/IEEE, Post Deadline paper 3, 1985.

[12] A. Ebberg and G. Schiffner, "Closed-loop fiber-optic gyroscope with a sawtooth phase-modulated feedback," *Optics Letters*, 10, pp. 300-302, 1985.

[13] R.A. Bergh, "Dual-ramp closed-loop fiber-optic gyroscope," *SPIE Proceedings*, Vol. 1169, pp. 429-439, 1989.

[14] A. Cordova, R. Patterson, J. Rahu, L. Lam, D. Rozelle, "Progress in navigation grade FOG performance," *SPIE Proceedings*, Vol. 2837, pp. 207-217, 1996.

[15] C. E. Laskoskie, B. Szafraniec and W. Trammell, "Depolarized interferometric fiber-optic gyro with improved polarization error suppression," *SPIE Proceedings*, Vol. 1795, Fiber Optic and Laser Sensors X, pp. 118-125, 1992.

[16] K. Böhm, P. Marten, K. Petermann, E. Weidel and R. Ulrich, "Low-drift fibre gyro using a superluminescent diode," *Electronics Letters*, Vol. 17, No. 10, pp. 352-353, 1981.

[17] R. Ulrich, "Polarization and Depolarization in the Fiber-Optic Gyroscope", *Fiber-Optic Rotation Sensors and Related Technologies*, pp. 52-77, 1982.

[18] B. Szafraniec, J. Feth, R. Bergh and J. Blake, "Performance Improvements in Depolarized Fiber Gyros," *SPIE Proceedings*, Vol. 2510, Fiber Optic and Laser Sensors XIII, pp. 37-48, 1995.

[19] K. Hotate and K. Tabe, "Drift of an optical fiber gyroscope caused by the Faraday effect; Influence of earth's magnetic field," *Appl. Opt.*, Vol. 25, pp. 1086-1092, 1986.

[20] J. N. Blake, "Magnetic field sensitivity of depolarized fiber optic gyros," *SPIE Proceedings*, Vol. 1367, Fiber Optic and Laser Sensors VIII, pp. 81-86, 1990.

[21] C. D. Poole, J. H. Winters and J. A. Nagel, "Dynamic Equation for Polarization Dispersion," *Optics Letters*, Vol. 16, No. 6, pp. 372-374, 1991.

[22] K. Killian, M. Burmenko, W. Holliger,"High performance fiber-optic gyroscope with noise reduction," *SPIE Proceedings*, Vol. 2292, July, 1994.

[23] D. C. Hall, W. K. Burns and R. P. Moeller, "High-Stability Er^{3+}-Doped Superfluorescent Fiber Sources," *Journal of Lightwave Technology*, Vol. 13, No. 7, pp. 1452-1460, July, 1995.

[24] H. J. Patrick, A. D. Kersey, W. K. Burns and R. P. Moeller, "An Erbium-doped Superfluorescent Fiber Source with Long Period Fiber Grating Wavelength Stabilization," *OSA Proceedings*, Vol. OFS-16, Fiber Optic Sensors, pp. 138-141, 1997.

[25] P. R. Morkel, R. I. Laming, and D. N. Payne, "Noise characteristics of high-power doped fiber superluminescent source," *Electron. Lett.*, Vol. 26, pp. 96-98, 1990.

[26] R. P. Moeller and W. K. Burns,"1.06–μm all-fiber gyroscope with noise subtraction," *Optics Lett.*, Vol. 16, pp. 1902-1904, 1991.

[27] R. C. Jones, "A New Calculus for the Treatment of Optical Systems. VII. Properties of the N-Matrices, " *JOSA*, Vol. 38, No. 8, pp. 671-685, 1948.

[28] G. N. Ramachandran, S. Ramaseshan, "Crystal Optics," a section in S. Flügge, ed., *Encyclopedia of Physics*, Vol. 25/1, Berlin, Springer, 1961

[29] W. A. Shurcliff, *Polarized Light*, London, Oxford University Press, 1962.

[30] M. V. Klein and T. E. Furtak, *Optics*, New York: Wiley, 1986.

[31] J. Blake and B. Szafraniec, "Polarization Impulse Response Function of a Single Mode Fiber Gyro Coil," *SPIE Proceedings*, Vol. 2837, Fiber Optic Gyros: 20th Anniversary Conference, pp. 295-302, 1996.

[32] R. Usui, A. Ohno and K. Okada, "Intermediate Grade FOG with Single Mode Fiber Coil," *SPIE Proceedings*, Vol. 2070, Fiber Optic and Laser Sensors XI, pp. 152-163, 1993.

[33] J. Blake, B. Szafraniec and J. Feth, "Partially polarized fiber optic gyro," *Optics Letters*, Vol. 21, No. 15, pp. 1192-1194, 1996.

[34] T. Saida and K. Hotate, "General Formula Describing Drift of Interferometric Fiber Optic Gyro due to Faraday Effect: Reduction of the Drift in Twin-depo-FOG," *OSA Proceedings*, Vol. OFS-16, Fiber Optic Sensors, pp. 134-137, 1997.

[35] J. Blake and B. Szafraniec, "Random Noise in PM and Depolarized Fiber Gyros," *OSA Proceedings*, Vol. OFS-16, Fiber Optic Sensors, pp. 122-125, 1997.

[36] J. W. Goodman, *Statistical Optics*, New York: Wiley, 1985.

[37] G. A. Sanders, B. Szafraniec, L. Strandjord, R. Bergh, A. Kaliszek, R. Dankwort, C. Lange and D. Kimmel, "Progress in High Performance Fiber Optic Gyroscopes," *OSA Proceedings*, Vol. OFS-16, Fiber Optic Sensors, pp. 116-121, 1997.

[38] E. C. Kintner,"Polarization control in optical-fiber gyroscopes," *Optics Letters*, Vol. 6, No. 3, pp. 154-156, 1981.

[39] S. L. A. Carrara, B. Y. Kim and H. J. Shaw, "Bias drift reduction in polarization-maintaining fiber gyroscope, " *Optics Letters*, Vol. 12, pp. 214-216.

[40] G. H. Chin, A. Cordova and E. L. Goldner, "Extended Environmental Performance of Attitude and Heading Reference Grade Fiber Optic Rotation Sensors," *Fiber Optic and Laser Sensors VIII*, SPIE Vol. 1367, pp. 107-120, 1990.

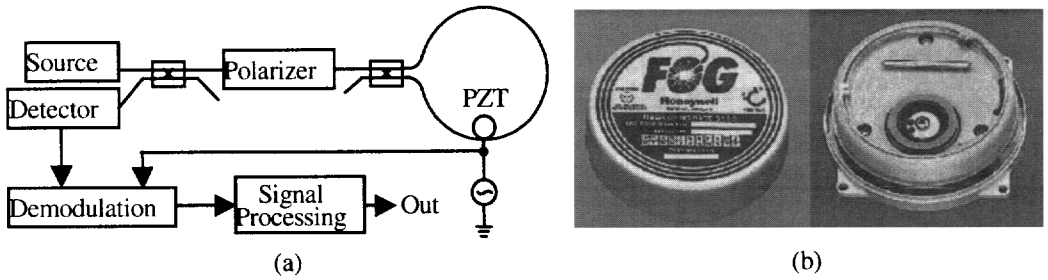


Figure 1: Honeywell's all-fiber open-loop FOG design. a) schematic, b) photograph

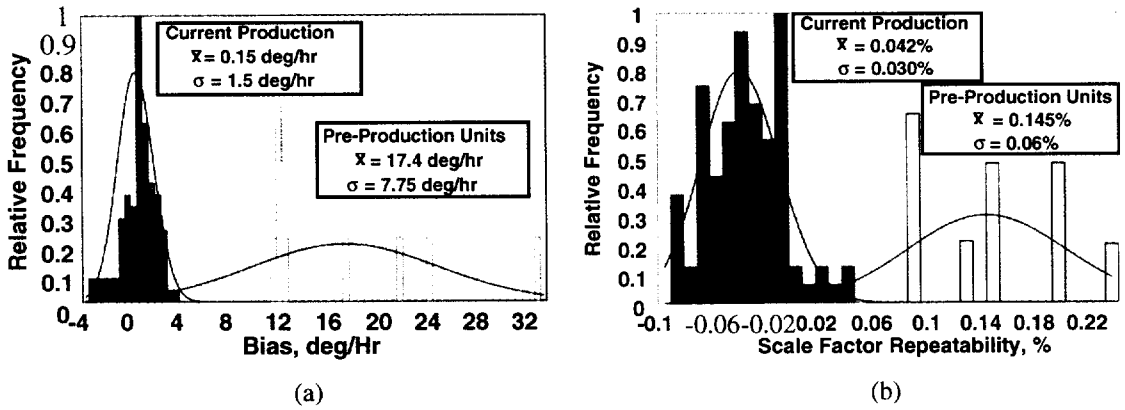


Figure 2: Open-loop FOG design comparing current performance with that of the preproduction units. a) bias stability, b) scale factor repeatability

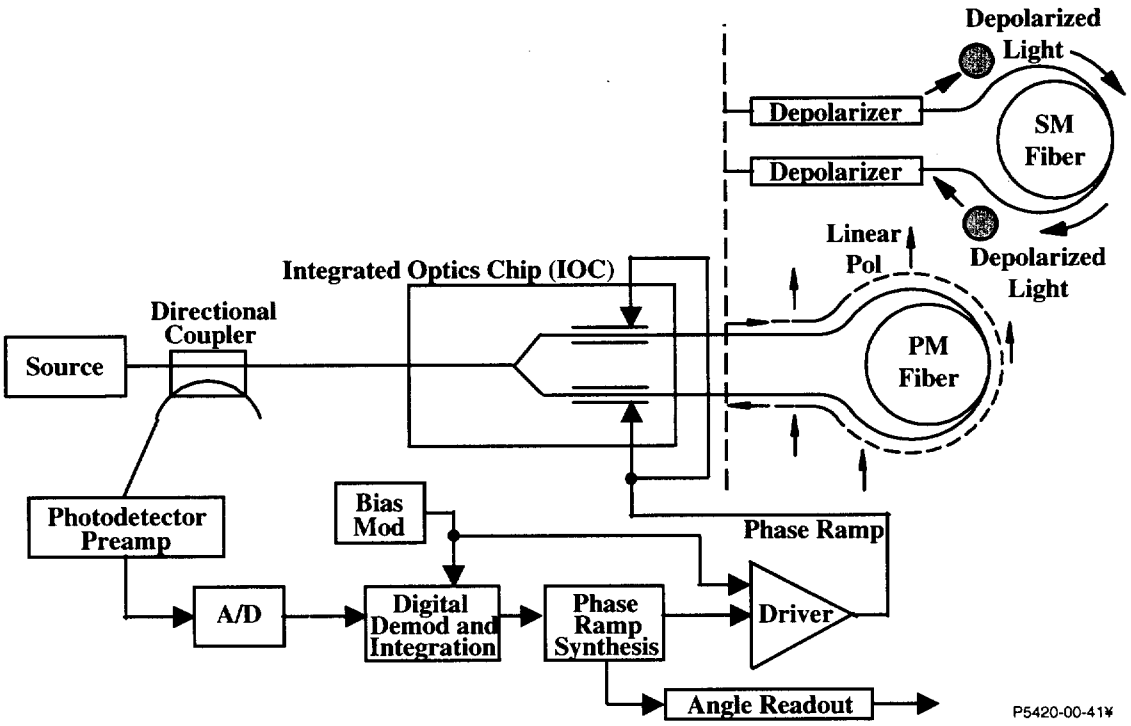


Figure 3: Schematic of typical closed-loop navigation-grade FOG utilizing a standard PM fiber coil or SM coil

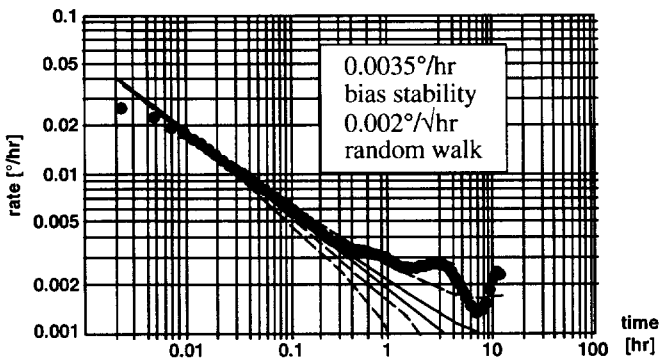


Figure 4: Rate uncertainty versus integration time for a bias stability test of Honeywell's navigation-grade depolarized FOG design

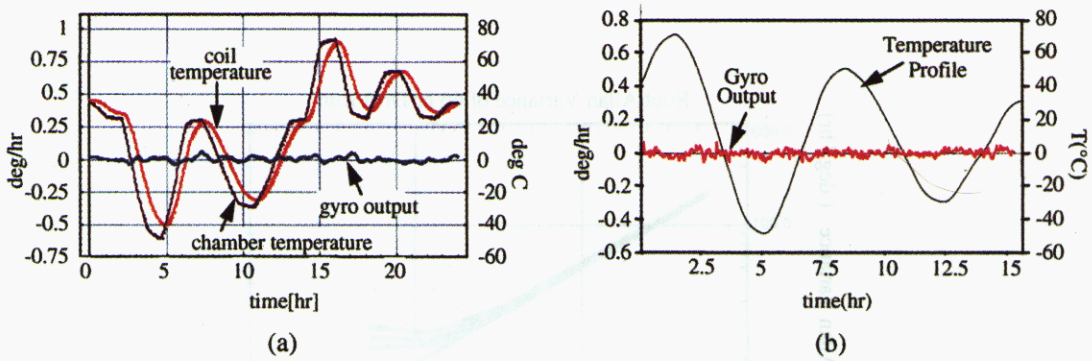


Figure 5: Thermal response of Honeywell depolarized FOG

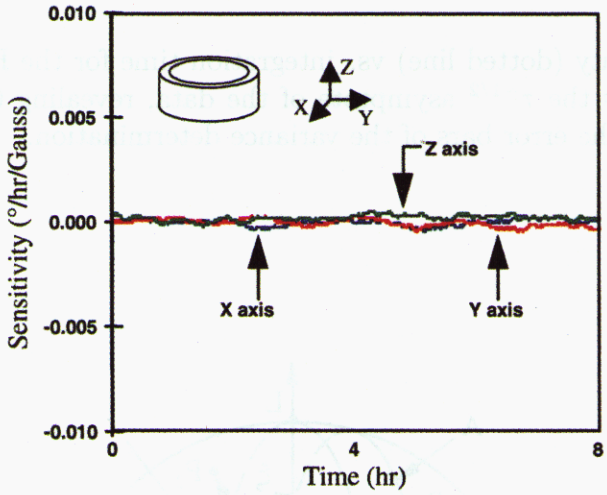


Figure 6: Magnetic sensitivity of navigation grade depolarized FOG

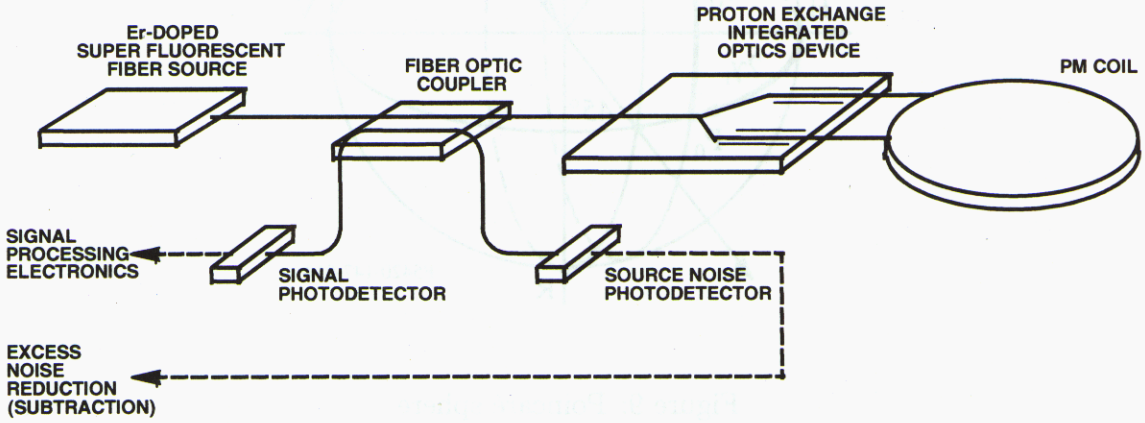


Figure 7: Typical precision FOG design

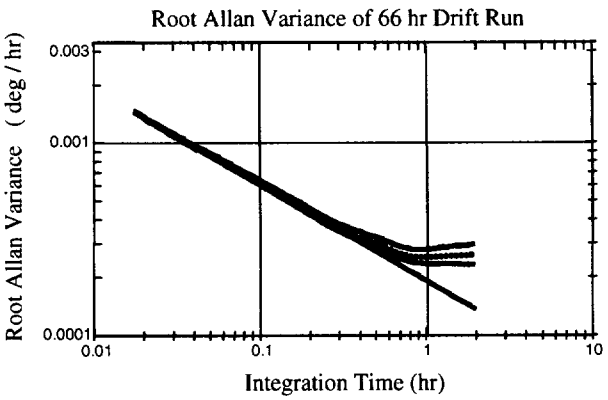


Figure 8: Rate uncertainty (dotted line) vs. integration time for the Honeywell precision FOG. The solid straight line is the $\tau^{-1/2}$ asymptote of the data, revealing $0.00019°/\sqrt{hr}$ ARW. The curved solid lines show the error bars of the variance determination.

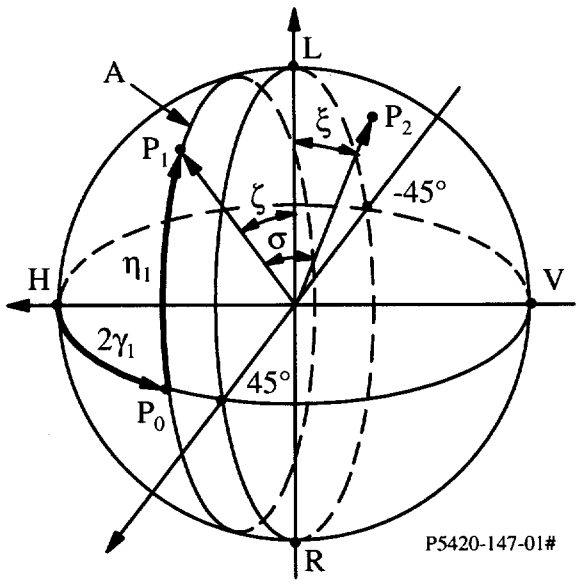


Figure 9: Poincaré sphere

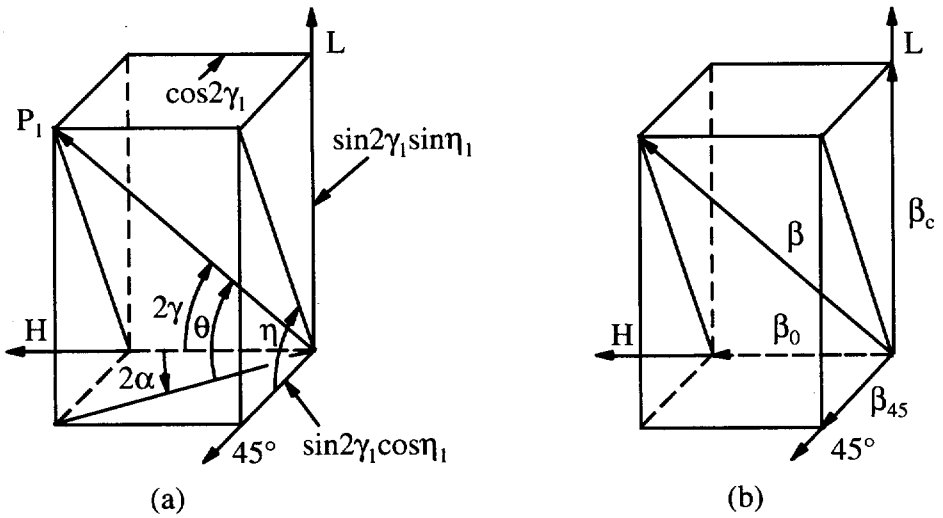


Figure 10: Poincaré sphere description of a) polarization, b) birefringence

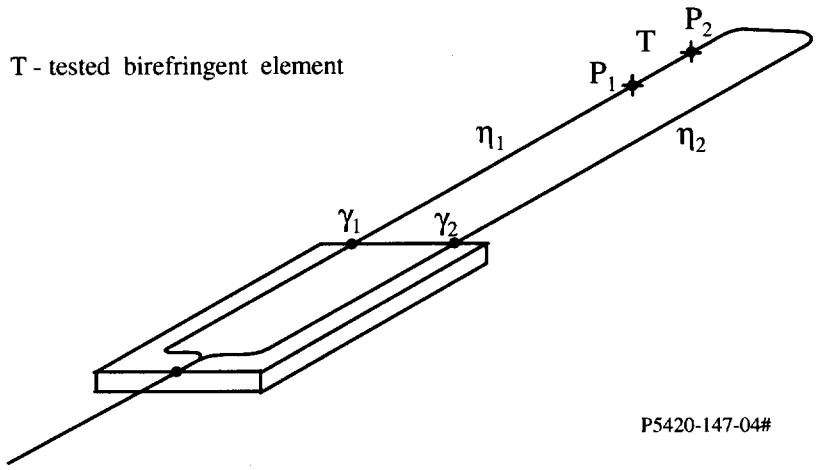


Figure 11: Sagnac interferometer with IOC and arbitrary polarization states P_1 and P_2

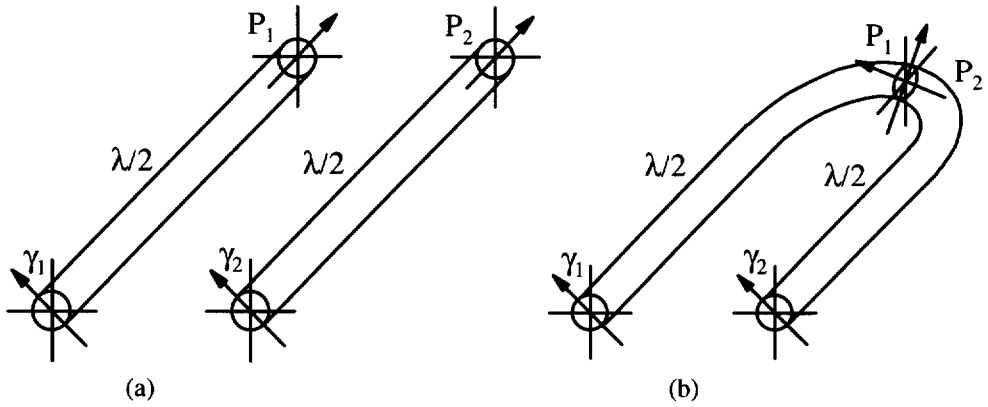


Figure 12: Polarization states P_1 and P_2 exiting IOC leads. a) leads open, b) leads joined into the loop

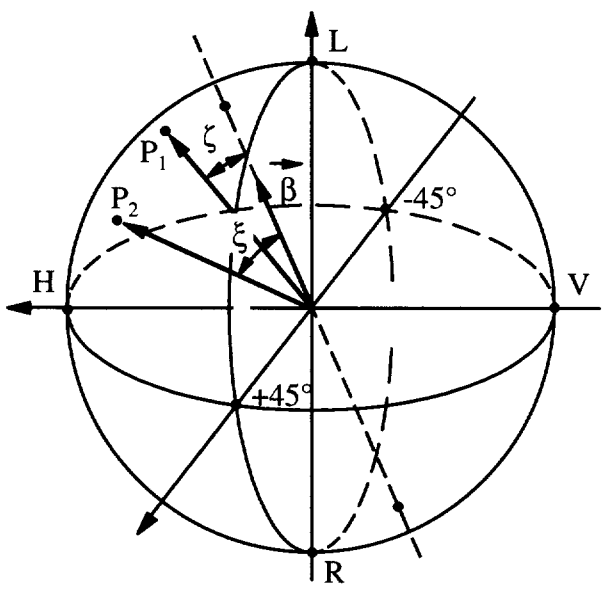


Figure 13: Sensitivity of polarization states P_1 and P_2 to arbitrary non-reciprocal birefringence

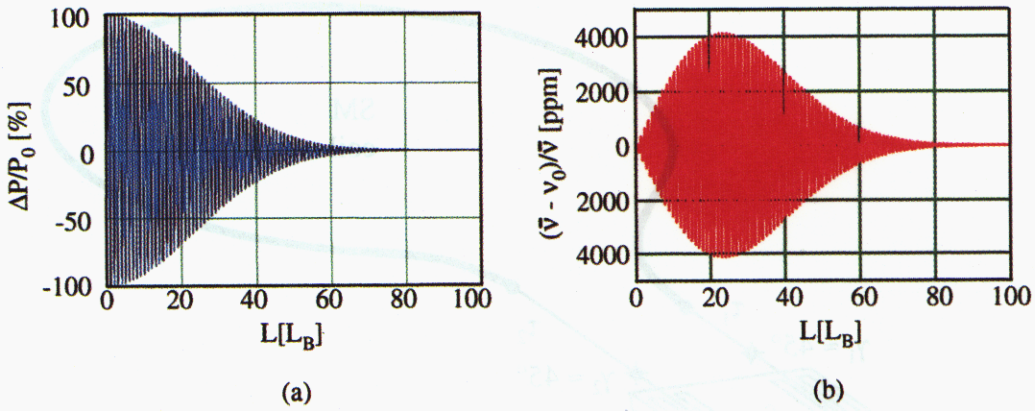


Figure 14: Fluctuations in depolarized gyro as function of depolarizer length. a) detected optical power, b) mean optical frequency

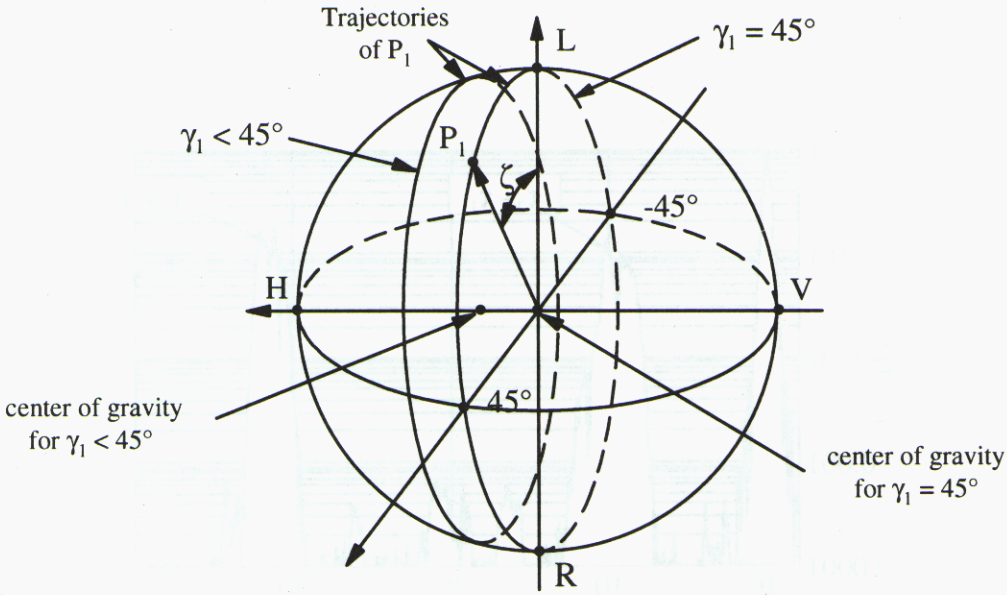


Figure 15: Distribution of polarization states in depolarized FOG using the broadband light source and Poincaré sphere interpretation of averaging processes

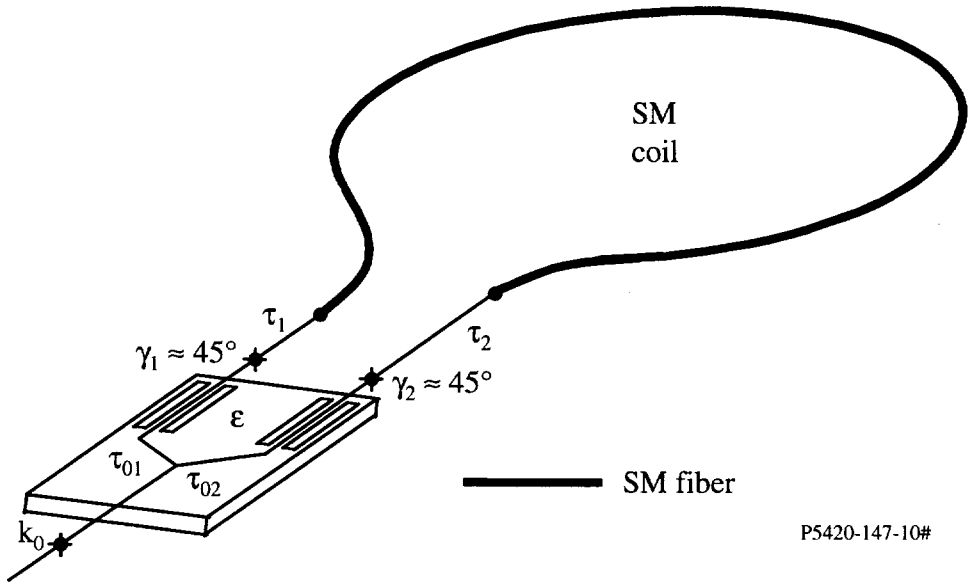


Figure 16: Low-drift configuration of depolarized gyro

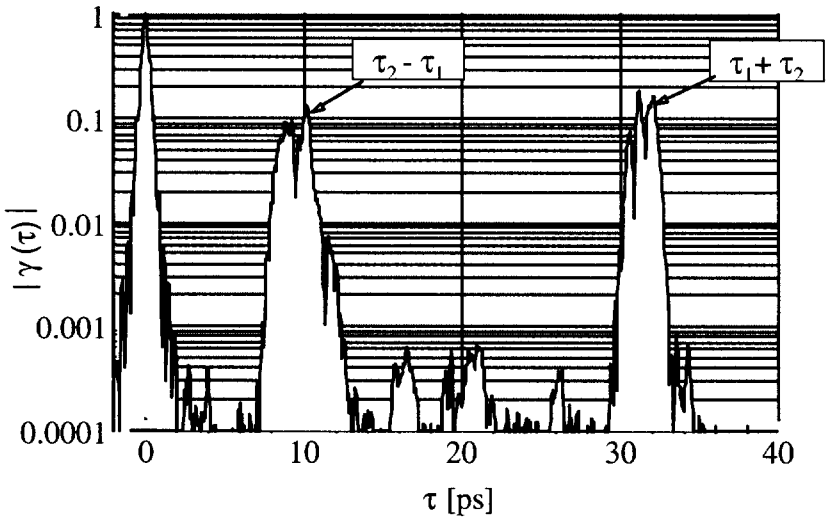


Figure 17: Coherence function of depolarized light exiting the gyro circuit

12. DYNAMIC RING LASER GONIOMETER

Yu.V.Filatov
D.P.Loukianov
P.A.Pavlov
M.N.Burnashev
Autonomous Navigation Department
St.Petersburg State Electrotechnical University
Russia
R.Probst
Physikalisch-Technische Bundesanstalt
Braunschweig, Germany

1.PRINCIPLE OF THE RING LASER GONIOMETER

1.1. Short history of dynamic laser goniometer (DLG)

In the mid-eighties, after more than ten years of research and design work, the performance of the ring laser gyro had been enhanced such that it could be used as a sensor in inertial navigation systems implemented in aircraft and other moving objects [1,2]. On the other hand, it became clear at a very early stage of the ring laser (RL) development that the RL can be successfully used in angle measuring technique as a tool offering high resolution and accuracy in the full angular range of 360 deg. In a French patent [3], Catherin and Dessus first suggested the goniometer with the RL as a sensor. The purpose of implementing the RL in a goniometer is the same as that of a circular scale in a conventional goniometer. Owing to the features of the RL, it can be used for circular measurements in a dynamic mode of operation with high speed and accuracy and offers the chance of automating the whole process of measurement. With the RL goniometer it is, for example, possible to carry out the calibration of a circular optical encoder with a number of lines of several thousands within a few seconds. As the main metrological characteristic, the RL offers the advantage of a nearly perfect circular division as it is based on the laser wavelength. It may therefore be used as an angle standard just as the linear laser as a length standard.

To use the ring laser as a laser gyro in navigation systems, it is a prerequisite to eliminate the lock-in effect; to achieve this, some special means (for example the dithering of the RL) should be used. In the goniometer the RL is continuously rotated at a constant rate (so-called "rate-bias" technique) so that, in this case, the RL does not need any additional means to avoid the lock-in effect. This mode of operation is also an important advantage of using the RL in a goniometer, because it enables the RL scale factor to be continuously calibrated against the angular reference of 360 degrees.

The first realization of the RL goniometer was made at the Metrology Institute VNIIM (St.Petersburg) by the authors of [4], to measure an optical polygon angle standard with an accuracy of about 0.5 arcs. The first commercial RL goniometer GS-1L was then produced by the company "Arsenal" in Kiev [5] in the early eighties. This system is employed in the countries of the former Soviet Union and at some other places, for example at the Slovakian Metrology Institute, Bratislava [6], as a standard of the plane angle unit. The further development of the RL goniometer was carried out by some companies and universities [7,8], particularly at the St.Petersburg Electrotechnical University (SPEU) where such systems were developed, for example, for noncontact measurements of an object's angular position, producing circular scales and other applications[9,10]. The developments by the SPEU ultimately brought forth the commercial laser goniometer system EUP-1L.

1.2. Principle of measurement

As it was shown above the frequency, detected as frequency f of the RL's sinusoidal output signal, is proportional to the rotation rate or instantaneous angular velocity $\Omega = d\varphi/dt$ in the inertial space:

$$\omega = K(\Omega)\cdot\cos\theta\cdot\Omega \tag{12.1}$$

where $K(\Omega)$ is the RL scale factor which slightly depends (if Ω is far enough from lock-in zone) on the rotation rate and θ is the angle formed by the normal to the RL cavity plane (measuring axis) and the rotation axis. In a first-order approximation, K is

$$K \cong 2\pi\frac{4\cdot S}{\lambda\cdot L} \tag{12.2}$$

where S and L are the RL area and perimeter and λ is the RL wavelength.

Integration of (12.1) over a time interval 0....t yields

$$2\pi\cdot N_{\varphi} = K\cdot\cos\theta\cdot\varphi, \tag{12.3}$$

where N_φ is the number of periods of the RL output signal and φ is the angle of the RL rotation during the integration time (it is assumed here that the dependence of K on Ω may be neglected). From (12.3) it follows that the angle of rotation φ is determined by the number of periods N_φ of the RL output signal together with K and $\cos\theta$:

$$\varphi = 2\pi \frac{N_\varphi}{K \cdot \cos\theta} \tag{12.4}$$

It follows from (12.4) that the RL scale factor K can be easily determined from the number of signal periods during one revolution $N_{2\pi}$ ($\varphi = 2\pi$):

$$K = \frac{N_{2\pi}}{\cos\theta} \tag{12.5}$$

For the angle φ, (12.4) and (12.5) give the basic relation

$$\varphi = 2\pi \frac{N_\varphi}{N_{2\pi}} \tag{12.6}$$

For simplicity, we assume for the following $\cos\theta$ to be equal to one. From (12.5) we get the angle interval $\Delta\varphi_{RL}$ corresponding to one signal period of the RL which is equal to

$$\Delta\varphi_{RL} = 2\pi/K \tag{12.7}$$

Fig.12.1 shows the principle of operation of the RL goniometer for the calibration of any type of angle converter. It includes the ring laser 1, the rotary table 2, the spindle with the drive 3, the angle converter consisting of the rotating 4 and the non-moving 5 parts, indicating the angular positions of the rotary table, and the counting unit 6. The devices 4 and 5 may, for example, be represented by an optical polygon together with an optical null indicator or by a circular encoder disc and its reading head.

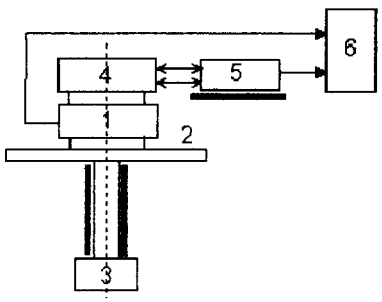


Fig.12.1. Schematic of the ring laser goniometer.
1 ring laser, 2 rotary table, 3 spindle with drive, 4 turning and 5 non-moving parts of angle converter, 6 counter.

1.3. Comparison of dynamic and static methods of measurements

At least two measuring routines may be used to realise this schematic. While providing the first routine the rotary table stops in a certain angle positions defined by the angle converter. The RL output signal is integrated during the interval between two non-moving positions of the table. This method may be called static [11]. While running the second procedure the rotary table is always rotating with the quasi-constant rate. Certain angular positions of the table are fixed in dynamic by choosing pulses of the angle converter output signal. The RL output signal is integrated in that case between the chosen pulses. Such method is usually called dynamic.

Static and dynamic procedures have some principle differences concerning not only the rotary table type of movement, but also some peculiarities of the RL's work. These peculiarities determine precision of the methods and their appropriate application areas.

First we'll discuss the static method more in detail. Here, the RL moving together with the rotary table from one position to another rests at every of them long enough to fix theangular position. Typical curve of the platform angle movement one can see in fig.12.2 [11}

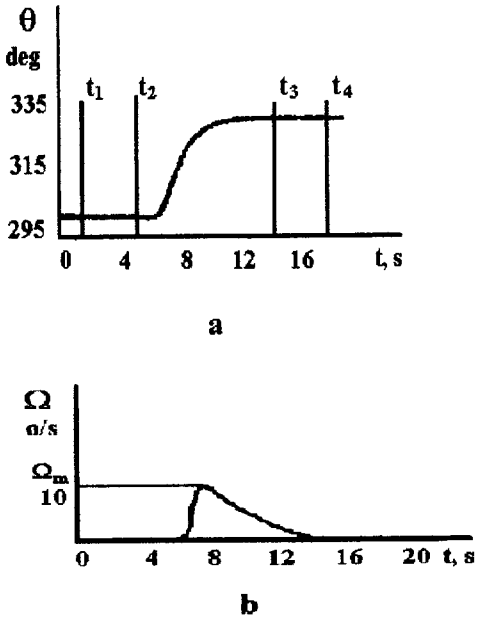


Fig.12.2. The platform angle movement curve in static procedure of angle measurements.

As it is shown in fig.12.2a during the time period t_1, t_2 the rotary table is in the stage of the angular position determination. After that the table moves to the new angular position in time period $t_3 - t_2 \approx 11$ s. The table's rotation rate increases quickly from 0 to the maximum value Ω_{max} and then slowly decreases to 0 at the moment t_3 as it's shown in fig.12.2b.

So the RL moves from one stable angular position to another together with the rotary table with a big change of the rotation rate. In the case represented in fig.12.2b the rotation rate of the table changes from 0 to 12°/s. Wide range of angular velocities, including the value close to 0, requires to bias the RL away from the rotation rates near to the lock-in zone. This may be achieved by using either a technique of dither or putting some nonreciprocal elements in the RL cavity [1]. The first decision is used normally now.

Long phases of the table non-movement and relative long rotation intervals causes increasing of the measuring time and in turn the necessity of taking into account the drift of the RL. Determination of the drift parameters takes normally long time and it is being made especially before the measurements. To determine the drift parameters in the work [11] the RL was placed on a stable base, it's measuring axis was combined with the local vertical and it's output signal was read with frequency of 1 Hz in order to estimate the vertical component of the Earth rotation rate Ω_E and compare it with the known value for the given latitude. The constant component of the drift was estimated by measuring several times the vertical component of Ω_E during several weeks, then averaging and comparing with the known value. In [11] for the RL of GG1342 type produced by Honeywell the constant component of the drift appeared to be equal 0.04 °/h. Several results of long (about 8h) tests of the RL are represented in fig.12. 3.

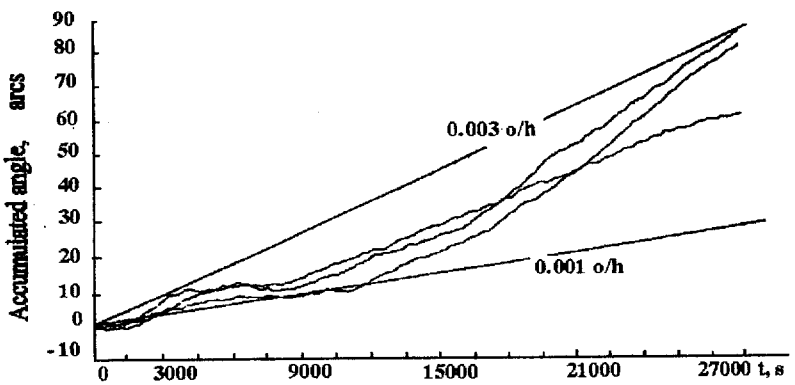


Fig.12. 3. Some results of the RL (GG1342 Honeywell model) tests

These results allow to estimate instability of drift from switch on to switch on and the random component of the drift (Random Walk coefficient (RW)). „Angular movement" of the RL in relation to the inertial space is plotted on ordinate in the figure. The earth rotation and the constant component of the drift are subtracted from the RL signal. Evaluation of the data represented in fig.12.3 showed that the drift instability from switch on to switch on is about $3 \cdot 10^{-3}$ °/h, and RW isn't bigger than $6 \cdot 10^{-3}$ °/ √h.

Those parameters are very important for the static procedure of measurement because the unknown and so uncompensated drift component may cause big errors while measuring time is long enough (about one minute).

The analysis shows that the measurement error is caused mostly by the random part of the drift (Random Walk), which appears because of quantum noise and using dithering in the RL [1]. The total error of the laser goniometer using static procedure is about 0.5 arcs and that is too much for calibration of precise converters and in some other cases of application. The reasons of this error are non-moving phases in measurements, use of the dithering and some other factors. That is why the dynamic procedure of the Ring Laser Goniometer, free from those shortcomings, seems to be more promising.

Dynamic method requires the rotary table to rotate constantly with the rate big enough to operate the RL on the linear part of it's output response. Normally this rate is in the range of 36...360 °/s. That makes possible not to use the technique of dither. So we get rid of the noise caused by passing through the lock-in zone and the RL random drift becomes much smaller. In this case it is possible to approach to the potential limit of the RL accuracy.

The most important advantage of the dynamic method is the possibility to calibrate the RL using a natural standard 2π at every revolution of the table (during every 1...10s). That allows to compensate errors caused by instability of the RL scale factor and drift components which have usually a period of changing longer than the DLG table revolution time.

Let us look at the DLG operation more in detail. Fig.12. 4 shows it's signal generation, where U_{RL} is the RL output signal and U_{AC} the angle converter output signal. The periods of the RL output signal are summed alternately by two counters C1 and C2 within time intervals formed by the output pulses of the angle converter. So $N^j_{\varphi_i}$ is the result of summing over the *i*-th time interval between t_i and t_{i+1}, which, according to (12.2), yields:

$$N^j_{\varphi_i} = \frac{1}{2\pi} \int_{t_{i-1}}^{t_i} \omega(t) \cdot dt, \tag{12.8}$$

where *j* is the number of revolutions of the platform from the beginning of the measurement.

The actual value of the scale factor K is derived according to (12.5) as the number of periods of the RL output signal during one revolution $N_{2\pi}$. Considering that this number may change in time due to scale factor drift, noise and other effects, it is useful to measure it as often as possible, best at every revolution. Also, the contribution of the earth rotation to the scale factor must be regarded. This, however, can be kept constant with a sufficiently constant rotation rate of the goniometer spindle or simply be corrected through a simultaneous measurement of time.

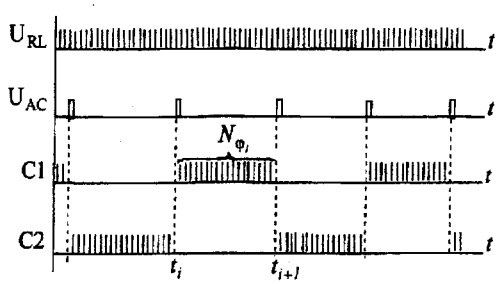


Fig. .12.4. Schematic of the signal generation of the DLG. U_{RL} - ring laser output signal, U_{AC} - angle converter output signal, C1 and C2 - counting results of two counters.

According to (12.6) and (12.8), the procedure of angle measurement during J revolutions of the table may be written as

$$\varphi_i = \frac{2\pi}{J} \cdot \sum_j^J \frac{\int_{t_{i-1}^j}^{t_i^j} \omega(t) \cdot dt}{\sum_i^n \int_{t_{i-1}^j}^{t_i^j} \omega(t) \cdot dt}, \tag{12.9}$$

where n is the number of the angle converter intervals in 2π and the denominator terms of (12.9) represent the determination of the scale factor number $N_{2\pi}$ at every revolution.

2. ACCURACY OF DLG

2.1. Model of DLG's errors

To analyse the errors of the DLG let's consider the measuring process more in detail. We will suppose here that due to instability of the scale factor and drift we should consider the numbers N_φ and $N_{2\pi}$ as a random functions depending on time. To operate with these functions we rewrite (12.8) to following form

$$N_\varphi(t) = \frac{1}{2\pi} \int_t^{t+t_\varphi} \omega \cdot dt, \tag{12.10}$$

where $t_\varphi = t_i - t_{i-1}$ - interval of integration or time of the table turn on the angle φ. Here we suppose that t_φ is a constant value. To get expression for an error of measurements let us consider $N_\varphi = N_\varphi^0 + \delta N_\varphi$ and $N_{2\pi} = N_{2\pi}^0 + \delta N_{2\pi}$. Now one can write from (12.6) with a precision up to the first order of $\delta N_{2\pi} / N_{2\pi}^0$

$$\varphi = 2\pi \cdot \frac{N_\varphi^0}{N_{2\pi}^0} + \frac{2\pi}{N_{2\pi}^0} \cdot \delta N_\varphi - \frac{2\pi}{(N_{2\pi}^0)^2} \cdot \delta N_{2\pi} \cdot N_\varphi^0. \tag{12.11}$$

The true value of the measured angle φ is equal

$$\varphi_0 = 2\pi \cdot \frac{N_\varphi^0}{N_{2\pi}^0}.$$

Then we will get the error of measurements $\delta\varphi = \varphi - \varphi_0$ as follows

$$\delta\varphi = \frac{1}{N_{2\pi}^0} \cdot (2\pi \cdot \delta N_\varphi - \varphi_0 \cdot \delta N_{2\pi}).$$

Let us write the RL output frequency as

$$\omega(t) = \omega_0 + \delta\omega(t),$$

where we understand $\delta\omega(t)$ as changes of the frequency due to changes of the scale factor and other parameters of the RL (the rotation rate is considered to be constant).

So N_{φ_i} may be written as

$$N_{\varphi_i} = \frac{1}{2\pi} \int_{t_{i-1}}^{t_i} \omega_0 \cdot dt + \frac{1}{2\pi} \int_{t_{i-1}}^{t_i} \delta\omega \cdot dt = N_{\varphi_i}^0 + \delta N_{\varphi_i}.$$

Let us consider $\delta\omega(t)$ as a stationary random process which a correlation function R_ω. One can calculate the correlation function of the number of periods, taking into account that integration is a linear transformation, as

$$R_{N_\varphi}(\tau) = \frac{1}{(2\pi)^2} \int_t^{t+t_\varphi} \int_{t+\tau}^{t+t_\varphi+\tau} R_\omega(u-v) \cdot du \cdot dv \tag{12.12}$$

and dispersion of the random function δN_φ as

$$D_{N_\varphi} = \frac{1}{(2\pi)^2} \int_t^{t+t_\varphi} \int_t^{t+t_\varphi} R_\omega(u-v)\cdot du\cdot dv. \tag{12.13}$$

From equations written above it is possible to calculate the correlation function and dispersion of the measured angle, taking into account that functions $N_\varphi(t)$, $N_{2\pi}(t)$ are correlating to each other,

$$R\varphi = (\frac{2\pi}{N^0_{2\pi}})^2 \cdot R_{N_\varphi} + (\frac{\varphi_0}{N^0_{2\pi}})^2 \cdot R_{N_{2\pi}} - \frac{\varphi_0 \cdot 2\pi}{(N^0_{2\pi})^2} \cdot (R_{N_\varphi N_{2\pi}} + R_{N_{2\pi} N_\varphi}), \tag{12.14}$$

$$D_\varphi = (\frac{2\pi}{N^0_{2\pi}})^2 \cdot D_{N_\varphi} + (\frac{\varphi_0}{N^0_{2\pi}})^2 \cdot D_{N_{2\pi}} - \frac{\varphi_0 \cdot 2\pi}{(N^0_{2\pi})^2} \cdot (D_{N_\varphi N_{2\pi}} + D_{N_{2\pi} N_\varphi}). \tag{12.15}$$

Now we will calculate the value of the cross-correlation moment $D_{N_\varphi N_{2\pi}}$ from the cross-correlation function of processes $N_\varphi(t), N_{2\pi}(t)$.

$$R_{N_\varphi N_{2\pi}} = M[\delta N_\varphi(t)\cdot \delta N_{2\pi}(t)] = \frac{1}{(2\pi)^2} \int_t^{t+T} \int_t^{t+\tau+t_\varphi} R_\omega(u-v)\cdot du\cdot dv, \tag{12.16}$$

$$D_{N_\varphi N_{2\pi}} = \frac{1}{2\pi} \int_t^{t+T} \int_t^{t+t_\varphi} R_\omega(u-v)\cdot du\cdot dv. \tag{12.17}$$

So the equation (12.15) together with equations (12.13) and (12.17) (equation for $D_{N_{2\pi}}$ one can get from (12.13) by changing t_φ to time of revolution T, and for $D_{N_{2\pi} N\varphi}$ from (12.17) by exchanging upper limits of integration) determine the dispersion of the angle measurements with stationary fluctuations of the frequency of the RL output signal. In this case it is necessary to know the correlation function of the RL output frequency.

Usually as a model of variations of the RL output frequency they use the wide band fluctuations with the following correlation function

$$R_\omega(\tau) = D_\omega \cdot e^{-\alpha|\tau|}. \tag{12.18}$$

Let us analyse the angle measurement error in that case. One will get dispersion of the number of periods by substituting equation for $R_\omega(\tau)$ in (12.13) and making the integration.

$$D_{N_\varphi} = \frac{2D_\omega \cdot t_\varphi}{\alpha\cdot(2\pi)^2} \cdot [\frac{1}{\alpha\cdot t_\varphi}\cdot(e^{-\alpha t_\varphi} - 1) + 1]. \tag{12.19}$$

We will consider fluctuations of ω to have wide band enough for $\alpha \cdot t_\varphi >> 1$. Then

$$D_{N_\varphi} \approx \frac{2D_\omega \cdot t_\varphi}{\alpha\cdot(2\pi)^2}. \tag{12.20}$$

In the same way for $\delta N_{2\pi}(t)$ considering $\alpha \cdot T >> 1$ one gets

$$D_{N_{2\pi}} \approx \frac{2D_\omega \cdot T}{\alpha\cdot(2\pi)^2}. \tag{12.21}$$

The cross-correlation moment $D_{N_\varphi N_{2\pi}}$ can be obtained by substituting (12.18) into (12.17) and integrating according to appropriate rules

$$D_{N_\varphi N_{2\pi}} = D_{N_{2\pi} N_\varphi} = \frac{2D_\omega \cdot t_\varphi}{\alpha\cdot(2\pi)^2} \cdot [1 + \frac{1}{\alpha\cdot t_\varphi}\cdot(e^{-\alpha t_\varphi} + e^{-\alpha\cdot T} - e^{-\alpha(T-t_\varphi)} - 1)]. \tag{12.22}$$

Considering $t_\varphi < T$ one gets

$$D_{N_\varphi N_{2\pi}} = D_{N_{2\pi} N_\varphi} \approx \frac{2D_\omega \cdot t_\varphi}{\alpha \cdot (2\pi)^2} \tag{12.23}$$

and substituting (12.23), (12.21) and (12.20) into (12.15) one gets the dispersion of the angle measurement as

$$D_\varphi = \frac{2 \cdot D_\omega}{(N^0_{2\pi})^2 \cdot (2\pi)^2 \cdot \alpha} \cdot [2\pi \cdot (2\pi - \varphi_0) \cdot t_\varphi + \varphi_0 \cdot (\varphi_0 \cdot T - 2\pi \cdot t_\varphi)].$$

If the rotation rate is constant then the second term in square brackets is zero, because $t_\varphi / T = \varphi_0 / 2\pi$. Then finally we get

$$D_\varphi = \frac{D_\omega \cdot t_\varphi}{(N^0_{2\pi})^2 \cdot \alpha} \cdot (1 - \frac{\varphi_0}{2\pi}) = \frac{D_\omega \cdot t_\varphi}{(N^0_{2\pi})^2 \cdot \alpha \cdot \Omega} \cdot (1 - \frac{\varphi_0}{2\pi}). \tag{12.24}$$

The dependence of $\sqrt{D_\varphi}(\varphi_0)$ is represented in fig.12.5. Dispersion of the angle measurement is proportional to D_ω and comes to it's maximum when the measured angle is π. It comes to zero when $\varphi_0 \to 2\pi$ due to calibration process.

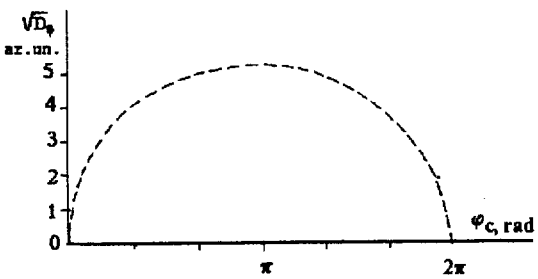


Fig.12.5. Sdandard deviation of angle measurements in the case of wide band fluctuations of the RL output frequency.

2.2. The DLG model accounting the real output response of the RL

The RL output response, connecting output signal frequency ω with the rotation rate Ω, is determined by the equation (12.1) only approximately. In reality it contains also additive members, causing angle measurement errors. Taking into account these members the RL output response may be written as

$$\omega = K_1 \cdot \Omega + K_0 + K_{-1}/\Omega, \tag{12.25}$$

where K_1 - the scale factor, K_0 - represents the bias or null-shift which is responsible for the RL drift, K_{-1} - determines non-linearity of the output response.

Representation of the RL output response in the form of (12.25) is convenient because every coefficient can be determined experimentally. Besides it is logical from the physical point of view to divide ω into 3 parts, because changes of the coefficient K_1 are determined mostly by some equivalent refractive index of the material in the cavity (mostly the active medium). Coefficient K_0 unites all negative non-reciprocal effects. Coefficient K_{-1} is caused by the coupling between counter-propogating waves due to radiation backscattering.

Changes in time of the output response coefficients and distinction of K_0 and K_{-1} from zero (together with the influence of the earth rotation) in the presence of the rotation rate changes cause the error of the angle measurements. Let us analyse these effects in detail.

Let us assume that coefficients of the output response and the rotation rate are as follows

$$K_i = K_{i0} + \delta K_i \tag{12.26}$$

$$\Omega(t) = \Omega_0 + \delta\Omega(t) \tag{12.27}$$

where $\delta K_i(t)$ and $\delta\Omega(t)$ are random processes, K_{i0} and Ω_0 - are constant values, i = 1,0,-1.

We should also keep in mind that the RL is sensitive to the angle movements i n relation to the inertial space and so the

equation for ω accounting the earth rotation (if $\theta = 0$) should be as follows

$$\omega = K_1 \cdot (\Omega + \Omega_E \cdot \sin\gamma) + K_0 + \frac{K_{-1}}{\Omega + \Omega_E \cdot \sin\gamma}, \tag{12.28}$$

where γ - latitude, Ω_E - earth rotation rate.

In such conditions the equation for the number of periods N_φ becomes

$$N_\varphi = \frac{1}{2\pi} \int_t^{t+t_\varphi} [K_1 \cdot (\Omega + \Omega_E \cdot \sin\gamma) + K_0 + \frac{K_{-1}}{\Omega + \Omega_E \cdot \sin\gamma}] \cdot dt'. \tag{12.29}$$

While substituting (12.26) and (12.27) into (12..29) one should consider fluctuations of the time integration interval caused by the rotation rate instability. Let us analyse one part of the equation (12.29)

$$\int_t^{t+t_\varphi} K_0(t') \cdot dt'. \tag{12.30}$$

The integration interval here may be determined according to the measured angle from identity

$$t_\varphi = \frac{\varphi_0}{\frac{1}{t_\varphi} \cdot \int_t^{t+t_\varphi} \Omega(t') \cdot dt'}. \tag{12.31}$$

Substituting (12.31) in (12.30) and expanding in series around Ω_0 one will get with a precision up to the first order members of $\delta K_0 / K_0$, $\delta\Omega / \Omega_0$

$$\int_t^{t+t_\varphi} K_0(t') \cdot dt' \cong \int_t^{t+t_\varphi^0} K_0(t') \cdot dt' - K_{00} \int_t^{t+t_\varphi^0} \frac{\delta\Omega(t')}{\Omega_0} dt',$$

where $t_\varphi^0 = \frac{\varphi_0}{\Omega_0}$. Finally we get for N_φ with a precision of the first order members

$$N_\varphi(t) = N_\varphi^0 + \frac{\Omega_0}{2\pi} \int_t^{t+t_\varphi} \delta K_1(t') \cdot dt' + \frac{1}{2\pi} \int_t^{t+t_\varphi^0} \delta K_0(t') \cdot dt' + \frac{1}{2\pi \cdot \Omega_0} \int_t^{t+t_\varphi^0} \delta K_{-1}(t') \cdot dt' -$$
$$\frac{1}{2\pi} \cdot (K_{10} \frac{\Omega_E}{\Omega_0} \sin\gamma + \frac{K_{00}}{\Omega_0} + 2\frac{K_{-10}}{\Omega_0^2}) \int_t^{t+t_\varphi^0} \delta\Omega(t') \cdot dt', \tag{12.32}$$

where $N_\varphi^0 = \frac{1}{2\pi} \cdot (K_{10}\varphi_0 + K_{10}\Omega_E \sin\gamma \cdot t_\varphi^0 + K_{00} t_\varphi^0 + \frac{K_{-10}}{\Omega_0} \cdot t_\varphi^0)$ - number of periods in the case of the absence of fluctuations.

The equation for $N_{2\pi}$ we get from (12..32) by chanting t_φ^0 to $T = \frac{2\pi}{\Omega_0}$. In equations, we've got, the real output response of the rotating RL is considered.

Let us discuss the DLG errors, caused by fluctuations of the RL output response coefficients. We will suppose fluctuations of coefficients $\delta K_i(t)$ to be stationary random processes with zero mean values. That means that we can use the correlation analysis results obtained above. To use equations (12..7), (12..8), (12.11) one should know correlation functions of processes $\delta K_i(t)$. Without considering pure technical effects such as voltage pulsation and others and taking into account the results of study of phase fluctuations of the rotating RL [12] it is possible to suppose that fluctuations of coefficients are wide band random processes characterised by correlation functions

$$R_{K_i} = D_{K_i} \cdot e^{-\alpha_i |\tau|}, \qquad i = 1, 0, -1. \tag{12.33}$$

In that case, using equations (12..20-12..24), we will get contributions of fluctuations of output response coefficients into the angle measurement errors as

$$D_{\varphi}^{K_0} = \frac{D_{K_0} \cdot \varphi_0}{(N_{2\pi}^0)^2 \alpha_0 \Omega_0} \cdot (1 - \frac{\varphi_0}{2\pi}) , \tag{12..34}$$

$$D_{\varphi}^{K_1} = \frac{D_{K_1} \cdot \Omega_0 \cdot \varphi_0}{(N_{2\pi}^0)^2 \alpha_1} \cdot (1 - \frac{\varphi_0}{2\pi}) , \tag{12..35}$$

$$D_{\varphi}^{K-1} = \frac{D_{K-1} \cdot \varphi_0}{(N_{2\pi}^0)^2 \alpha_{-1} \Omega_0^3} \cdot (1 - \frac{\varphi_0}{2\pi}) . \tag{12.36}$$

So the dependence of the error upon the value of the measured angle is the same for all three output characteristic coefficients. The same function we have obtained before and it is represented in the fig. 5. It's necessary to note that in these equations for the different coefficients of output characteristic we have different dependencies on the rotation rate. The dispersion of angle measurement errors caused by the RL's scale-factor fluctuations is proportional Ω, the dispersion caused by the bias instability varies in inverse proportion of the rotation rate, the dispersion caused by the nonlinearity of the output response reduces with Ω even faster. The analysis of the contribution of the RL's rotation rate instability into the angle measurement error is based on features of RL's output response. From (12.32) we can obtain the equation which will be the basis for future analysis

$$\delta N_{\varphi}(t) = \frac{1}{2\pi} \cdot (K_{10} \frac{\Omega_E}{\Omega_0} \sin\gamma + \frac{K_{00}}{\Omega_0} + 2\frac{K_{-10}}{\Omega_0^2}) \int_t^{t+t_{\varphi}^0} \delta\Omega(t') \cdot dt' . \tag{12.37}$$

As well as before the function we are integrating is a random process. In that case if the process is stationary we should know its correlation function. In spite that instability of Ω is determined by the type of bearing and drive, experiments show that the spectral density $S_{\Omega}(f)$ is concentrated mostly in the first or higher harmonics and has also slight wide band background. An example is shown on fig. 12..6.

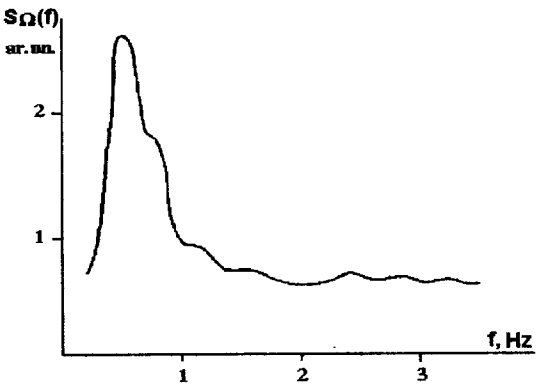


Fig.12. 6. The spectral density of the rotation rate instability.

That means that the rotation rate instability should be represented as a sum of wide band and narrow band random processes with correlation function of the following kind

$$R_{\Omega}(\tau) = D_{\Omega}' \cdot e^{-\beta|\tau|} \cdot \cos(\nu_0 \tau) + D_{\Omega}'' \cdot e^{-\alpha_0|\tau|} . \tag{12.38}$$

For the wide band part of the rotation rate fluctuation, using (12.20) -(12.24), we obtain

$$D_{\varphi}'' = \frac{D_{\Omega}'' \cdot \varphi_0}{(N_{2\pi}^0)^2 \cdot \alpha_{\Omega}} (\frac{K_{00}}{\Omega_0^2} + K_{10} \frac{\Omega_E}{\Omega_0^2} \sin\gamma + 4\frac{K_{-10}}{\Omega_0^3})(1 - \frac{\varphi_0}{2\pi}) . \tag{12.39}$$

The contribution of the narrow band part of the rotation rate fluctuations we can obtain by substituting the first part of the equation (12.38) into equations determining the dispersion of the periods number (12.13) and (12.17). To make the writing shorter let us take $K_{00}^{\sim} = K_{10} \cdot \Omega_E \sin\gamma + K_{00}$. Calculating an appropriate integral and substituting the results into equation (12.15)

and supposing for the narrow band random process (12..3 8) $\beta T >> 1, \beta << \nu_0$ we obtain dispersion of the measured angle, caused by narrow band fluctuations of Ω as

$$D_\varphi^{\Omega'} \cong \frac{2D'_\Omega}{(N_{2\pi}^0)^2}\left(\frac{K_{00}^{\sim}}{\Omega_0} + 2\frac{K_{-10}}{\Omega_0^2}\right)^2 \left\{\frac{1}{\nu_0^2}\left(1 - \cos\frac{\nu_0}{\Omega_0}\varphi_0\right) + \right.$$
$$+ \frac{\varphi_0^2}{(2\pi)^2}\cdot\frac{1}{\nu_0^2}\left(1 - \cos 2\pi\frac{\nu_0}{\Omega_0}\right) - \frac{\varphi_0^2}{2\pi\nu_0^2}\left[1 - \cos\frac{\nu_0}{\Omega_0}\varphi_0 - \right. \tag{12.40}$$
$$\left.\left. - \cos 2\pi\frac{\nu_0}{\Omega_0} + \cos\frac{\nu_0}{\Omega_0}(2\pi - \varphi_0)\right]\right\}.$$

The most interesting situation is when the rotation rate oscillation frequency coincides with one of the rotation rate harmonics. For example if oscillation frequency is equal to the harmonic n of Ω_0 , i.e. $\nu_0 = n\Omega_0$, then the equation (12.40) becomes

$$D_\varphi^{\Omega'} = \frac{2D'_\varphi}{(N_{2\pi}^0)^2}\left(\frac{K_{00}^{\sim}}{\Omega_0} + 2\frac{K_{-10}}{\Omega_0^2}\right)^2 \cdot\left[\frac{1}{n^2\Omega_0^2}(1 - \cos n\varphi_0)\right] . \tag{12.41}$$

The dependence $D_\varphi^{\Omega'}(\varphi_0)$ represents a harmonic process with a frequency determined by ratio of Ω_0 and ν_0. When n=2 this dependence has two periods in one revolution. In this case the angle measurement error due to the narrow band fluctuations has no maximum on $\varphi_0 = \pi$ in the contrast to the case of wide band fluctuations. .The maximum value of this error was calculated. It was considered in these calculations that the greatest contribution into the error due to the rotation rate instability is caused by the rotation of the Earth. The value of $K_1\Omega_3 \sin\gamma$ normally for traditionally used RL and for the latitude of St.Petersburg (γ=60°) is equal to $\approx$10 Hz. Values of K_0 and K_{-1}/Ω (with a rotation rate Ω =0.2 ...1 Hz) is not bigger then 1Hz. So in these calculations we assumed that $K_{00} = K_{-10} = 0$. The other values were taken as follows $N_{2\pi}^0 = 10^6$, $\sigma'_\Omega/\Omega_0 = 0{,}01$, i.e. arbitrary instability of the rotation rate is about 1 %. The result of the calculation is that the double amplitude of harmonic process equal to 0,08 arcs.

2.3. Consideration of the DLG base angle vibrations

We should especially stop on the influence of the DLG's base angle vibration on the measurement accuracy. The RL is sensitive to angle movements in relation to the inertial space, that is why it measures not only the table rotation rate in relation to the base, but also angle movements of the system base, caused for example by the vibration of the building. Considering this fact the equation for the RL output signal (12.28) should be supplemented by the component $K_1\xi'$, where $\xi' = \frac{d\xi}{dt}$ is the angular velocity of the DLG base. Then the equation for the number of periods should be also supplemented by the component

$$\delta N_\varphi = \frac{1}{2\pi}\int_t^{t+t_\varphi} K_1\xi'(t')\cdot dt'. \tag{12.42}$$

Using the approach, developed above, we will write down an equation for dispersion of period number of the RL output signal, caused by the base vibration as

$$D_{N_\varphi} = \frac{K_1^2}{(2\pi)^2}\int_t^{t+t_\varphi}\int_t^{t+t_\varphi} R_{\xi'}(u - v)\cdot du\cdot dv,$$

where $R_{\xi'}$- correlation function of angular velocity of the base. Using known relations we are coming to the correlation function of the base angle movement $R_\xi(\tau)$

$$D_{N_\varphi} = \frac{2K_1^2}{(2\pi)^2}\int_t^{t+t_\varphi}(t - \tau)R_{\xi'}(\tau)\cdot d\tau = \frac{2K_1^2}{(2\pi)^2}\int_t^{t+t_\varphi}(\tau - t)\frac{d^2R_\xi}{d\tau^2}\cdot d\tau .$$

After some simple transforms we obtain

$$D_{N_\varphi} = \frac{2K_1^2}{(2\pi)^2}\left[R_\xi(0) - R_\xi(t_\varphi) + t_\varphi\frac{dR_\xi}{d\tau}\Big|_{t=0}\right]. \tag{12.43}$$

The dispersion of the period numbers on the whole revolution one can get by means of changing t_φ by the revolution time T.

The cross-correlation moment $D_{N_\varphi N_{2\pi}}$ we can write in the same manner as in (12.17) and after some transformations get

$$D_{N_\varphi N_{2\pi}} = \frac{K_1^2}{(2\pi)^2}[\, R_\xi(0) - R_\xi(T) - R_\xi(t_\varphi) + R_\xi(T - t_\varphi) + 2t_\varphi \frac{dR_\xi}{d\tau}|_{t=0}\,]. \qquad (12.44)$$

We can obtain now the angle measurement dispersion by substituting (12.43) and (12.44) into (12.15).
Experiments [12] (see fig.12.7) show that the basis angle vibration is concentrated mostly on discrete frequencies, therefore we should take as an angle movement correlation function the sum of the correlation functions of the narrowband random processes.

$$R_\xi(\tau) = \sum_i d_i e^{-\alpha_i|\tau|} \cos \nu_i \tau .$$

Substituting that form of the correlation function in the equation of the dispersion of angle measurements we obtain

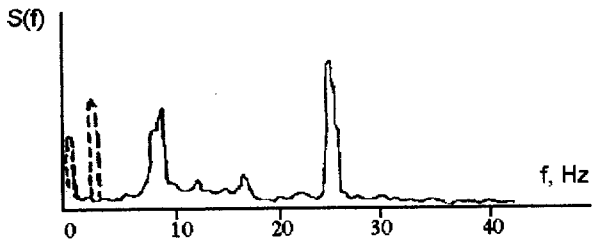
$$D_\varphi = \frac{2K_1^2}{(2\pi N_{2\pi}^0)^2} \sum_i d_i \{(2\pi)^2 (1 - \alpha_i t_\varphi)(1 - \cos \nu_i t_\varphi) +$$
$$+ \varphi_0 (1 - \alpha_i T)(1 - \cos \nu_i T) - \varphi_0 2\pi [(1 - \alpha_i t_\varphi)(1 - \cos \nu_i t_\varphi) +$$
$$+ (1 - \alpha_i T)(1 - \cos \nu_i T) - (1 + \alpha_i t_\varphi - \alpha_i T)(1 - \cos \nu_i (T - t_\varphi))]\}.$$

In the case when the vibration is caused by the rotation of the DLG table and concentrated on frequencies multiple of the rotation frequency $\nu_i = n^{(i)}\Omega$, where $n^{(i)}$ - integer, expression for D_φ becomes simpler

$$D_\varphi = \frac{2K_i^2}{(N_{2\pi}^0)^2} \sum_i d_i [(1 - \alpha_i t_\varphi) - \frac{\varphi_0}{2\pi}(\alpha_i T - 2\alpha_i t_\varphi)](1 - \cos n^{(i)} \varphi_0). \qquad (12.45)$$

One can see here that the error caused by vibrations depends on the value of the measured angle and the frequency of the main vibration.

The research of angle vibrations of the DLG base let us to get its spectral density with unmovable and rotating platform. That is represented in fig.12.7 by solid and dotted lines. When the table is unmovable there are two main lines in the spectrum : 8.3 Hz and 25 Hz. When the DLG table is rotating with the frequency of 0.5 Hz, components 0.5 Hz and 3.5 Hz are appearing (the first and seventh harmonics)



Fig,12.7. The spectrum of the angle vibration of the DLG base

The determination of the vibration spectrum allows using the analysis results to estimate the vibration contribution into the DLG error. Numeric values of the vibration spectrum were substituted into (12.45), which allowed to get the dependence of the measurement error on the measured angle value.

Fig.12. 8 represents the dependence, obtained with calculations. It shows a rapid growth of the measurement error when the measured angle value is in order of several degrees, i.e. when the value of φ/Ω approximately equals to the minimum vibration period. For these angles the error grows up to 0.1arcs which corresponds to the vibration amplitude. With the further growth of the measured angle influence of vibrations decreases and the error value stops growing and remains on the approximately constant level because of the random nature of the vibrations and the averaging of their influence.

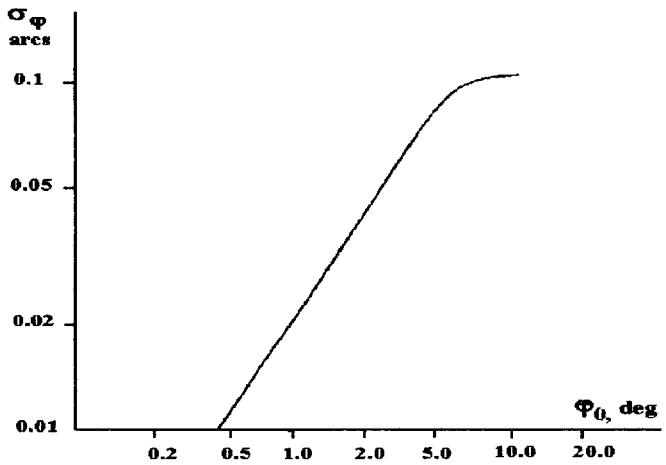


Fig.12. 8. The DLG error caused by angle vibration of the base

2.4. Quantization error

One of the most important components of the random angle measurement error in DLG is the quatization error. To be more correct it is an error caused by digital form of the RL output signal. Counters, summating the RL signal periods during intervals, determined by the angle converter, can count only integer number of periods. That causes a certain random error. Let us discuss that error. Fig.12.9 shows the summation interval (a) and RL output signal pulses (b).
In case shown in the figure the counter would count 4 periods of the RL output signal inside the summating interval. One can see in the figure that as a matter of fact only about 3.5 periods are in the interval.

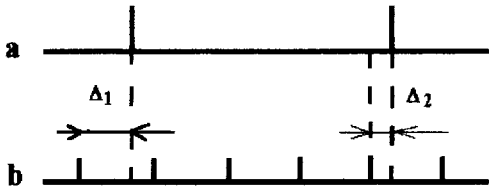


Fig.12. 9. The pulse diagram to the quatization error estimation.

The exact number of periods in that case is equal to

$$N = 4 + \frac{\Delta_2}{\tau} - \frac{\Delta_1}{\tau},$$

where τ - the RL output signal time period (the RL rotation rate supposed to be constant). The error of the number of periods counting in that case is equal to

$$\Delta_d = \frac{1}{\tau}(\Delta_1 + \Delta_2).$$

The time interval between the pulse starting the periods counting and the first counted pulse of the RL signal is normally considered to be random and even distributed over the RL signal period. The same is supposed for the end of the summating interval. Such assumption should be considered right because of the independence of random factors defining the angle position of output pulses of the angle converter and RL. In that case the digital count error has a triangle distribution with dispersion

$$D_d = \frac{\Delta\varphi_{RL}}{6},$$

where $\Delta\varphi_{RL}$ - angle per RL signal period, defined by (7). With a typical dimensions of the RL (square with side length 11cm) $\Delta\varphi_{RL}$ is equal 1.26 arcs and the RMS of the digital count error is σ_d = 0.51arcs. This error may be reduced by using the multiplication of the RL output frequency or another interpolation technics.

2.5. Potential limit of the DLG accuracy

The potential accuracy of the DLG is defined by the quantum noise of the RL. Owing to the spontaneous emission of radiation the phase of the RL output signal has a random component. This random component is accumulated during the measurement time resulting in the random drift (random walk) of the number of the pulses.

Using results of [1,13] it is possible to write the dispersion of the RL output frequency fluctuations provided by quantum noise in the form

$$D_\omega = \frac{\delta\omega_n}{t_\varphi},$$

where $\delta\omega_n$ - "natural" linewidth of the frequency difference of counter-propagating waves. Passing to the number of the RL output pulses and taking into account calibration of the RL one can get, with the use of the obtained above results, the error of the DLG

$$D_\varphi = \frac{\delta\omega_n \varphi_0}{(N^0_{2\pi})^2 \Omega}\left(1 - \frac{\varphi_0}{2\pi}\right).$$

The maximum of the angle measurement error corresponds in this case to $\varphi_0 = \pi$ and is equal to

$$\left[D_\varphi\right]_{max} = \frac{\delta\omega_n \pi}{2(N^0_{2\pi})^2 \Omega}.$$

To make estimation of this error it is necessary to know the value of the natural linewidth $\delta\omega_n$. This value as a function of the RL parameters can be written as

$$\delta\omega_n = \frac{8\pi h\nu}{P}(\Delta\nu_c)^2,$$

where P - the power of the output beam, ν - the light frequency and $\Delta\nu_c$ - the cavity linewidth. Let us consider the RL of the square cavity with the side of about 11cm, the losses in the cavity about 10^{-4} and the output power of 50 μW. In this case for the value of the rotation rate Ω of about 0,7rev/s the standard deviation of angle measurements can be estimated on the level of 0,004arcs. Certainly we can call this value "potential limit of the DLG accuracy" only arbitrarily, because it depends on the values of the RL parameters and therefore can be improved by changing the RL quality.

3. MAIN MODES OF DLG OPERATION AND KEY ELEMENTS

3.1. Optical polygon calibration

Calibration of discrete angle converters (such as magnet converters, inductosines, optical encoders etc.) and angle gauges (optical polygons, limbs etc.) was the main aim of development of the first laser goniometers [5-9]. We will analyse the DLG on that purpose consideringfirst the laser goniometer for optical polygon calibration.

The DLG's schematic for optical polygon calibration is presented in fig. 12.10. Mostly this scheme is equal to the general one presented in the fig.12.1, but this scheme has its own features determined by this certain purpose. That DLG consists of the RL 1; rotary table 2; drive 3; optical polygon (OP) 4 mounted on the rotary table together with the RL; optical null-indicator (NI) 5 mounted on the system base; interpolator (for instance, frequency multiplier) of the RL output signal 6; counters 7; computer 8. One of the key element of that scheme is the optical null-indicator, which is connected optically with the OP faces. While rotation of the OP in the moments of coincidence of the normals to the OP faces and the NI axis the NI generates short pulses, forming

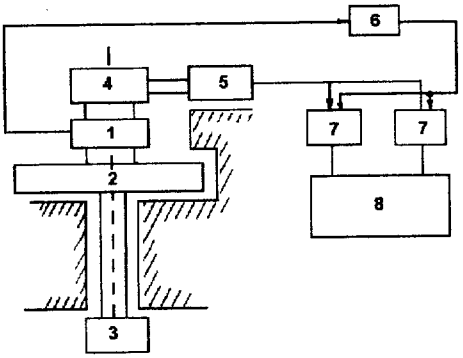


Fig.12. 10. The DLG's schematic for optical polygon calibration

intervals of summating the RL output signal periods. The DLG operation diagram corresponds to the one shown in fig.12.4. The only difference is that the NI signal is used instead of the angle converter signal. Besides, the RL output signal passes the interpolator.

The operation of the DLG accounting the peculiarities of the RL output response and instability of the rotation rate was analysed above in detail that is why the most attention here will be attracted to the NI and some peculiarities of the laser goniometer as a system for the OP calibration.

The NI in that system is used to form intervals of summating the RL output signal periods. Errors of the summating interval formation, caused by the NI, contribute directly to the measuring system error. We should analyse in detail the process of forming of summating intervals to determine main requests to NI performance.

3.1.1. Null-indication of angle directions

As it was said above the optical NI generates a pulse in the moment of coincidence of its axis with the normal to the reflecting surface (the OP face). Summating interval bounds are formed at the moments, when the pulse value reaches a certain level fixed by the trigger. The trigger is switched by the pulse edge or by the edge of the signal formed after differentiation of the pulse. The second routine is better because it allows to work near the pulse peak where the derivative has normally the biggest slope. Any way in both cases error appears due to the fact that the real NI output signal contains of the determined pulse signal and some additive noise. The task of determining the error of the null-indication of the angle direction may be solved by using results of pulse coming time estimation in the radiolocation systems [14]. Following the mentioned work, one can determine the dispersion of the pulse coming time as

$$D_\tau = \frac{1}{\beta_u q}, \qquad (12.46)$$

where q - the signal-to-noise ratio; b_u - the pulse form parameter

$$\beta_u = \frac{\int_{-\infty}^{\infty} \left(\frac{du(t)}{dt}\right)^2 \cdot dt}{\int_{-\infty}^{\infty} u^2(t) \cdot dt}. \qquad (12.47)$$

We should operate here with the angle variables, so let us pass from time variables to angle variables, supposing the rotation rate quasi-constant and changing in (12.46), (12.47) t to φ (angle changes in time linearly). Suppose the NI output pulse has an bell-like shape

$$u(\varphi) = \exp[-\pi(\frac{\varphi}{\alpha_0})^2], \qquad (12.48)$$

where α_0 - angle "width" of the pulse on the level $e^{-\frac{\pi}{4}}$. Then the equation (12.46) will transform to

$$D_\varphi = \frac{\alpha_0^2}{\pi q}, \qquad (12.49)$$

i.e. the dispersion grows with the square of the pulse angle width and reduces with the signal-to-noise ratio.
According to obtained result we will discuss some possibilities of NI's design. A traditional way of NI's design is using an autocallimation system. Fig. 12.11 represents a simplified scheme of the autocallimation NI. The light source 1 illuminates the regulating slot diaphragm 2, located into the focal plane of the lens 3, and as a result forms a parallel light beam which impinges to the rotating OP face. After reflection the light beam is directed by the beam-splitter cube 4 to the slot diaphragm 5 also located in the lens focal plane. Behind the diaphragm the photo-detector (PD) is situated. It detects the intensity of the light passed the diaphragm. While the OP is rotating the PD forms output pulses corresponding to the coincidence of the OP face normals and the NI axis. The linear shift Δx of the image of slot 2 upon the analysing slot 5 is defined by the OP turn φ as follows

$$\Delta x = F \cdot tg\varphi,$$

where F- focus of the lens. If both slots have the same width h the NI output pulse would be of the width α_0 equal (assumingthe angle φ small)

$$\alpha_0 = \frac{h}{2F}.$$

Let h=20 μm, F = 0.5 m then the pulse angle width would be 4 arcs. Substituting that value into (12.49) and considering the signal-to-noise ratio equal 100 we calculate the null-indication error σ_φ = 0.23 arcs.

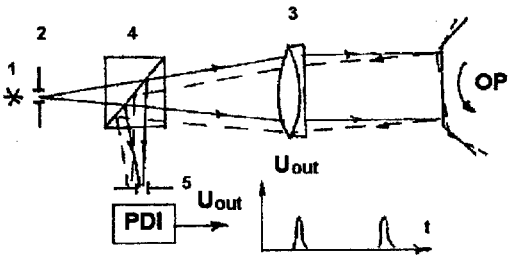


Fig. 12.11. The scheme of the null-indicator using the autocollimation system.

That accuracy normally is not sufficient for the purposes of precision OP calibration. It seems that it can be improved by using narrower slot, but it causes lowering the signal-to-noise ratio due to diffraction.

Some improvement may be achieved by using slot masks with pseudo-random distribution of transparent arias instead of slot diaphragms [15]. This method allows to rise the signal-to-noise ratio with the same angle width of the pulse and so to improve the null-indication accuracy.

Another possibility of developing the precise NI is to use interference schemes. Let us analyse the interference NI operation. The analysis can be carried out, for example, for a null indicator with the beam-splitter as a Koesters prism. The schematic of the null indicator is shown in fig.12.12, where 1 is the light source, 2 is a Koesters prism, 3 is a reflecting mirror and 4 is the photo-receiver aperture. The electric fields of the beams interfering at the output of the Koesters prism can be represented in the form:

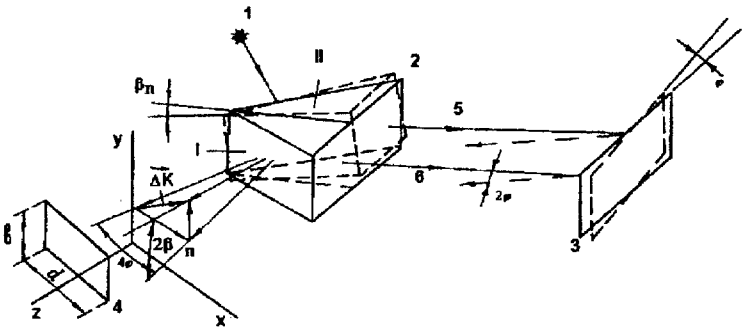


Fig. 12.12. The scheme of the interference NI with the beam-splitter as the Koesters prism.

Here it is assumed, for simplicity, that the light source is monochromatic, coherent (for example, a He-Ne laser in single-mode operating regime) and producing a beam having a plane front. We will take into account difference in the optical length for the two beams by means of the phase terms $\delta_1(\bar{r})$ and $\delta_2(\bar{r})$.

Let us consider the mutual displacement of two halves of the Koesters prism and mutual tilt of both faces of the Koesters prism. The light intensity at the output of the Koesters prism can be written as

$$I = \frac{1}{2} <(\bar{E}_1 + \bar{E}_2)(\bar{E}_1 + \bar{E}_2)^* > = \frac{1}{2}(E_{10}^2 + E_{20}^2) + E_{10}E_{20}\cos[\bar{r}\Delta\bar{k} + \Delta\delta],$$

where $\Delta\bar{k} = \bar{k}_1 - \bar{k}_2$, $\Delta\delta = \delta_1 - \delta_2$.

In the coordinate system of the photo-receiver (see fig.12.12) the difference components of the wave vectors have the form

$$\Delta k_x = 2|k|\sin 2\varphi, \qquad \Delta k_y = 2|k|\sin\beta,$$

where φ is the angle of turn of the reflecting mirror 3 from the normal position, and β is the mutual displacement angle of the two halves of the Koesters prism in the plane of the beam-splitter surface. We will assume now for simplicity that the fronts of the interfering waves are plan. In this case the phase difference $\Delta\delta$ includes a phase change caused by the turn of the reflecting mirror as well as a phase change due to the mutual linear displacement s of the two halves of the Koesters prism in the plane of the beam-splitter surface in the horizontal direction

$$\Delta\delta = \frac{4\pi}{\lambda}[\varphi \cdot h + s(\frac{n}{2} - 1)],$$

where h - is the distance between the centres of the beam incidence on the reflecting mirror and n is the index of refraction of the prism. Assuming the angles φ and β small, we can write

$$I = \frac{1}{2}(E_{10}^2 + E_{20}^2) + E_{10}E_{20}\cos[\frac{2\pi}{\lambda}(4\varphi x + 2\beta y + 2\varphi h + 2s(\frac{n}{2} - 1)].$$

By integrating the intensity over the rectangular aperture of the photo-receiver and assuming that $E_{10} = E_{20} = \sqrt{I_0}$, we obtain the output signal in the form

$$U(\varphi) = \gamma_P \int_{-\frac{b}{2}}^{\frac{b}{2}} \int_{-\frac{d}{2}}^{\frac{d}{2}} I(x,y,\varphi)dxdy =$$

$$= \gamma_P I_0 bd[1 + \frac{\sin(\frac{2\pi}{\lambda}\beta b)}{\frac{2\pi}{\lambda}\beta b} \cdot \frac{\sin(\frac{4\pi}{\lambda}\varphi d)}{\frac{4\pi}{\lambda}\varphi d} \cos[\frac{4\pi}{\lambda}(\varphi h + s(\frac{n}{2} - 1)]]. \quad (12.51)$$

where γ_P is the photo-receiver sensitivity, b and d are, respectively, the height and width of the aperture. The difference between this form of the output signal and the ideal case (β =0, s=0) lies in the decrease of amplitude caused by the presence of the factor $\sin(\frac{2\pi}{\lambda}\beta b) \Big/ \frac{2\pi}{\lambda}\beta b$ and in the distortion of the shape due to the phase shift $s(\frac{n}{2} - 1)$.

Fig. 12.13a shows the interference NI real output signal. One can see its wide base and a short interference pulse formed at the moment when a pair of light beams incident normally on the reflecting surface. The form of the interference pulse itself is shown on fig. 12.13b in a bigger scale. One can see the asymmetry of the side wings which is caused by the Koesters prism halves shift. The interference pulse width can be estimated in accordance to the equation (12.51). For d = l = 10mm it is about 0.6 arcs. Substituting this value into the equation for the null-indication error (12.4 9) with q = 100 we get σ_φ = 0.03 arcs.

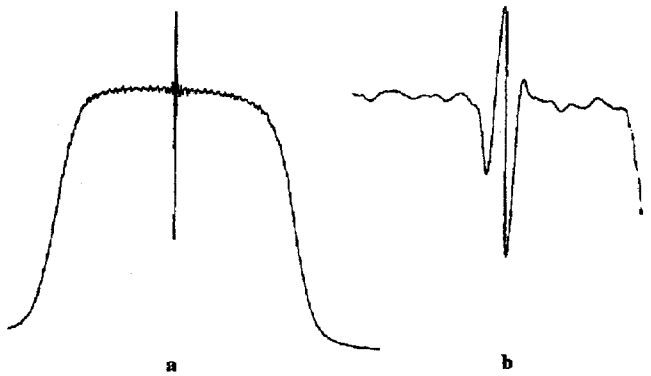


Fig.12.13. The output pulse of the interference NI (a - the whole output pulse, b - the central part of the pulse in a bigger scale).

The NI with the Koesters prism and Dove biprism as the beam-splitter have practically no differences, therefore the equation for the output signal (12.51) is the same for both variants.
The shortcoming of described NI types is a very high request to parameters β and b in the technological process of beam splitter prism production. Such prisms are usually combining in a special technology process with checking up the interference pattern.

3.1.2. The influence of the system elements adjustment on the calibration error

Let us analyse the peculiarities of the DLG operation, connected with errors of the adjustment of goniometer elements. We will start our analysis from the influence of the inclination of the OP axis in relation the rotation axis.

Two coordinate systems should be considered. Axes xyz are connected with the rotary table so that the axis z is the rotation axis (fig. 12.14). Axes $\xi\zeta\eta$ are connected with the OP. Suppose that the prism is inclined around x axis on angle β. Let one of normals to OP faces be the x axis. Angle between that normal and the normal to the next OP face (line OB in the figure) is equal φ_0 The connection between angle φ_0 and its component on the measuring plane (xy plane) φ_0' can be found from the triangle AOB. It follows from this triangle that OA=Rcos φ_0, AB=Rsin φ_0. The triangle AOC shows that

$$\mathrm{tg}\varphi_0' = \frac{AB}{AO} = \mathrm{tg}\varphi_0 \cos\beta.$$

If the angle between the OP faces is equal φ_0 the turn angle of the platform between two NI pulses is equal φ_0' . Then the measurement error is

$$\Delta\varphi = \varphi_0 - \varphi_0' \quad ,$$

$$\Delta\varphi = \varphi_0 - \mathrm{arctg}(\mathrm{tg}\varphi_0 \cdot \cos\beta) \quad . \qquad (12.52)$$

Fig. 12.15 represents the dependence of the error of angle measurements $\Delta\varphi$ on the value of the measured angle φ_0. This is the systematical harmonic error with a period π. It corresponds to the second harmonic of the rotation frequency .

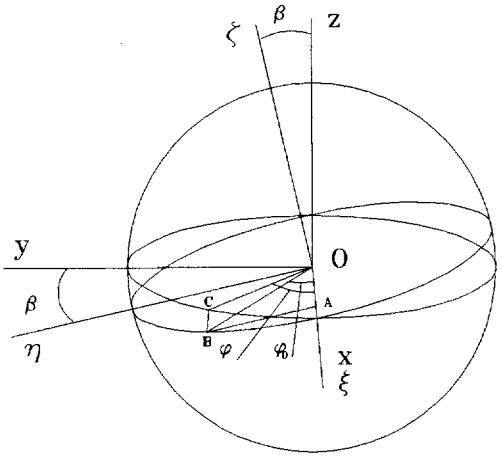


Fig. 14. The scheme to estimation of the measurement error due to the system elements disadjustment.

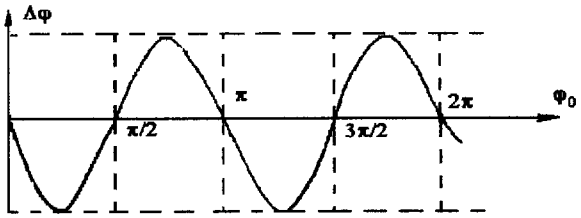


Fig.12. 15. The dependence of angle measurements error $\Delta\varphi$ on the value of the measured angle φ_0.

Let us analyse the dependence of $\Delta\varphi$ on the prism inclination angle β. Angle β normally is not bigger than the several arcminutes, therefore we can write (12..52) as

$$\mathrm{tg}(\varphi_0 - \Delta\varphi) = \mathrm{tg}\varphi_0 \cdot (1 - \frac{1}{2}\beta^2) .$$

After some transformations we obtain

$$\Delta\varphi = \frac{\mathrm{tg}\varphi_0}{2(1 + \mathrm{tg}^2\varphi_0)} \cdot \beta^2 \quad . \qquad (12.53)$$

Estimation of the error according to equation (12.53) shows that if the OP inclination angle β is about 2 arcmin then the maximum value of Δφ (for $\varphi_0 = 45^\circ$) is about 0.01 arcs, i.e. it is considerable only if the OP inclination angle is much more than one arcminute.

Another source of the systematic error is an inclination of the NI around its axis. Null-indicators used in the DLG should be sensitive only for angle movements of the reflecting surface around one axis. We will use a term NI sensitive plane. The inclination of the NI sensitive plane in relation to xy plane is equivalent to sensitivity of the NI for angle movements in the plane orthogonal to the rotation plane. Suppose the NI sensitive plane is inclined on a certain angle to the rotation plane. If at the same time there is no inclination angle of the OP, there wouldn't be any measurement error, because all OP angles have the same components on the NI sensitive plane. If there is the OP inclination, then components of angles between normals to

different OP faces are different and therefore a measurement error appears. This error is changing harmonically with the DLG platform rotation with the same period, i.e. it is concentrated on the first harmonic of the rotation frequency. Maximum of that harmonic dependence is determined by

$$\Delta\varphi_{max} = \beta \sin\gamma, \qquad (12.54)$$

where γ - angle of NI tilt. The same effect is known for autocollimators [18]. This source of error strongly influences the measurement accuracy. Let β = 1 arcmin, γ = 10 arcmin (adjustment of the NI sensitive plane is rather difficult operation), then $\Delta\varphi_{max}$ = 0.18 arcs, which is more than one order bigger then the previous estimation.

A considerable error may be caused by the radial wobbling of the bearings of the rotation axis. In that case the rotation axis makes conical movements of the angle α

$$\alpha \cong \frac{r}{l},$$

where r - the value of the radial wobbling , l - distance between bearings holding the axis. This error is of the same type as one caused by the OP inclination and may be estimated with the use of (12.53) and (12.54) if we assume $\beta = \alpha$. Let r = 10 μm and l =200 mm, then we have the inclination of OP on the angle β = 10 arcs. In this case for γ = 10 arcmin $\Delta\varphi$ = 0.03 arcs.

3.1.3. Results of DLG use for optical polygon calibration

The results of the research work which are considered in this chapter have been obtained using the DLG system EUP-1L developed at the St.Petersburg Electrotechnical University.

The EUP-1L allows angle calibrations to be carried out on optical polygons in accordance with the above description. The EUP-1L system uses an RL with He-Ne active medium, λ=633 nm, monolithic square cavity structure with four total reflecting prisms ([16], also chapter 5 of this monograph), 44 cm perimeter and an angle interval $\Delta\varphi_{RL}$ (discretness of the RL) of about 1,3 arcs. To increase the resolution, the RL output signal frequency is multiplied with a PLL circuit by a factor of 10. The constant rotation rate of the RL goniometer spindle is approximately 0,7 rev/s. The spindle is carried in ball bearings and driven by a DC motor via a belt drive. Slipring contacts are used to transmit the RL signal and the supply voltage. The size of the system is about 300x300x300 mm. Fig.16 shows the design of the EUP-1L. At the top of the instrument, an optical encoder is installed for the purposes described in the following. The EUP-1L goniometer is connected to a personal computer which ensures data processing using specific programs. For this investigation, the "polygon" program was used which allows angles between OP faces to be measured.

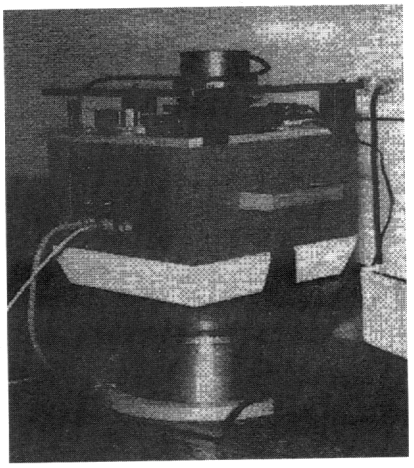


Fig. 12.16. The design of the laser goniometer EUP-1L.

For the investigation, a 12-faced precision polygon of the PTB (Physikalish-Technische Bundesanstalt, Braunschweig, Germany) was used which had been calibrated against an angle standard turntable together with an autocollimator and a phase-shifting interferometer. It was shown in [17] that these two methods can lead to slight differences between the results, which are due to unflatness of the OP faces .

The measurements with the EUP-1L were carried out using an interference null indicator with Dove biprism described above. The NI contained an incandescent lamp as light source and a photodiode as receiver. The cross-section of the light beam consisting of both beams of the NI interferometer was a rectangle 20 mm in height and 16 mm in width. It should be noted that for the OP calibration at the PTB a light beam with a circular cross-section 25 mm in diameter was used.

The measurements with the EUP-1L were carried out by repeated RL counts over 50 to 100 revolutions of the goniometer rotary table. The typical standard deviation of the mean values from these measurements was about 0,02 to 0,03 arcs. The main source of a systematic error was found to be the inclination of the OP in relation to the rotation axis of the goniometer and the tilt of the NI with respect to its optical axis. We largely eliminated this influence by mounting the OP on a precisely

adjustable device and obtained a coincidence with the rotation axis of the order of 5 arcs. The cross calibration method [18] was used, so the whole measurement procedure consisted of 12 steps: In each step the OP was turned by 30 deg in relation to the rotary table and measured again. The final results for the 12 polygon angles were then obtained from the well-known evaluation of 12x12 measurements.

Fig.12.17 shows the angular differences in arcs obtained from a comparison between the EUP-1L and the PTB methods with autocollimator (line EUP.AC) as well as with phase-shifting interferometer (line EUP.PI). It can be seen that the differences from the interferometer data are smaller and do not exceed ±0,15 arcs. It is assumed that the differences are due to the use of the interference null indicator and also to the above-mentioned differences between the cross-sections of the light beams.

The EUP-1L measurements on optical polygons were also compared with measurements of the VNIIM (Institute of Metrology, St.Petersburg, Russia). In this case, an 8-faced OP with faces size 24mm x 56 mm had been used. It should be noted that the flatness error of the faces of this OP was better than that of the 12-faced OP of the PTB. This had an average flatness error of about 11 nm (RMS) on the square of the NI light beam cross section, whereas the 8-faced OP had about 6 nm (RMS) on the same square. Fig.18 shows the angular differences between the results obtained by the VNIIM and those obtained by use of the EUP-1L. CW and CCW lines in the figure correspond to results at different directions of rotation of the EUP-1L's rotary table. As can be seen, the angle difference does not exceed ±0,035 arcs on an average. It is assumed that these differences which are much smaller than the data in Fig.17 are mainly due to the better flatness quality of the 8-faced OP.

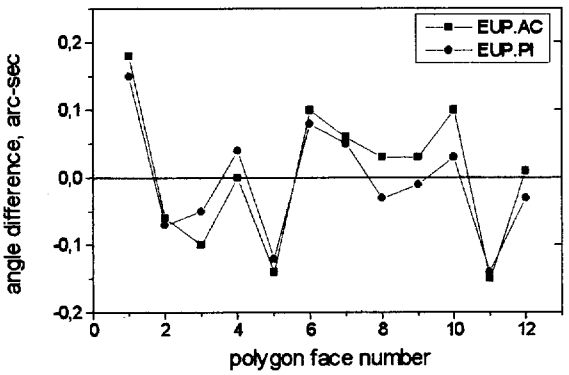


Fig. 12.17. Angular differences between angle measurements on a twelve-faced OP with the EUP-1L and the PTB methods with autocollimator (line EUP.AC) as well as with phase-shifting interferometer (line EUP.PI).

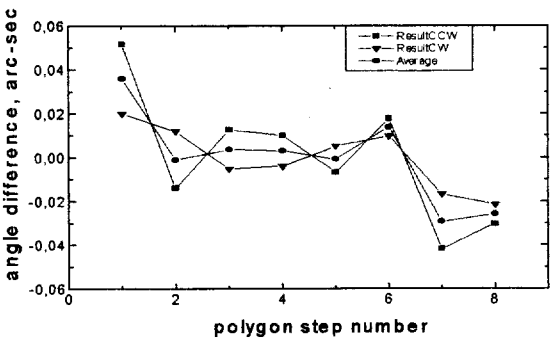


Fig.12. 18. The angular differences between angle measurements on an eight-faced OP with the EUP-1L and the VNIIM autocollimator method.

As a result we can conclude, that the error of optical polygon calibration by means of the EUP-1L is mainly dependent on the optical quality of the polygon faces, whereas the uncertainty of the calibration is equal 0,02 to 0,03 arcs.

3.2. Optical encoders calibration

Angle gauges (OP, limbs) calibration is a rather complicated procedure, requiring special NI and very accurate adjustment of the system elements. In case of limb calibration instead of NI we should have a microscope, forming a pulse signal when the limb edge is passing its view.

Calibration of optical encoders (incremental or absolute type) is easier in realisation because they usually consist of stator and rotor parts connected by its own bearing. In that case it is necessary to connect the moving part of the encoder with the DLG shaft by a solid coupling and fix the stator part of the encoder on the DLG base with middle accurate adjustment.

As an example of the DLG implementation for optical encoder calibration we can use the results obtained during investigation of the EUP-1L system. These measurements were aimed at obtaining an estimate for the repeatability and accuracy of the EUP-1L in this mode of operation. The commercial device Heidenhain RON255 was examined for this purpose. This device is an incremental angle encoder of medium accuracy with 18000 signal periods in 360 deg. The accumulated error for this encoder is specified within 5 arcs. The RON255 was mounted on the EUP-1L (it can be seen in fig.12.18) in coaxial direct coupling with the shaft of the goniometer spindle. The integral coupling device of the RON255 allowed the influence of shaft wobbling and the non-coincidence of EUP- and RON-axes to be almost completely eliminated. Our electronic measuring system included a frequency divider for the encoder output signal, intended to reduce the number of angular intervals to be calibrated. The number of intervals defined by the division factor varied in the range from 30 to 360 (division factor 600 to 50).

As an example, some measurement results are presented with 30 encoder intervals. 25 repeated measurements were performed, each covering 25 revolutions of the goniometer table. Fig.12.19 shows the results of four measurements of this series. It can be seen that the 30 encoder intervals deviate from their nominal angles of 12 deg by values between -0,8 arcs and +0,6 arcs, showing a typical first-harmonic distribution due to the own encoder disc eccentricity. The measurement results were also processed with a view to obtaining the accumulated error of the encoder, i.e. the difference from the nominal value of an angle referred to an initial angle. Fig.12.20 shows the differences between the accumulated angles of the encoder and the mean values over all the measurements of the series, for the same four measurements as in Fig.12.19.

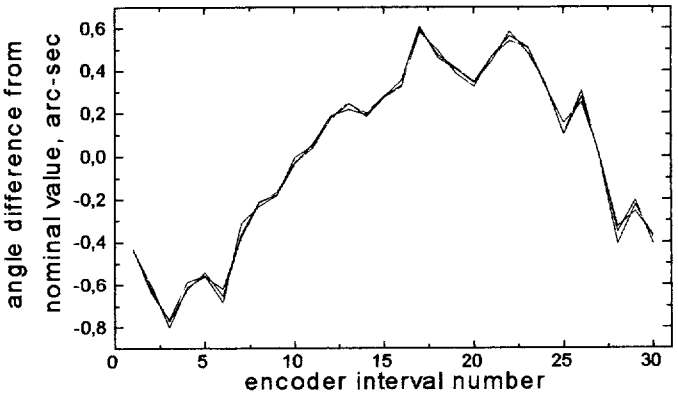


Fig.12.19. The results of the optical encoder RON255 calibration by means of the laser goniometer EUP-1L.

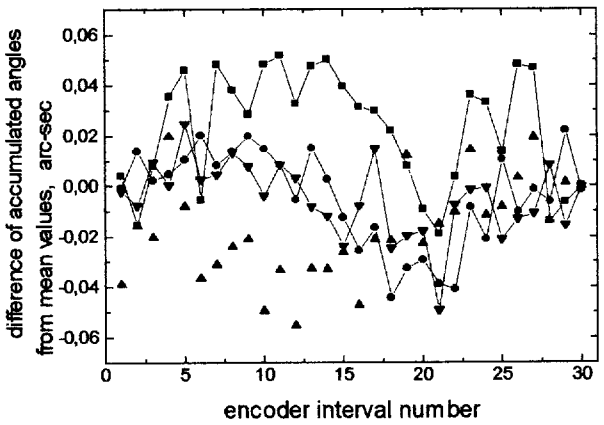


Fig.12.20. The differences between the accumulated angles of the encoder and the mean values over all the measurements of the series, for the same four measurements as in Fig.12.19.

The characteristics of these differences which are within ± 0,06 arcs are partly of a random nature but partly also of a systematic first-harmonic nature. The latter may be caused by the axial instabilities of the EUP-1L spindle or of the encoder bearing. The essential results of the optical encoder calibration series are summarized in the following table. From these data it can be concluded that the uncertainty of the RL goniometer in this mode of calibration is of the order of 0,02 to 0,03 arcs.

Table 12.1

Maximum deviation of 30 angles intervals from the nominal value	Error of angle interval measurements (standard deviation)	Maximum deviation of 30 accumulated angles from the nominal value	Error of accumulated angle measurements (standard deviation)
0,78 arcs	0,019 arcs	2,68 arcs	0,025 arcs

3.3. External noncontact angle measurements

One of the most important tasks for angle measurements is development of noncontact systems for measurements of angle movements of various type objects. Among aims of such systems one should note determination of angle position of inertial stabilised objects (mostly gyrosystems) and measurement of angle movement parameters of the test tables. Both tasks suppose a very complicated movements of controlled objects and require measuring systems with very high precision, data reading frequency and measuring range.

Traditional means (autocollimators, interferometers) normally do not satisfy those requests. The development of the DLG allowed to solve some of these tasks.

3.3.1. Principle of the DLG for external noncontact measurements

Let us consider the DLG's operation in the mode of external angle measurements. It allows to estimate the angle position of an object in respect to a certain basic direction. For this mode of operation (fig.12.21), the DLG depends on the rotation of the RL together with an optical polygon (OP). With this rotation, the light beam of a null indicator (NI), mounted on the base of the system, is scanned in a horizontal plane by its reflection from the faces of the rotating OP. The reference direction of the system is formed by the normal to the reference mirror (RM) which is fixed to the base. The NI generates pulses whenever the beam impinges normal to RM and to the control mirror (CM) which is mounted on the movable object to be measured.

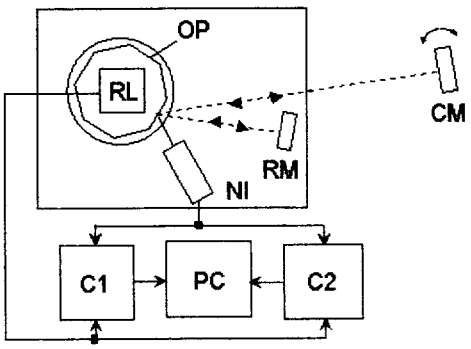


Fig.12.21. The scheme of the DLG in the mode of operation of the external angle measurement

Consequently the light beam reflection from each face of the OP leads the NI to give two output pulses (from RM and CM), and the interval between these pulses defines the angle formed by RM and CM. To measure this angle, the number N_φ of the RL output pulses given in the interval are to be counted. To obtain these numbers two counters (C1, C2) alternately totalize the RL output pulses within the intervals, similar to the scheme in Fig.4. For the scale factor, it is also necessary to count the total number of pulses during one revolution $N_{2\pi}$. When these numbers are known, the CM angular position in relation to the RM position can be calculated using the formula

$$\varphi = 4\pi \left[N_\varphi / N_{2\pi} \right]$$

which is equivalent to (12.6) except the factor 2, taking into account the doubling of the reflection angle at the OP.

As this measurement of the angle φ is carried out on each OP face, the measurement is repeated with the frequency f=n/T, where n is the number of the OP faces and T the time of one RL revolution.

An important feature of this mode is an extremely wide angle measuring range with a high accuracy. If the faces of the used OP are rather wide (about 50 mm), this range $\Delta\varphi$ can be approximately defined by the following formula

$$\Delta\varphi \cong 2\tan^{-1}(D/L)$$

where D is the diameter of the control mirror (CM) whose angular position is measured and L the distance between OP and CM. If the distance is rather small (about 0,2 to 0,5 m), the measuring range $\Delta\varphi$ can extend up to a value of 30 deg. In this mode of operation, with the usual for DLG accuracy of about 0,1 arcs, the DLG has an important advantage over photoelectric autocollimators with their very small measuring range.

3.3.2. Investigation of DLG in external measurement mode

The investigation of the DLG's accuracy into this mode of operation was carried out using the precision standard turntable of the PTB. This turntable which was used in the measurements together with the EUP-IL, was mounted on the granite table in a laboratory room at a controlled ambient temperature of $(20 \pm 0{,}1)^{\circ}C$. The turntable contains a high precision air-bearing spindle

with cylindrical and axial support. The spindle is rotated by a precise drive system consisting of a coarse friction wheel drive with a controlled DC motor and a piezoelectric fine-positioning drive. For angle measurement, the table is equipped with a Heidenhain RON 905 incremental rotary encoder system coupled inside the housing to the bottom of the spindle. The angular resolution of the encoder system, together with electronic interpolation, is about 0,035 arc-sec. The overall uncertainty of the system amounts to 0,1 arcs.

To avoid errors caused by the unflatness of the CM due to the displacements of the scanning beam, the EUP-1L was mounted on the turntable and the control mirror was fixed to the granite plate at a distance of about 20 cm from the turntable. The EUP-1L was mounted such that the point of incidence of the NI beam on the OP face lay in the rotation axis of the turntable. In this case, the point of incidence of the beam on the CM surface is not displaced when the EUP-1L is rotated on the table.

Several series of measurements were carried out rotating the table with the EUP-1L by various angles. For each angle of rotation, a measurement was made over 25 revolutions of the goniometer table. The uncertainty of each measurement, evaluated as SD of the mean over 25 revolutions of the table, was of the order of 0,05 to 0,08 arcs. Fig.12.22 shows a typical measurement result obtained when the turntable changed its position in an automatic run in the range of 25 deg in steps of 1 deg. In the figure, the step error (line *diff.*) and the accumulated error (line *accum*) of the measurements are given as deviations from the values indicated by the turntable. The data processing program allowed an angular resolution of 0,01 arcs to be obtained only for angles within 10 deg, whereas for angles exceeding this limit, a resolution of only 0,1 arcs was obtained (cf. figure). As it follows from the figure, the step measurement error is not greater than 0,2 arcs in the whole range of 25 deg and the accumulated error is within of 0,35 arcs in the same range of angles.

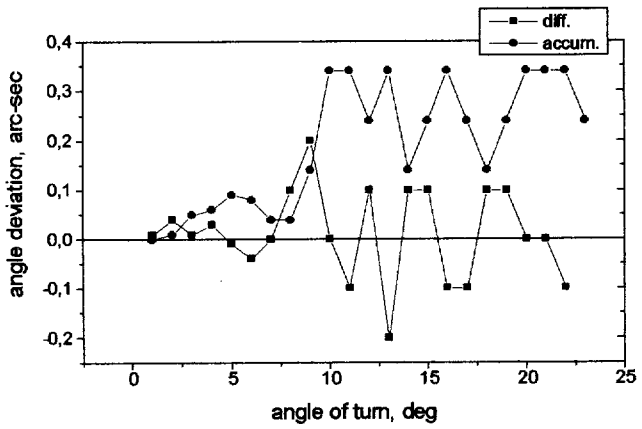


Fig12..22. The results of the tests of the DLG in the external measurement mode by means of comparison with the PTB standard turntable.

As an example of the method using we will discuss the angle position measurement of the three-axis test-bed in the German Aerospace Research Establishment (DLR, Braunschweig, Germany). Such measurements allow to calibrate inductosines controlling the table position. The measurement scheme is shown on fig.12.23. The OP 2 is fixed on the turning table 1. The OP face is used as a controlled mirror. The laser goniometer 4 is situated on a separate basis. Rotating polygon of the DLG provides the beam scanning and its incident on the surface A of the polygon 2. The table was turned consequently to several angle positions fixed by the inductosine and measured by the laser goniometer.

The angle scanning range on the distance of 1785 mm over the control reflecting surface (the polygon face) 50*50 mm was 48 arcmin. The error of measurements appeared to be on the level of one tenth of arcs (SD). These results clearly showed advantages of the laser goniometric method comparing to the very hard in using autocallimation method.

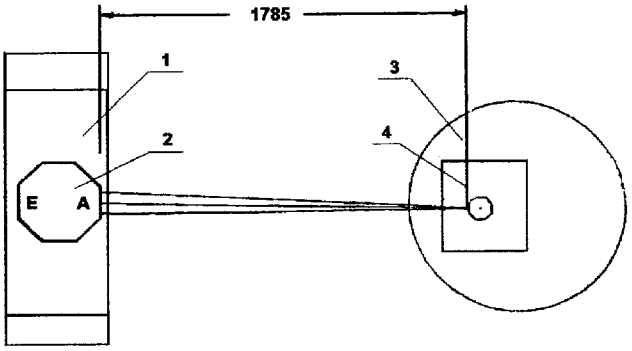


Fig.12.23. The scheme of the DLG use for calibration of the inductosine of the DLR tree-axis test table

3.3.3. Using DLG for estimation of the object angular rate

Development of research and design of the DLG were mostly defined by needs of the industry to estimate the gyro-system drift. Previously used means for estimation of the angle rate of the object were based on method of fixed angle or fixed time [19].

In both cases the object angle rate is estimated using the equation $\overline{\Omega}_\rho = \frac{\Delta\rho}{\Delta t}$, where $\Delta\rho$ - object rotation angle, Δt - measurement time. The precision of time intervals measurements is high, therefore the error of the rate $\overline{\Omega}_\rho$ estimation is mostly caused by the angle measurement error. If the object movement is "smooth" the results of modern autocollimator implementation (the error $\Delta\rho$ in this case is about several tenth of arcs) are good. To get the error of estimation of $\Delta\overline{\Omega}_\rho$ equal 0.001deg/h, while $\Delta\rho$ is equal to 0.1arcs, it's enough to have measurement time about 100 s. The situation becomes much more complicated in the presence of angle oscillations of the object, which are usual for gyro-stabilised objects. If the amplitude of the object angle oscillation is about 5-10arcs, then the measurement time which is necessary to have for the same error would be (0.5-1)10^4 s. That is impossible for gyro-devices applied on short-living objects (for example, rockets).

Therefore the main advantage of the DLG comparing to conventional measuring instruments is high data reading frequency combined with a wide range of angle measurements. That allows to get big arrays of the object angle position values as a function of time $\rho_i(t_i)$. These features of the DLG provide two advantages: they allow to determine the spectrum of the object angle rate oscillations and to use various statistic methods to estimate the average angle rate, the least square method in particular. In that case when the value of the measured angle is written as $\rho_i = \overline{\Omega}_\rho \cdot t_i$, where ρ_i, t_i - current angle and time, the estimation of the average angle rate $\overline{\Omega}_\rho$ can be done according to the equation

$$\overline{\Omega}_\rho = \frac{m \cdot \sum_i t_i\rho_i - \sum_i t_i \cdot \sum_i \rho_i}{m \cdot \sum_i (t_i)^2 - (\sum_i t_i)^2},$$

where m - number of measurements in a series.

Dispersion of the angle rate estimation of such method can be defined as

$$D_{\overline{\Omega}} = \frac{12D_\rho}{T_m^{\,3} \cdot F_u},$$

where T_m - measuring time, F_u - data reading frequency. If angle oscillations of the object are considered as the angle measurement error, then the required measurement time as a function of the angle measurement error has a shape, presented in fig.12.24. Here we supposed the required error of the rate estimation to be equal to $\sigma_{\overline{\Omega}}$=0.001 arcs/s. Shown curves prove considerable advantages of the DLG in comparison with traditional means of measurement. If the object movement is smooth and the error is caused by the own error of the DLG, the result of estimation process is described by the fig.12.25. The figure proves that the required precision of the angle rate estimation can be achieved even with relative big errors of the angle measurement (about 1arcs). The fig.12.25 shows also that increasing of the data read frequency more than 30-40 Hz does not improve effectively characteristics of the process of the object rate estimation. That means that a high data reading frequency (more than 50 Hz) is necessary only to get the information about the spectrum of the object angle movement.
Analysed characteristics of the DLG are influenced mostly by the random error of the angle measurement. Now we will analyse shortly the influence of the systematic part of the angle measurement error. It is obvious that the systematic component, constant in time and not depending on the measured angle value, does not contribute in the angle rate estimation error.

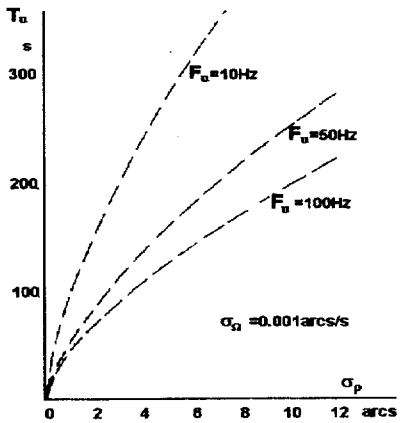


Fig.12.24. The required measurement time of angular rate estimation as a function of the angle measurement error.

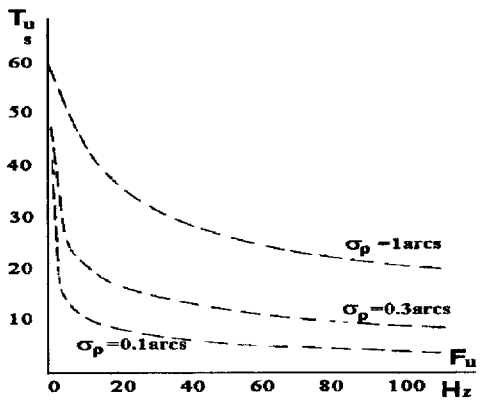


Fig.12.25. The required measurement time of angular rate estimation as a function of the data read frequency.

If the systematic component is changing during the measurement and a value of change is $\Delta\rho'$, then we have the direct contribution to the estimation error $\Delta\overline{\Omega}_\rho$

$$\Delta\overline{\Omega}_\rho = \frac{\Delta\rho'}{T_m} .$$

With the integration time about 100 s to get the error 0.001 deg/h the value of $\Delta\rho'$ should be in the limits of 0.1arcs. That makes important the requirement to the mirror flatness.

3.4. DLG in X-ray two crystal diffractional spectrometer of St.Petersburg Nuclear Physics Institute

One of the main directions of research into the determination of fundamental physical constants is the study of π-mesoatoms in a synchrocyclotron. The purpose of this research is to measure hyperfine structure of 2p-levels of π-atoms of various elements with high accuracy. One of the best ways of such research is the use of the X-ray radiation of the energy $E_x \leq$ 15keV on the two-crystal diffractional spectrometer operating in reflection mode. Measurement of radiation diffraction angles in spectrometers usually is carried out by means of optical encoders or laser interferometers [20]. The best optical encoders (for example RON 905 of Heidenhain) have an accuracy of not better than 0.2 arcs and ttis is not sufficient for the spectrometer angle measurement system. The shortcomings of laser interferometers are non-linear scale and restricted range of angle measurements, that doesn't allow measurements with 180°-turn of crystals.

On the other hand the results obtained during the last years in the development of the laser goniometer systems show remarkable capabilities of such systems in various applications [9,10]. As it was shown in the previous chapters the DLG provides the accuracy of the angle measurements on the level of several handreds of arcsecond in the range of the measurements 0...360 deg. It was the reason to undertake an attempt to develop the DLG for diffractional spectrometers.

3.4.1. Principle of measurement

Unlike published earlier designs, the laser goniometer in this case includes two coaxial turn tables (fig.12.26). One of them is the goniometer rotary table with the ring laser and the null-indicator (NI). Both of them rotate together with the table platform. The other is the table of the diffractional spectrometer with the crystal (Cr) positioned on it. The control mirror (CM) is situated on the same table. The reference direction of the system is formed by the normal to the reference

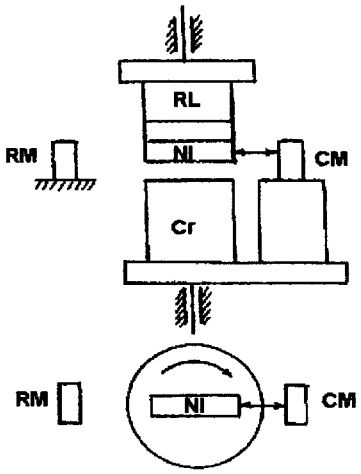


Fig.12.26. The scheme of the laser goniometer system in the diffractional spectrometer.

mirror (RM) which is fixed to the base. The NI generates pulses whenever its light beam impinges normal to RM and to CM and the interval between these pulses defines the angle formed by RM and CM. To measure this angle, the described above procedure of counting the RL output pulses is used. This angle can be considered as the angle between some direction of the crystal and the reference direction. The development of this system with the NI rotating
together with the RL allows to remove any restriction on the angular range of the measurements. Another advantage of the NI's location on the rotary table is that its optical axis crosses the axis of rotation and as a result the light beam impinges to the mirror surface always at the same place. In this case there is no displacement of the beam upon the mirror surface during the measurement and the nonflatness of the mirror does not cause the error of the measurements.

3.4.2. Results of measurements

The main sources of the measurement error in this system are the errors of the RL induced by its parameter fluctuations and the random and systematic errors of the NI. The random error of the NI is mostly determined by the signal to noise ratio of the NI output signal. This source of the error was reduced using a special procedure which consisted of the use of a fast analogue-to-digital converter (FADC) of NI and RL output signals. The FADC scheme provides the registration of the NI and RL output signals with the frequency of 20 MHz that corresponds to an angular resolution of about 0.04 arcs for the rotation rate of 240 deg/s. The output signals of the FADC were fed to a personal computer for corresponding processing. Fig.12.27 shows the results of the digital registration of the RL signal (curve 1) and the NI signal (curve 2). The signals were analysed in the real time. The processing of the NI signal consisted of the determination of the pulse centre of gravity.

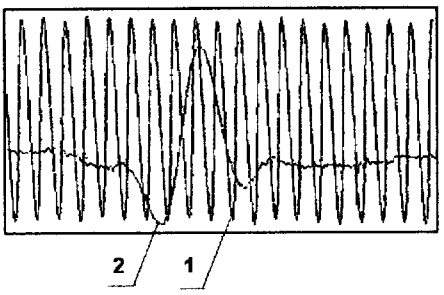


Fig. 12.27. The results of the digital registration of the NI signal (curve 1) and the RL signal (curve 2).

Such processing allowed to reduce the influence of the noise in the NI output signal. The purpose of the RL output signal processing was to determine the value of the time interval between the NI pulse centre of gravity and the nearest edge of the RL pulse. Simultaneously the integer number of the RL output signal periods in the interval between the NI pulses was determined. Such processing allowed to get the random error of the null-indication about 0.015 arcs.

The random error induced by fluctuations of the RL parameters was estimated with the special experiments. The measurement set-up in these experiments consisted of the airbearing rotary table, The Heidenhain optical encoder RON 905 and the ring laser. The RON 905 was implemented to form the angular intervals of various length. These angles were measured by the RL in series of measurements. The standard deviations of every series was accepted as the random error of the RL (one should take into account that in reality such standard deviation consists of the random errors of the RL and also of the RON 905). The value of measured angle was varied from 10 deg to 360 deg. The RL used in the measurements had He-Ne active medium, $\lambda = 633$ nm, monolithic square cavity structure with four total reflecting prisms, 80 cm perimeter and an angle resolution α about 0,6 arcs. To improve the resolution the described above technics was used. Fig.12.28 shows the dependence of the measurement random error (standard deviation) on the value of measured angle. It is worth to be noted that for small angles (about 10 deg) a SD does not exceed 0,03 arcs. Our estimations show that this value can be considered as the random error of the RON 905.

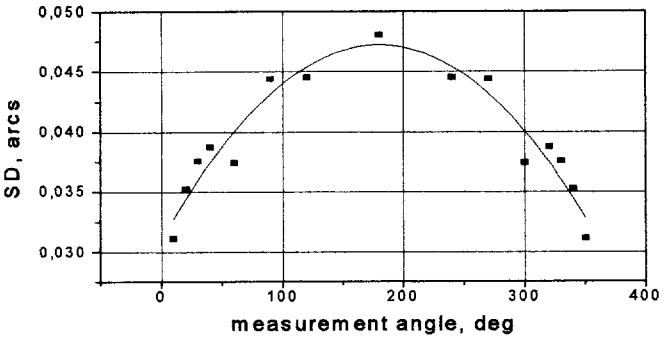


Fig. 12.28. The dependence of the measurement random error (standard deviation) on the value of measured angle.

From the measurement results it can be concluded that the utmost random error of the RL does not exceed the level of 0,05-0,06 arcs. One should take into account that this value correspondents to a SD of a single measurement. For the series of 100 measurements we will have a SD of the mean on the level of 0,005 arcs (the time of 100 measurements in our case is about 2 min).

The described above (fig.12.26) system was implemented in the consist of the two crystal diffractinal spectrometer. In experiments we used the X-ray line of radiation with the energy of radiation E = 17,47334 keV. The result of the measurement is presented in the Fig. 12.29. Every point of the curve is the mean of the 10 measurement series.

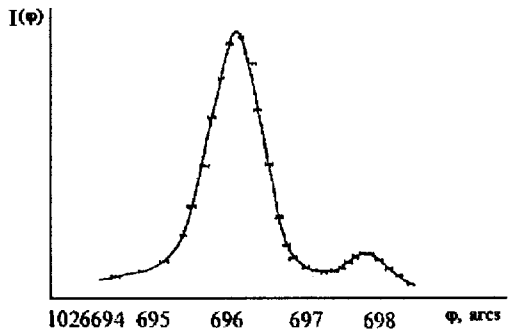


Fig.12.29. The results of the measurement of the X-ray line by means of the diffractional spectrometer with the laser goniometer system.

4. PERSPECTIVE OF DEVELOPMENT AND APPLICATYION OF THE DLG

4.1. Inertial angle measurements by means of the DLG

In the "inertial mode of operation", owing to the feature of the RL as an inertial rotation sensor, the DLG measures directly angular displacements of the object on which it is mounted. This measurement method offers the advantage that it is free from any need for external coupling or reference elements and independent of the location of the rotary axis.

The first tests of this method was made using the DLG, equipped with the Heidenhain RON 255 optical encoder. The instrument was placed on the standard rotary table of the PTB as described above, which allows positioning with an angular accuracy of 0,1 arcs.

The principle of inertial angle measurement, as applied in this special case, is elucidated by the diagram of Fig.12.30.

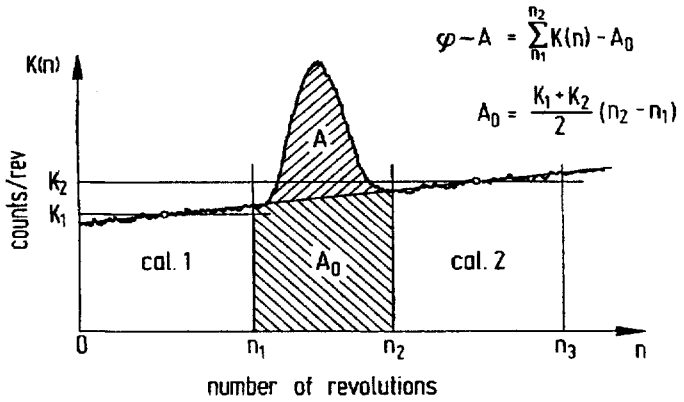


Fig.12. 30. The principle of inertial angle measurement.

First the ring laser scale factor is determined when the rotary table is in the rest position. This is achieved by counting the RL pulses during several revolutions of the goniometer spindle, using the 360 deg reference from the pulses of the optical encoder coupled to the spindle. Immediately after this calibration, the table rotates the goniometer through a certain angle and again comes to the rest position. The integral number of the RL pulses counted during the rotation minus the stationary calibration value is a measure of the angle of rotation as is made clear by Fig.12.30: The integral $\sum K(n)$ of the RL counts in the interval between the spindle revolutions n_1 and n_2 minus the interpolated stationary integral A_o yields the integral A as a value proportional to the inertial angle of rotation φ. In the new rest position after the inertial measurement, another calibration of the scale factor can be carried out. If a linear drift of the scale factor between the first and the second calibration is assumed, the mean scale factor may be taken into account for subtraction during the measurement interval. Especially if the calibrations performed before and after the measurement are taken over equal numbers of revolutions (n_3-n_2=n_1), the mean calibration values K_1 and K_2 can be used to calculate A_o= ½ (K_1+K_2)(n_2-n_1), thus correcting completely a linear drift.

This principle shows that the stability of the RL scale factor and the linearity of its drift determine largely the accuracy which can be achieved in inertial measurements. In our measurements, the RL had been operated for a sufficiently long time under laboratory conditions, the scale factor usually had a relative drift of about 10^{-7} per hour and a random noise with a relative SD of the order of $(1{,}5 \text{ to } 2{,}0) \cdot 10^{-7}$, which is, in units of angle, ≈ 0.2 arcs. It is worth noting that this scale factor noise not only is due to the variation of the RL parameters but can be caused by variations of the optical encoder readout due to vibrations and lateral motions of the spindle axis.

We checked the inertial angle measurements using the DLG in series of repeated angular steps of 1, 2, 5 and 10 deg of the standard rotary table. The deviations of the inertial measurements from the reference values of the rotary table were evaluated (deviation = goniometer readout minus reference value). The numbers of revolutions for the calibration and measurement intervals were selected according to the automatic run of the rotary table, and maintained for all measurements during one series.

Fig.12.31 shows the results of four series of 1 deg measurements made in both directions of rotation. In all cases, the calibration times in the range from 40 to 90 s were longer than the measurement times. Due to different speeds of rotation, these varied between 12 and 30 s. The resulting angle deviations had a total mean value of -0,01 arcs and a standard deviation SD = 0,40 arcs as indicated in Fig.12.33 by dotted lines.

Other results obtained for 2,5 and 10 deg step measurements had a quality similar to that of the 1 deg measurements with a slightly lesser accuracy. For 2 deg we got a mean deviation of 0,10 arcs with SD = 0,46 arcs, for 5 deg a mean deviation of 0,51 arcs with SD = 0,50 arcs and for 10 deg a mean deviation of 0,07 arcs with SD = 0,77 arcs.

Let us consider potential sources of the uncertainties obtained. First they can be connected with the RL scale factor drift. As was pointed out above, the relative drift of the scale factor was about 10^{-7} per hour. For the time interval of our measurements (usually not more than 100 to 150 s, together with the calibration time), the uncertainty connected with the scale factor drift should therefore be of the order of 0,1 arcs or smaller.

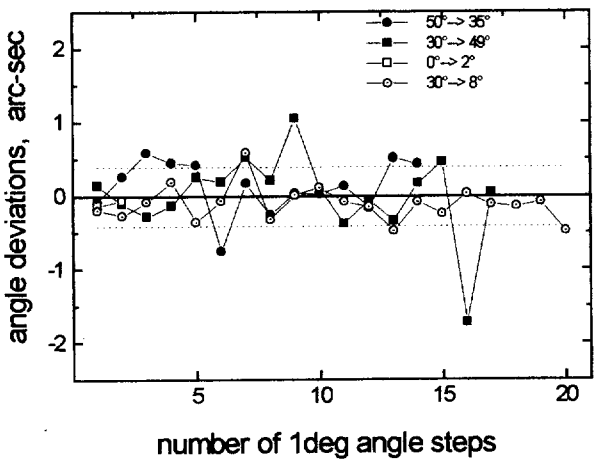


Fig.12.31. The results of four series of 1 deg measurements made in both directions of rotation.

The second source of uncertainty is the scale factor noise. It consists of a noise from the readout variations of the optical encoder forming the counting interval and the random noise of the RL output signal. The noise from the former is not accumulated during the measurement. However, the integration of the RL noise during the inertial angle measurement gives rise to a random walk uncertainty to be expressed as a standard deviation following a square root law with time. This assumption was confirmed experimentally by repeated integrations of the RL signal in the stationary condition over time intervals of different length. The curve of the standard deviation obtained was fitted by the expression

$$SD = 0{,}14\text{arc-sec} + (0{,}047\text{arc-sec}/\sqrt{s})\sqrt{t}.$$

The uncertainty of the inertial angle measurement due to the RL random noise was estimated on the basis of scale factor data recorded over a long time. The recorded data were processed in accordance with the previously described scheme of the measurement procedure. Fig.32 shows the calculated mean standard deviation of the simulated measurements as a function of the measurement time in seconds, for calibration times of 30 s, 60 s and 80 s.

The SD rises as the square root of the measurement time, with the calibration time showing a significant nonlinearity for the decrease of the SD. The values are experimentally confirmed by the results of our inertial angle measurement. It is clear from these results, that the uncertainty of this method depends significantly on the time duration of measurement and calibration as well as on the level of noise in the RL output signal.

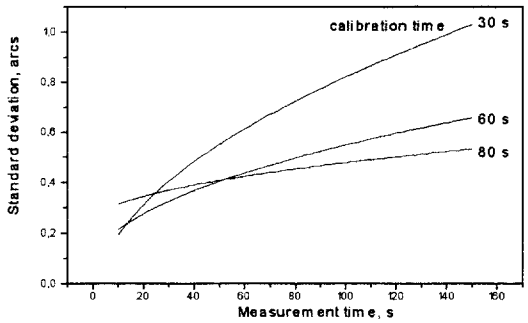


Fig.12.32. The calculated mean standard deviation of the simulated measurements as a function of the measurement time.

4.2. DLG in technology of circular scales production

The errors in angular measuring devices, in particular in optical encoders, are mainly caused by graduation or dividing errors. Conventional ways of producing high precision circular scales or optical gratings are based on static methods. In these methods gratings are produced with diamond cutting or photolithography techniques. The main point of the producing process from the point of view of accuracy is the dividing procedure or the turn of an instrument on the required angle. Usually this procedure is provided with a precision dividing machine or a precision turntable [21]. Disadvantages of static methods are connected with the need of using standard circular scales for measurement of angles in dividing machines or turntables. So the accuracy of circular scales production is provided with the accuracy of the circular scale which is chosen as a standard but also was produced with the same procedure and has the same type of errors.

The use of the ring laser can be considered as new concept of the circular scale recording process. The ring laser offers the advantage of a nearly perfect circular division due to the laser wavelength as its base. In the case of the ring laser we have due to the interaction of two opposite propagating waves a structure of a standing wave in the cavity of a ring laser [22].This standing wave structure in ideal is immovable in reference to inertial space and so it can be used as a high precision circular scale with resolution for usual size of a ring laser and without any interpolation about 1 arcs.

Made above description of the rotating RL operation makes clear that this instrument can be implemented successfully in the technology of the circular scales production. In this procedure the RL can be used as a reference encoder with perfect angular division.

It is possible to realise several kinds of systems on the base of the RL for producing precision circular scales. In one version the method consists in direct recording the interference pattern formed by waves counterpropogating in the ring laser cavity [23] (Fig.12.33).

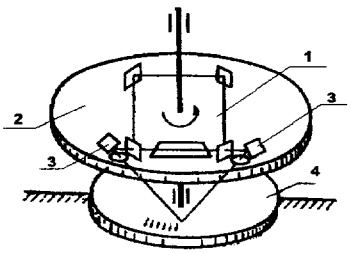


Fig.12.33. The method of direct recording the circular scales by means of the interference pattern formed by waves counterpropogating in the ring laser cavity.

In that case the ring laser is set up on the rotary platform 2, and the output beams are handled by an optical system 3 and brought to the photosensitive surface of the immovable scale blank 4. On the photosensitive surface of the scale blank there is light spot with the interference fringes of the radial direction. The dependence of light intensity in the spot in tangential direction can be expressed as

$$I = I_0(1 + \cos(\omega t + \frac{4\pi}{\lambda} x \sin\theta)),$$

where θ - angle of incidence of the RL beams on the photosensitive surface.

From this expression it is easy to get the velocity of the fringe in relation to the light spot:

$$v_x = \frac{\omega\lambda}{4\pi\sin\theta}.$$

In the same time the light spot moves upon the photosensitive surface with the velocity $v_x = \Omega r$ due to rotation of the RL (where r - distance from the center of the spot to the rotation axis). These two velocities have the different signs. When radius r is equal to

$$r = \frac{\omega\lambda}{4\pi\Omega \sin\theta}$$

velocities are equal to each other, have different signs and owing to it the interference fringes are immovable in relation to photosensitive surface. When the RL rotates, immovable interference fringes provides at the photosensitive layer the structure of radial lines, whose number is determined by the value of the radius r. The number of lines can achieve the value compared with the scalefactor of the RL. So as the result the circular scale produced by this method could have the resolution on the level of 1-2 arcs without any interpolation. As no ring laser signal conversion is involved, one has a simple system that eliminates various sources of error. Also, one can use a photosensitive medium with a high resolution (over 1000 mm^{-1}), which avoids the loss of part of the information arising from ring laser signal frequency division.

In the second version the technological powerful laser is used. In this system (Fig.12.34) the ring laser 1 and the blank for the circular scale 2 are mounted on a common rotary platform 3. During the rotation of the RL its output signal is fed to an instrument 4 for producing lines on the scale blank. A powerful laser can be applied as such instrument. The radiation of this laser is formed by optical system 5 for producing lines of desirable size and form. For example it can be a nitrogen laser with ultra-violet radiation and photoresist layer on the scale blank. The required scale resolution is provided by incorporating a frequency divider 6.

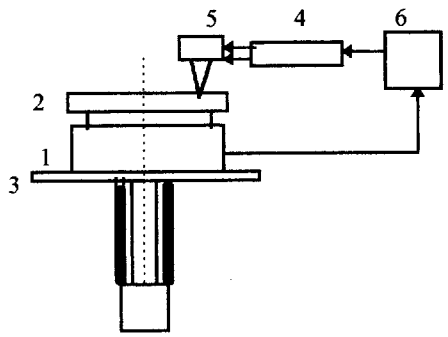


Fig.12.34. The scheme of the laser goniometric system for the circular scales recording with the use of the technological powerful laser.

The advantage of this version is the use of the photoresist as a photosensitive medium. In comparison with the first version, where emulsion is used as the photosensitive medium, the use of the photoresist provides the firm and reliable structure. The circular scales recorded with this method can be used in production of the optical encoders manufactured by Heidenhain and other firms.

4.3. Necessity and perspectives of creation of plane angle standard in dynamics on the base of the DLG

The results of the DLG investigations have proved that angle measurement with use of a RL is an effective and versatile method for applications such as calibration of optical polygons or optical encoders, external angle measurement of a rotating object or inertial angle measurement. The errors of these methods compared with standard angle measurements are in the range from 0,03 to 0,2 arcs. Except the inertial method, where the error depends strongly on the time of the measurement, three other applications described above show values of error in the range from 0,02 to 0,08 arcs The quantum noise level of the ring laser is up to now not approached as the limit of accuracy so that an improvement can be expected from optimisation of the ring laser parameters. The most perspective way of increasing the measurement accuracy is the development of the special metrological ring laser.

It was shown in some papers [24,25] that the powerful mean to increase the DLG accuracy is the use of the ring laser with the large perimeter. Extending the RL perimeter leads to diminishing of all components of the laser goniometer error. The dependences of the error components via the perimeter L are:

1. Quantisation error 1/L
2. Bias instability $1/L^3$
3. Output response non-linearity due to lock-in $1/L^5$
4. Random drift $1/L^2$

The estimations of error for the ring laser, for example, with perimeter L=160 cm (usually we have in the DLG the RL with perimeter L=40 cm) give us the next values:

1. Quantisation error will be reduced to 4 times and with the output signal frequency multiplication (k=100) will be about 0,003 arcs.
2. Bias instability will be reduced to 64 times and can be neglected.
3. Output response non-linearity due to lock-in also can be neglected.
4. Random drift will be reduced to 16 times and will be about 0,003 arcs/√s.

All this will give the opportunity of angle measurements with the accuracy at the level of 0,001 arcs. Owing to reduction of the output response non-linearity it will be possible to reduce the rotation rate of RL to the values about 0,3 rad/s.

In this case that ring laser, which can be called „Metrology Ring Laser" (MRL), should be investigated at the best angle comparator (for example new angle comparator of the PTB) as high-accuracy standard of a circular division. According to the angle comparator, the MRL will attain an angle resolution and an uncertainty of angle measurement of 0,001-0,005 arcs.

REFERENCES

1. Chow W.W., Gea-Banacloche J., Pedrotti L.M., Sanders V.E., Schleich W., Scully M.O., Rev of Modern Phys. 1985, 57, pp 61-104.
2. Anderson R., Bilger H.R., Stedman G.E., Am. J. Phys. 1994, 62, pp 975-985.
3. Catherin J.M., Dessus B., French Patent No.1511089; issued 26.01.68.
4. Blanter B.E., Filatov Yu.V., Metrology (USSR), 1979, pp 3-9.
5. Vanyurikhin A.I., Zaitsev I.I., Sov. J. Opt. Technol., 1982, 49, pp 566-569.
6. Mokros J., Vu K.X., Jemna mechanika a optika, 1987, 32, pp 18-19.
7. Xu Y., Liao F., Dai R., Shen S., Situ Z., Acta Metrologica Sinica, 1985, 6, pp 261-262.
8. Gelashvili N.V., Birkadze Sh.V., Kuryatov V.N., Orlov M.V., Measurem. Techn., 1988, pp 851-853.
9. Filatov Yu.V., Lukyanov D.P., Pavlov P.A., VDI-Berichte Nr.1118, 1994, pp 123-128.
10. Lukyanov D.P., Pavlov P.A., Filatov Yu.V., Symposium Gyro Technology, Stuttgart, 1991, pp 4.0-4.12.
11. Ulrich D., Lubeck U., Wetzig V. Simp. Gyro Tehnology, 1988, Stuttgart.
12. Filatov Yu.V. Doct. Dissert. Thesis, St.Petersburg, 1991.
13. Dorschner T.A., Haus H.A., Holz M., Smith I.W., Statz H., IEEE J.of QE, 1980, *16*, pp 1376-1379.
14. "Space radiotechnical complexes", ed. by S.I.Bychkov, Moscow, Sov.radio, 1967 (in russian).
15. "Laser measurement systems", ed. by D.P.Loukianov, Moscow, Radio i svjaz, 1981 (in russian).
16. Zhuravleva E.N., Kuryatov V.N., Semenov B.N., The 2nd Saint Petersburg Internat. Conf. on Gyroscopic Technology and Navigation, Part 2, St.Petersburg, SCRI "Electropribor", 1995, pp 57-59.
17. Probst R., VDI-Berichte Nr.1118, 1994, pp 173-178.
18. Sim P.J., in *Modern Techniques in Metrology* (edited by P.L.Hewitt), Singapore, World Sientific, 1984, pp 102-121.
19. Ivanov V.A. "Metrological support of gyro devices", Leningrad, Sudostroenie, 1983 (in russian).
20. W.Schwitz, Nuclear instruments and methods, 1978, 154, pp 95-104.
21. Ernst A. "Digital linear and angular metrology", Die Bibliothek der Technik, Band 34, Verlag Moderne Industrie, 1989.
22. Schulz-DuBois E.O., IEEE J. of QE, 1966, *2*, pp 299-305.
23. Krivtsov E.P., Loukianov D.P., Filatov Yu.V., Opt. Spekrosk. July 1986, *61*, pp 144-148,.
24. Stedman G.E., Bilger H.R., Li Ziyuan, Poulton M.P., Rowe C.H., Vetharaniam I., Wells P.V. Austral. J. Phys., 1993, 46, pp 87-102.
25. Bilger H.R., Stedman G.E., Li Ziyuan, Poulton M.P., Rowe C.H., Wells P.V., IEEE Trans. on Instrum. and Measur., 1993, 42, pp 407-411.

13. Gyroscopic Goniometry During Quasistatic and Kinematic Conditions Using a New and Conventional Methods for Aiding

B. Stieler

Institute of Flight Guidance,
German Aerospace
Research Establishment (DLR)

and

Institute A of Mechanics,
University of Stuttgart, Germany

Abstract

Gyroscopic goniometry necessitates the calibration of scalefactor error and drift. Conventional on-line and off-line methods are discussed based upon external angle and angular rate information - zero velocity update (ZUPT) in the latter case. A new method is also presented - the Differential Inertial Measurement Technique (DIMT) - which allows to calibrate time-dependent measurement errors during recurring but identical measurement conditions of the inertial sensor or system. No absolute external references are required, since DIMT handles measurement differences at different points in time. DIMT can be regarded as an extension to ZUPT, but without the constraint of the sensor or system being at rest. It opens a great variety of applications. All methods are verified with the quasistatic measurements of the angles of a quadrangle, the measurements of angles during recurring motions like in a robot or during tests of railway tracks and during measurements of angular characteristics like those of a windtunnel model, i.e. the dependence of lift upon the angle of incidence. A fiberoptic gyro is used as sensor.

1. Introduction

Inertial measurement accuracy improvement is based on calibration procedures before and/or after the mission or on aiding procedures during the mission.

Calibration procedures before and/or after the mission make use of the fact that known external references are then conveniently available. A disadvantage of this calibration procedure lies in the fact that the accuracy of the subsequent mission will be limited, if sensors are used whose error parameters are a function of time. The effect of this kind of calibration relies on a certain stability of those values during the mission.

Aiding procedures during the mission can successfully be applied even if sensors of inferior quality, i.e. with instable error parameters are used. A disadvantage of these procedures is, of course, the provision of such measurement data either from external instruments, or from "zero velocity updating (ZUPT)", i.e. stopping the motion for taking the measurement signal gradient with respect to zero velocity as calibration information. Such measurements are often out of range.

"Differential Inertial Measurement Techniques (DIMT)" described below and documented in more detail in [St 94] can be regarded as an extension to ZUPT. It handles the difference between measurements of one sensor or system at different points of time when identical measurement conditions (same location, velocity or angle) exist. With zero velocity this is always the case, but identical measurement conditions may also occur during motion. During repetitive motions identical locations or directions will be passed, for instance.

Thus time-dependent errors of sensors and systems can be calibrated and compensated as soon as a second measurement is at hand. Since measurement differences are used for this kind of calibration, no absolute reference values are required, they remain unknown. This opens the way for a great variety of applications. DIMT thus constitutes a valuable supplement to the calibration and aiding procedures especially with respect to ZUPT.

In order to illustrate the procedures just mentioned, the measurement of angles and angular characteristics with a fiberoptic gyro (FOG) of medium accuracy is taken as an example. "Angular characteristics" is understood as the dependence of angles upon forces and moments as in elastic deformations or the dependence of forces and moments upon angles as in aerodynamics.

2. The Gyro Error Modelling

The sensor employed is a fiberoptic gyro kindly provided by ALCATEL SEL, Stuttgart, Germany. Its development is now taken over by SFIM Germany, GmbH. The sensor and a test setup are shown in Fig. 2.1. Type and performance data are listed in Tab. 2.1. The data are typical for a sensor of the late 1980 ies.

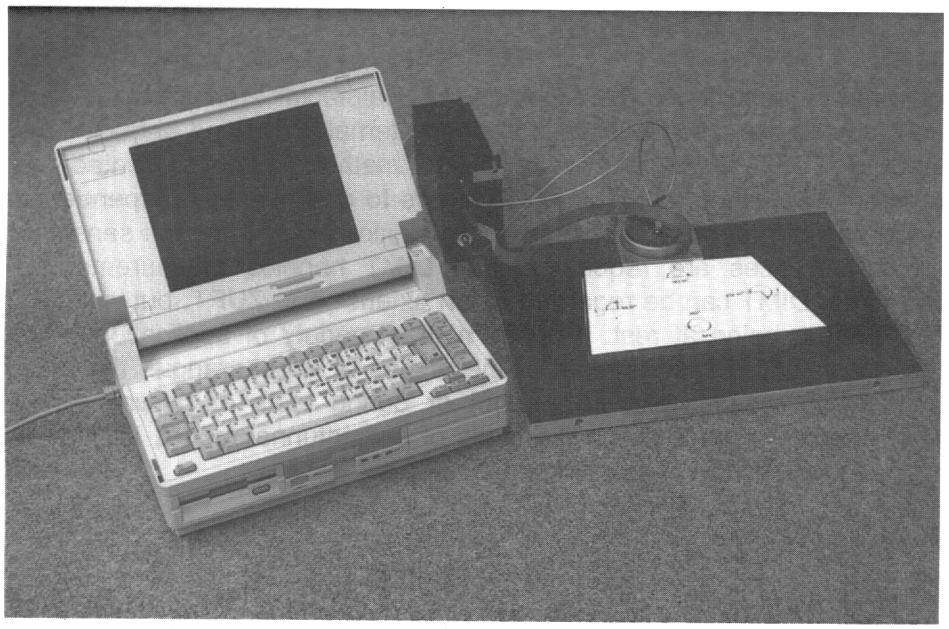


Figure 2.1 Setup for Measuring the Angles of a Quadrangle with a Fiberoptic Gyro (FOG)

Fiber length	100 m	
Open loop phase modulation		
Drift	< 10 deg/h	(1σ)
Random walk	< 0,2 deg/$\sqrt{h}$	(1σ)
Scalefactor error	< 5000 ppm	(1σ)

Table 2.1 Performance Data Published by Alcatel SEL for the Fiber-Optic Gyro PM1DE 9100147.

When a gyro is used as a goniometer for measuring the angle α its output signal $\omega^* = \dot{\alpha}^*$ is fed to a computer for integration. The * denotes a computed or measured value which includes errors. We assume the following model for this process:

$$\alpha^* = \alpha + \varepsilon = \int \left[(1+\kappa)\,\dot{\alpha} + d\right] d\tau, \tag{2.1}$$

with κ = scalefactor error and d = gyro drift.

Assuming κ to be quasi-constant and d to consist of a bias plus contribution of earth rate D_1 a drift gradient D_2 and stochastic noise w, we obtain for the angle-dependent and the time-dependent measurement errors:

$$\varepsilon_\alpha = \kappa\alpha \tag{2.2a}$$

$$\varepsilon_t = D_1 t + D_2 t^2 / 2 + \int w(t)\, d\tau \tag{2.2b}$$

and

$$\varepsilon = \varepsilon_\alpha + \varepsilon_t \,. \tag{2.2c}$$

The integral of the stochastic noise is known as random walk, whose $\pm$ 1σ bands follow the relationship:

$$\sigma(\varepsilon) = r\sqrt{t}\,. \tag{2.3}$$

(r = random walk coefficient).

As example for the growth of the integrated FOG error the results for 10 runs over 1h are shown in Fig. 2.2 with a constant value of 10 deg/h due to earth rate and drift being subtracted. In Figure 2.3 the residues of two runs against a first order slope due to constant earth rate plus drift D_1 plus a second order slope due to drift gradient D_2 are plotted together with the $\pm$ 1σ bands based on an estimated random walk coefficient r. D_1 and D_2 are estimated by taking for one run a great number of $\Delta\varepsilon_{j,i} / \Delta t_{j,i}$ [$\Delta\varepsilon_{j,i} = \varepsilon(t_j) - \varepsilon(t_i)$, $\Delta t_{j,i} = t_j - t_i$] at varying times $t_{j,i}$ and feeding them into a regression algorithm with $(t_j + t_i)/2$ as independent variable. From the differences of $\Delta\varepsilon_{j,i}$ against the angle increments of the model divided by $\sqrt{\Delta t_{j,i}}$ the random walk coefficient was found as the standard deviation. Fig. 2.3 indicates that the error model chosen above hardly comprises reality in all cases, but it is the only practicable one for conventional Kalman filtering, especially since an additional state in the error model does not furnish a considerable improvement.

Table 2.2 shows the results of the 10 runs in terms of mean and standard deviation for drift D_1

Drift (mean $\pm$ 1σ)	[deg/h]	:	-0,4 $\pm$ 1,2
Driftgradient (mean $\pm$ 1σ)	[deg/h^2]	:	1,9 $\pm$ 3,4
RMS [Random Walk]	[deg/$\sqrt{h}$]:		0,084
RMS [$\varepsilon(t)$]	[deg]	:	0,069
RMS [ε_{max}]	[deg]	:	0,134

Table 2.2 Performance Data for the SEL Fiber-Optic Gyro PM1DE 9100147 as Result from 10 Runs over 1h

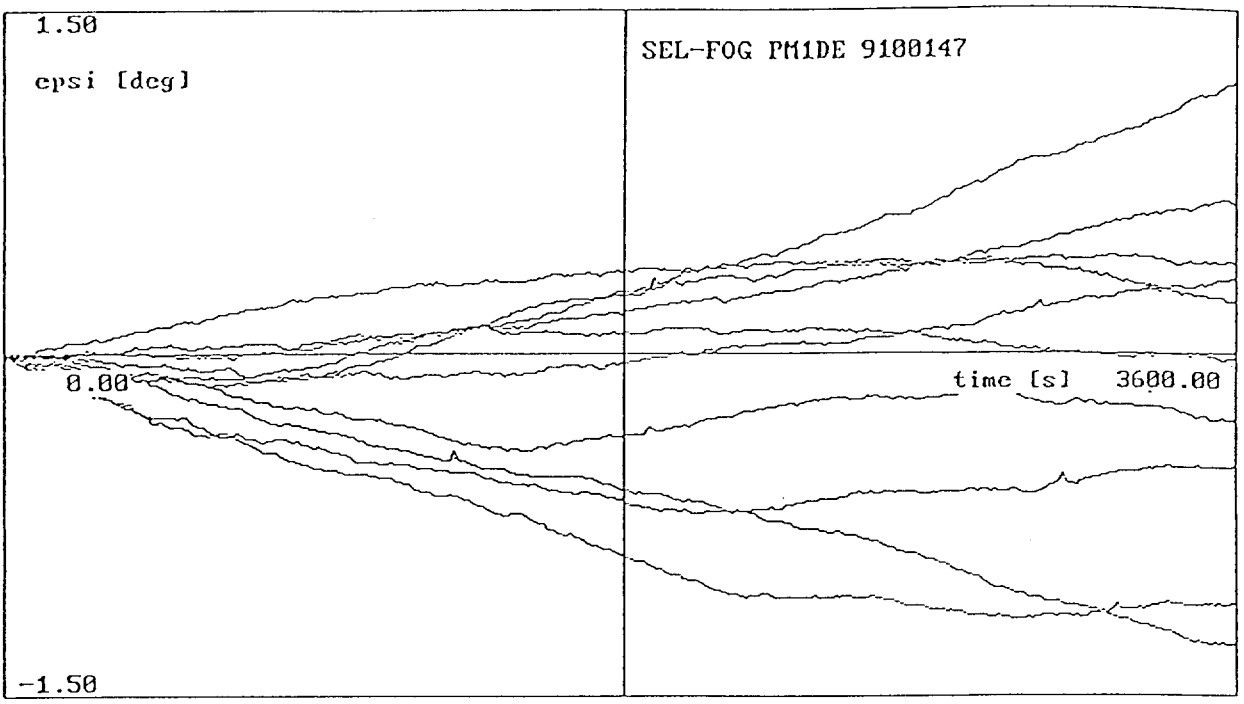


Figure 2.2 Integrated Output Signals of the Stationary FOG for 10 Runs

and driftgradient D_2 and in terms of the RMS values for the random walk r, for the angle residues $\varepsilon(t)$, over time as well as for the peaks ε_{max} of the residues. D_1, D_2 and r will serve us for choosing proper input parameters for the Kalman filter and RMS $[\varepsilon(t)]$ and RMS $[\varepsilon_{max}]$ for judging the estimation results in Section 6.

For better understanding DIMT error modelling discussed in the next section, the basics of system error prediction will be reviewed. Eq. 2.2. describing the system error dynamics written in matrix-vector notation is:

$$\begin{pmatrix} \dot{\varepsilon} \\ \dot{D}_1 \\ D_2 \\ \dot{\kappa} \end{pmatrix} = \begin{pmatrix} 0 & 1 & t & \dot{\alpha} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \varepsilon \\ D_1 \\ D_2 \\ \kappa \end{pmatrix} + \begin{pmatrix} w_d \\ w_{\dot{d}} \\ w_{\ddot{d}} \\ w_{\dot{\kappa}} \end{pmatrix} \tag{2.4a}$$

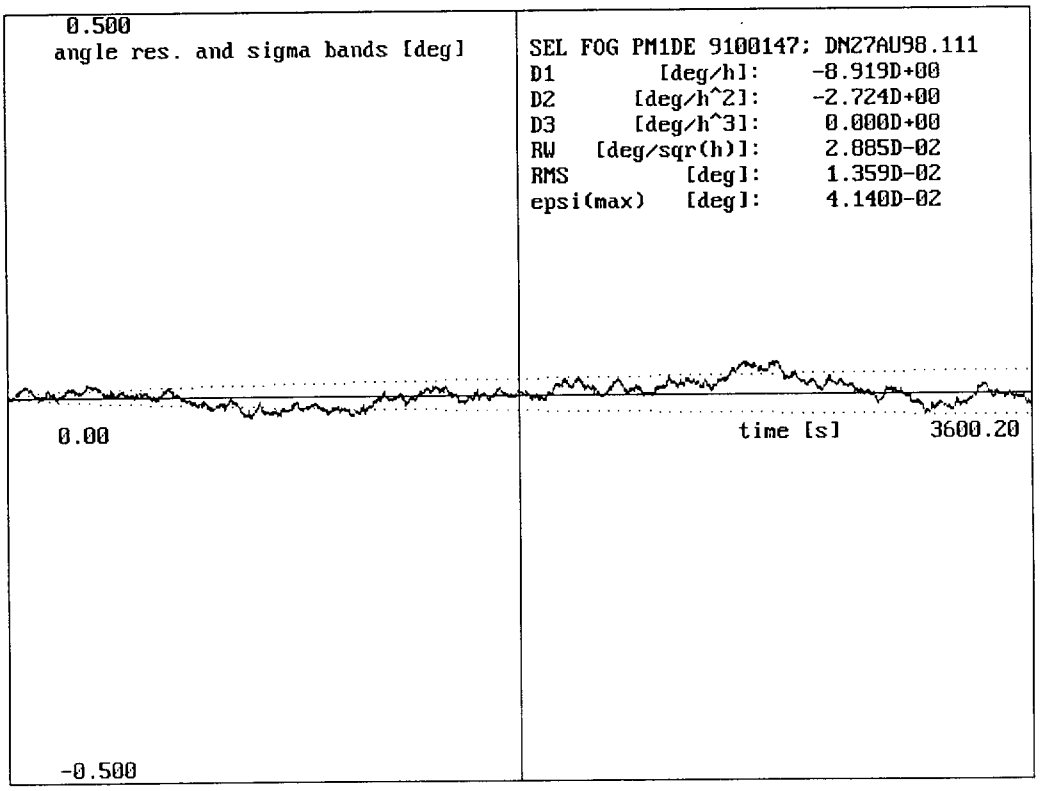
It is of the form:

$$\underline{\dot{x}} = \underline{\underline{F}}\,\underline{x} + \underline{w}\,. \tag{2.4b}$$

Its solution is:

$$\underline{x}(t) = \underline{\underline{\Phi}}(t, t_0)\,\underline{x}(t_0) + \int_{t_0}^{t} \underline{\underline{\Phi}}(t, \tau)\,\underline{w}(\tau)\,d\tau\,, \tag{2.5}$$

where $\Phi(t, t_0)$ is the transition matrix.



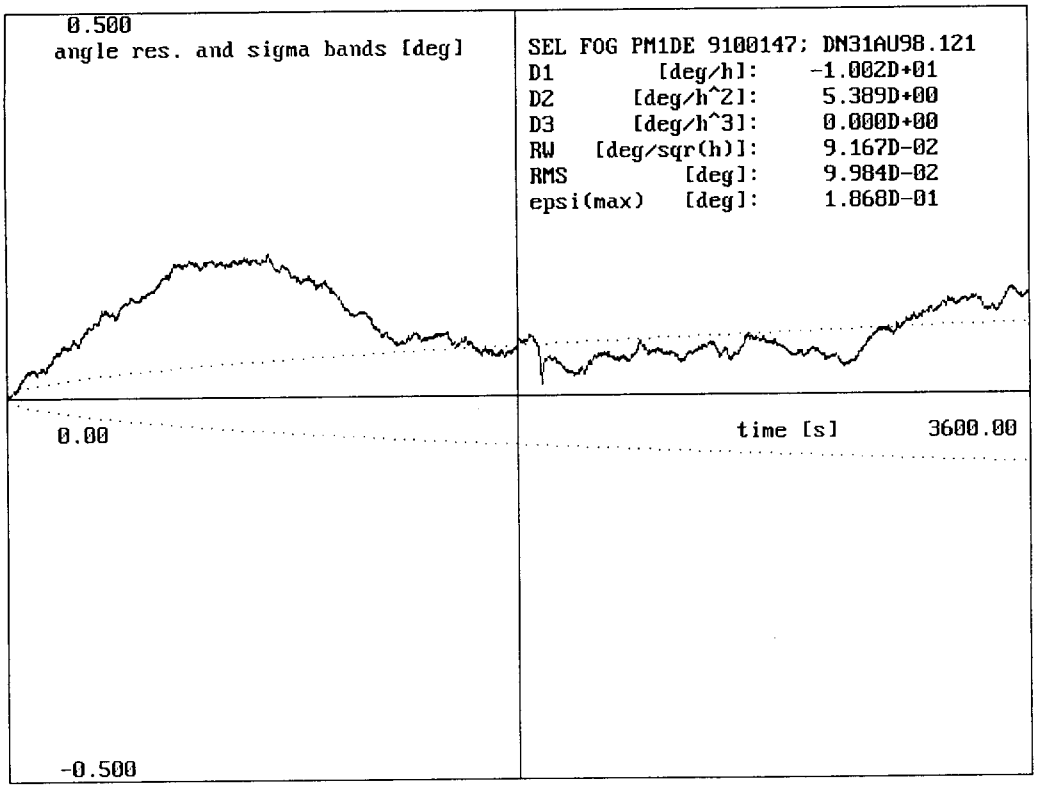


Figure 2.3 FOG Angle Residues With Respect to a Drift and Driftgradient Estimate and the ± 1σ Bands Due to a Random Walk Estimate (2 Runs of Different Quality Selected)

The estimation process is based upon the system error prediction at the measurement time t_k which again is based upon the optimal estimate $\hat{x}(t_j)$ at the previous measurement time:

$$\underline{x}'(t_k) = \underline{\underline{\Phi}}(t_k\, t_j)\, \hat{\underline{x}}(t_j). \tag{2.6}$$

Due to its non-predictability the effect of noise which has zero mean is left out. The transition matrix is in the present gyro example with $\Delta t = t_k - t_j$ and $\Delta\alpha^* = \alpha_k^* - \alpha_j^*$:

$$\underline{\underline{\Phi}}(t_k, t_j) = \underline{\underline{I}} + \int_{t_j}^{t_k} \underline{\underline{F}}\, dt$$

$$= \begin{pmatrix} 1 & \Delta t & \Delta t^2/2 & \Delta\alpha^* \\ 0 & 1 & \Delta t & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.7}$$

In this matrix the measured angle increment $\Delta\alpha^*$ is used instead of the unknown error-free value $\Delta\alpha$; the difference between both is negligible in the estimation process.

The uncertainty of the system error prediction is done in the conventional way:

$$\underline{\underline{P}}(t_k) = \underline{\underline{\Phi}}(t_k, t_i)\, \hat{\underline{\underline{P}}}(t_i)\, \underline{\underline{\Phi}}^T(t_k, t_i) + \underline{\underline{Q}}. \tag{2.8}$$

The first term on the right hand side is initialized by elements on the main diagonal of $\hat{\underline{\underline{P}}}(o)$ which describe the uncertainty in our knowledge of the initial system state; we choose:

$$\begin{aligned} \sigma_\varepsilon &= 10 \text{ arc min} \\ \sigma_\kappa &= 1000 \text{ ppm} \\ \sigma_{D1} &= 15 \text{ deg/h} \\ \sigma_{D2} &= 4 \text{ deg/h}^2 . \end{aligned} \tag{2.9a}$$

The second term on the right hand side of Eq. 2.8 describes the effect of system noise; for small Δt's it contains the following elements on the main diagonal:

$$Q_{ii} \approx r_{ii}^2\, \Delta t. \tag{2.9a}$$

The random walk coefficients are chosen as:

$$\begin{aligned} r_\varepsilon &= 0.15 \text{ deg}/\sqrt{h} \\ r_\kappa &= 500 \text{ ppm}/\sqrt{h} \\ r_{D1} &= 2 \text{ deg}/\sqrt{h} \\ r_{D2} &= 4 \text{ deg}/h^2/\sqrt{h} . \end{aligned} \tag{2.9b}$$

3. The Update Measurements and their Error Models

3.1 Updates with Position References and Zero Velocity Updating (ZUPT)

We will use the angle measurement of the quadrangle to be seen in Fig. 2.1 for the derivation of the model for angle measurement aiding.

The sides of the quadrangle are marked with the integers 0 to 4, which are used in the computer as measurement index m. All angles are related to the side with the integer 0. The integer 4 means a rotation by 360°. The signals to be registered in the computer are:

- the measurement time (t_k),
- the integrated gyro angles $\alpha^*(t_k)$ at
- the measurement position m = 0, 1, 2, 3, 4 and
- the start and end of a motion marked with m = 9.

With known references for a system state its error estimation is done in the conventional way. When the gyro is at rest in the positions m = 0 (reference $\alpha_m = 0$) and m = 4 (reference $\alpha_m = 360°$) the difference between the integrated gyro measurements $\alpha^*(t_k)$ and the reference α_m is the second information for the estimation process. The update measurement vector is thus:

$$\underline{y}(t_k) = \left(\alpha_m^*(t_k) - \alpha_m\right) + v(t_k) = \left(\varepsilon_m(t_k)\right) + v(t_k), \tag{3.1}$$

with the measurement error vector $\underline{v}(t_k) = \delta\underline{y}(t_k)$.

The update measurement is related to the system state by:

$$\underline{y}(t_k) = \underline{\underline{H}}\ \underline{x}(t_k) + \underline{v}(t_k), \tag{3.2}$$

with the measurement matrix:

$$\underline{\underline{H}} = (1 \quad 0 \quad 0 \quad 0). \tag{3.3}$$

From the update measurement noise $\underline{v}(t_k)$ the measurement noise covariance matrix in the positions m = 0 and m = 4 is derived as:

$$\underline{\underline{R}} = \left(\sigma_v^2(t_k)\right). \tag{3.4}$$

In theory $\underline{v}(t_k)$ should be uncorrelated noise. In the present case of measuring angles in the positions m = 0 and m = 4 with the gyro at rest we put the angular measurement noise equal to the manufacturing uncertainty of the quadrangle surface with $\sigma_v(t_k) = \sigma_v = 1$ arc min.

At the positions m = 1, 2, 3 of the quadrangle the reference angles α_m are assumed as unknown, but zero velocity updating (ZUPT) can be exploited, if the gyro is at rest during the measurement. For ZUPT we switch the measurement matrix in Eq. 3.2. to a velocity measurement:

$$\underline{\underline{H}} = (0 \quad 1 \quad 0 \quad 0) \tag{3.5}$$

and use

$$\underline{y}(t_k) = \left(\Delta\alpha^*(t_k) + v(t_k)\right) / \Delta t \tag{3.6}$$

as measurement with the covariance matrix:

$$\underline{\underline{R}} = (\sigma_v / \Delta t)^2. \tag{3.7}$$

3.2 Updates with Differential Inertial Measurement Techniques (DIMT)

DIMT extracts aiding information for time-dependent sensor errors from recurring identical measurement conditions. In the present example of measuring angles of a quadrangle such conditions are met if the gyro is at rest in an unknown position α_m with m = 1,2,3 or if the gyro is rotated again into this position. Aiding information is extracted from the difference of the measured angles $\alpha_m^*(t_i)$ - $\alpha_m^*(t_k)$ at the different points in time t_i and t_k. This is illustrated in Fig. 3.1 where the angles α_m to be measured are plotted in the upper diagram, the time dependent gyro error ε_t in the middle and the angles $\alpha_m^*(t_i)$ and $\alpha_m^*(t_k)$ for m = 1 including possible differences $\Delta\varepsilon_{ki}$ in the lower diagram.

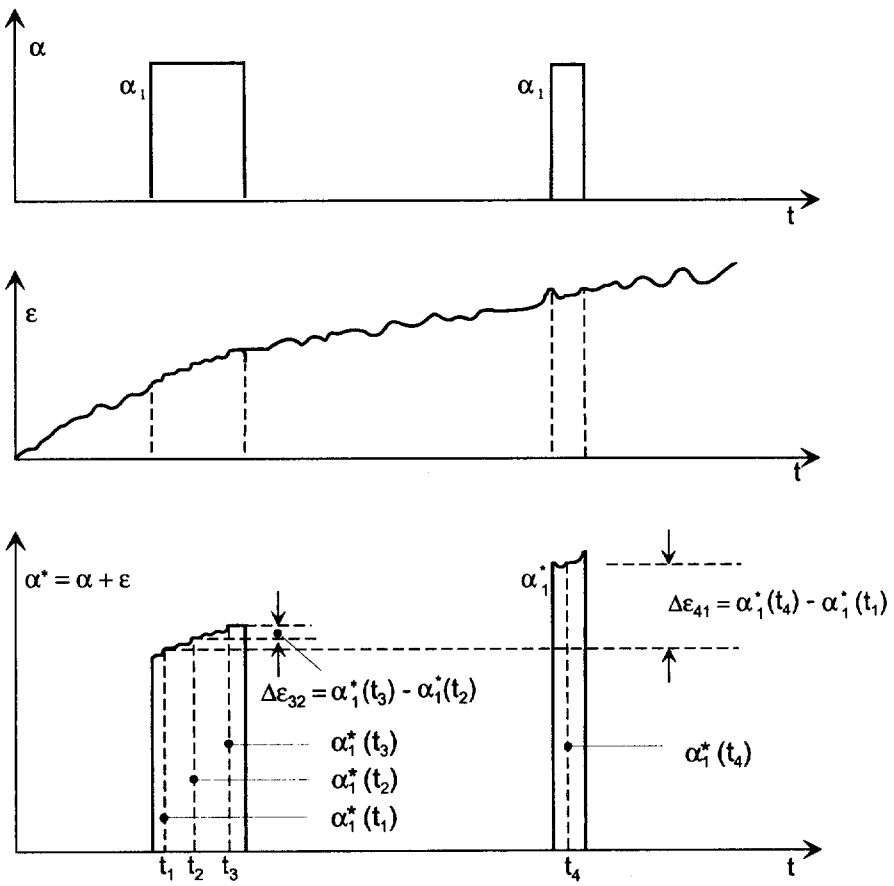


Figure 3.1 DIMT Illustrated for the Measurement of Angles

When the gyro is at rest in a certain position - m = 1 as indicated in Fig. 3.1 - DIMT processes information similar to ZUPT - the only difference being that ZUPT processes the instantaneous velocity or the position differences while DIMT can also process differences to measurements in this position m lying in the past. So also repetitive measurements at m are handled by DIMT even when the gyro is not at rest but has moved several times over the position m in the past.

Information about identical measurement conditions can be put manually as integer into the computer describing the identical but unknown position. If the computer registers a physical quantity standing in a fixed relationship to the angle to be measured - e.g. the distance travelled during the measurement of a railway track incidence or the lift measured as function of the angle of incidence of an aircraft model during windtunnel tests - information about identical measurement conditions can be derived from the measured distance or lift. These examples will be covered in Section 6.

As mentioned above, the DIMT measurement $y(t_k)$ is in the case of quadrangle angle measurement the difference of the measured angles $\alpha_m^*(t_k) - \alpha_m^*(t_i)$ in the same position α_m - as indicated through the same integers - but taken at different points in time t_k and t_i:

$$\begin{aligned} y(t_k) &= \left[\alpha_m^*(t_k) + v(t_k)\right] - \left[\alpha_m^*(t_i) + v(t_i)\right] = \left[\alpha_m + \varepsilon(t_k) + v(t_k)\right] - \left[\alpha_m + \varepsilon(t_i) + v(t_i)\right] \\ &= \Delta\varepsilon_{ki} + v(t_k) - v(t_i). \end{aligned} \tag{3.8}$$

The model for this measurement is:

$$\begin{aligned} \underline{y}(t_k) &= \underline{\underline{H}}\,\underline{x}(t_k) + \underline{v}(t_k) - \underline{\underline{H}}\,x(t_i) - v(t_k) \\ &= \underline{\underline{H}}\left[\underline{x}(t_k) - \underline{x}(t_i)\right] + \underline{v}(t_k) - \underline{v}(t_i). \end{aligned} \tag{3.9}$$

In order to relate within this model the system state at the previous time $\underline{x}(t_i)$ to the present one $\underline{x}(t_k)$ we have to extrapolate backward in time with:

$$\tau = t_k - t. \tag{3.10}$$

This leads to the differential equation:

$$\frac{d\underline{x}}{d\tau} = -\frac{d\underline{x}}{dt} = -\underline{\underline{F}}\,\underline{x} - \underline{w} = \underline{\underline{F}}^{-}\,\underline{x} - \underline{w}. \tag{3.11}$$

Its solution with $t = t_k$ or $\tau_k = 0$ taken as starting point is:

$$\underline{x}(t_i) = \underline{\underline{\Phi}}^{-}(t_i, t_k)\,\underline{x}(t_k) - \int_{t_k}^{t_i} \underline{\underline{\Phi}}^{-}(t_i, \tau)\,\underline{w}(\tau)\,d\tau, \tag{3.12}$$

with:

$$\underline{\underline{\Phi}}^{-}(t_i, t_k) = \begin{pmatrix} 1 & -\Delta t_i & \Delta t_i^2/2 & -\Delta\alpha_i^* \\ 0 & 1 & -\Delta t_i & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.13}$$

where we have used as abbreviation:

$$\Delta t_i = t_k - t_i$$
$$\Delta\alpha_i^* = \alpha_k^* - \alpha_i^* . \qquad (3.14a,b)$$

Introducing (3.12) into (3.9) the measurement model becomes:

$$y(t_k) = \underline{\underline{H}} \left\{ \left[\underline{\underline{I}} - \underline{\underline{\Phi}}^-(t_i, t_k) \right] \underline{x}(t_k) + \int_{t_k}^{t_i} \Phi^-(t_k, \tau)\, w(\tau)\, d\tau \right\} + v(t_k) - v(t_i) . \qquad (3.15)$$

We define as DIMT measurement matrix:

$$\underline{\underline{H}}^- = \underline{\underline{H}} \left[\underline{\underline{I}} - \underline{\underline{\Phi}}^-(t_i, t_k) \right]$$
$$= \begin{pmatrix} 0 & \Delta t_i & -\Delta t_i^2/2 & \Delta\alpha_i^* \end{pmatrix}. \qquad (3.16)$$

As distinguished from the direct measurement of $\underline{x}(t_k)$ using the model (3.2) with the uncertainty in (3.4), the DIMT measurement uncertainty is composed of

- the measurement uncertainties $\underline{v}(t_k)$ and $\underline{v}(t_i)$ and
- the uncertainty due to the system noise in "predicting" the state $\underline{x}(t_i)$ from the present state $\underline{x}(t_k)$.

These uncertainties are computed from the measurement error model:

$$\delta y(t_k) = \underline{\underline{H}} \int_{t_k}^{t_i} \underline{\underline{\Phi}}^-(t_i, \tau)\, \underline{w}(\tau)\, d\tau + \underline{v}(t_k) - \underline{v}(t_i) , \qquad (3.18)$$

which results in the following measurement covariance matrix:

$$\underline{\underline{R}}^-(t_k) = \underline{\underline{R}}(t_k) + \underline{\underline{R}}(t_i) + \underline{\underline{H}} \left[\int_{t_k}^{t_i} \underline{\underline{\Phi}}^-(t_k, \tau)\, \underline{\underline{q}}(t_k)\, \underline{\underline{\Phi}}^{-T}(t_k, \tau)\, d\tau \right] \underline{\underline{H}}^T . \qquad (3.19a)$$

For the quadrangle angle measurements the DIMT update measurement noise covariance matrix is:

$$\underline{\underline{R}}^-(t_k) = \left\{ 2\sigma_v^2 + r_\varepsilon^2\, \Delta t_i + r_{D1}^2\, \Delta t_i^3/3 + r_{D2}^2\, \Delta t_i^5/20 + r_\kappa^2\, \Delta\alpha^2\, \Delta t_i \right\}. \qquad (3.19b)$$

The measurement conditions described with ZUPT or with ZUPT-A in the following indicate that the sensor is at rest and instead of velocity measurements position differences were used as input into the DIMT algorithm. With ZUPT the difference to the last available position measurement in position m is taken which is equivalent to the conventional ZUPT described in Section 3.1. With ZUPT-A differences to all available position measurements in position m are taken. This has to be taken into account in the DIMT update algorithms in the following way:

When arriving at t_k the state prediction $\underline{x}'(t_k)$ and its covariance $\underline{\underline{P}}'(t_k)$ are done in the conventional way for the first DIMT update. For additional updates at t_k based upon additional DIMT measurements out of the past the prediction and its covariance consist of those of the last optimal update in t_k.

It was found out empirically that the measurement covariance should be raised in case of ZUPT-A. σ_v = 30 arc min was found adequate for ZUPT-A and DIMT updates during the way back. It was also found out empirically that for taking into account sensor error model defects (s. Section 2) the parameters in Eq. 3.19b can be raised with respect to the ones for the Kalman filter (s. Eq. 2.9b). For the evaluation discussed in the following this was not done, i.e. the figures of Eq. 2.9b were used.

4. Test Results for Gyroscopic Goniometry under Static Measurement Conditions - Measuring the Angles of a Quadrangle as Example

For evaluating the measurements Kalman filter and smoothing algorithms [Ge 74] will be applied. Remember, the former relies at time t_k on update measurements taken up to t_k and can be used for on-line estimation whilst the latter is used for off-line estimation and can thus rely on measurements taken prior and later with respect to t_k. This has an impact upon the measurement sequence to be chosen for obtaining best estimation results in the example mentioned above.

The gyro scalefactor error variation from run to run was observed to be up to 2500 ppm (s. Table 4.1). So the actual value for κ should be estimated during each run especially when big angles are to be measured. When Kalman filtering is used for estimation, κ has to be estimated first, i.e. each run has to begin with the sequence m = 0,4 which means a rotation by 360°. The time t_4 for this rotation should be short, because the Kalman filter estimation uncertainty in m = 4 depends upon the initial uncertainties in scalefactor error and drift according to: $(\sigma_{\kappa 0}\ 360°)^2 + (\sigma_{D1,0}\ t_4)^2$. When smoothing is applied to the measurements it is sufficient to include the position m = 4 into the rotation sequence. An obvious sequence is m = 0, 1, 2, 3, 4 or m = 0, 1, 2, 3, 4, 3, 2, 1, 0.

Gyro drift estimation requires the evaluation of measurements taken at different points in time. The following test and evaluation procedures are possible:

a) keeping the gyro at rest in the known positions m = 0 and m = 4 for a certain while with $\alpha_0 = 0$ and $\alpha_4 = 360°$ taken as external angular position references;

b) taking at least one measurement at the known positions m = 0 or m = 4 at different points in time;

c) keeping the gyro at rest in the unknown positions m = 1, 2 or 3 and making ZUPT or DIMT updates or

d) passing with the gyro through the positions m = 1, 2 or 3 at different points in time and making DIMT updates.

The results for angle residues, scalefactor error and drift listed inTable 4.1 indicate the following:

- the angle measurement uncertainty:
  - is below 0.02 deg,
  - seems not to depend on measurement time,
  - does not depend on the use of ZUPT or DIMT (not shown here) due to the short total length of the runs between angular position updates and the subsequent data smoothing during the evaluation;
- the FOG scalefactor error variation:
  - is below 1000 ppm during one switch-on,
  - is below 2000 ppm from switch-on to switch-on;
- the FOG drift:
  - is below 2 deg/h during one switch-on,

- is below 11 deg/h from switch-on to switch-on,
- the modelling of a drift gradient is not necessary - due to the relatively short total length of the runs.

no.	rotation sequence m	length of time at m [s]	total length of runs [s]	mean and sigma of angle residues [deg] ε_1	ε_2	ε_3	ε_4	total	scalefactor error [ppm]	drift [deg/h]
1	0 to 4	5	50	0.011 ±0.017	0.009 ±0.017	0.003 ±0.012	0 0	0.006 ±0.013	73 ± 395	- 6.9 ± 1.4
2	0 to 4	15	120	0.009 ±0.015	0.001 ±0.021	0.004 ±0.020	0 0	0.004 ±0.016	-1599 ± 758	-11.3 ± 1.6
3	0 to 4	3	110	0.001 ±0.018	0 ±0.022	0.019 ±0.019	0 0	0.005 ±0.017	-1685 ± 233	-22.3 ± 1.7

Table 4.1 Test Results from 9 Runs for the Quadrangle Angle Measurement in Fig 2.1 (Angle Residues, Scalefactor Error and Drift)

5. Test Results for Gyroscopic Goniometry under Kinematic Conditions with Quantitative Angular Position References - Machine Tools and Robots as Examples

Inertial technology is generally used for measuring the states of systems in motion. If external information for position, velocity or direction are available, aided inertial technology delivers information of high quality. This section deals with the application of inertial technology to kinematic goniometry which can supplement conventional sensors [St 98].

In machine tools or robots, the angular position of a moving part often has to be measured with respect to a fixed one. Conventional sensors are angle encoders which consist of circular optic or inductive sensing elements mounted concentrically to the axis of rotation. They deliver the angular position between both parts.

Gyros mounted for this purpose to the moving part offer the following advantages:

- they measure the angular rate of the moving part with respect to inertial space and not with respect to the non-moving part; thus no accurate concentric mounting is required - the gyro can be mounted to any point of the moving part - which alleviates fabrication and servicing,
- their angular rate output signal is of advantage for the angular position control loop.

Their disadvantage - the angle- and time-dependent measurement errors - can be overcome if the non-moving part can be regarded as remaining in a known position with respect to inertial space and carries angular position references.

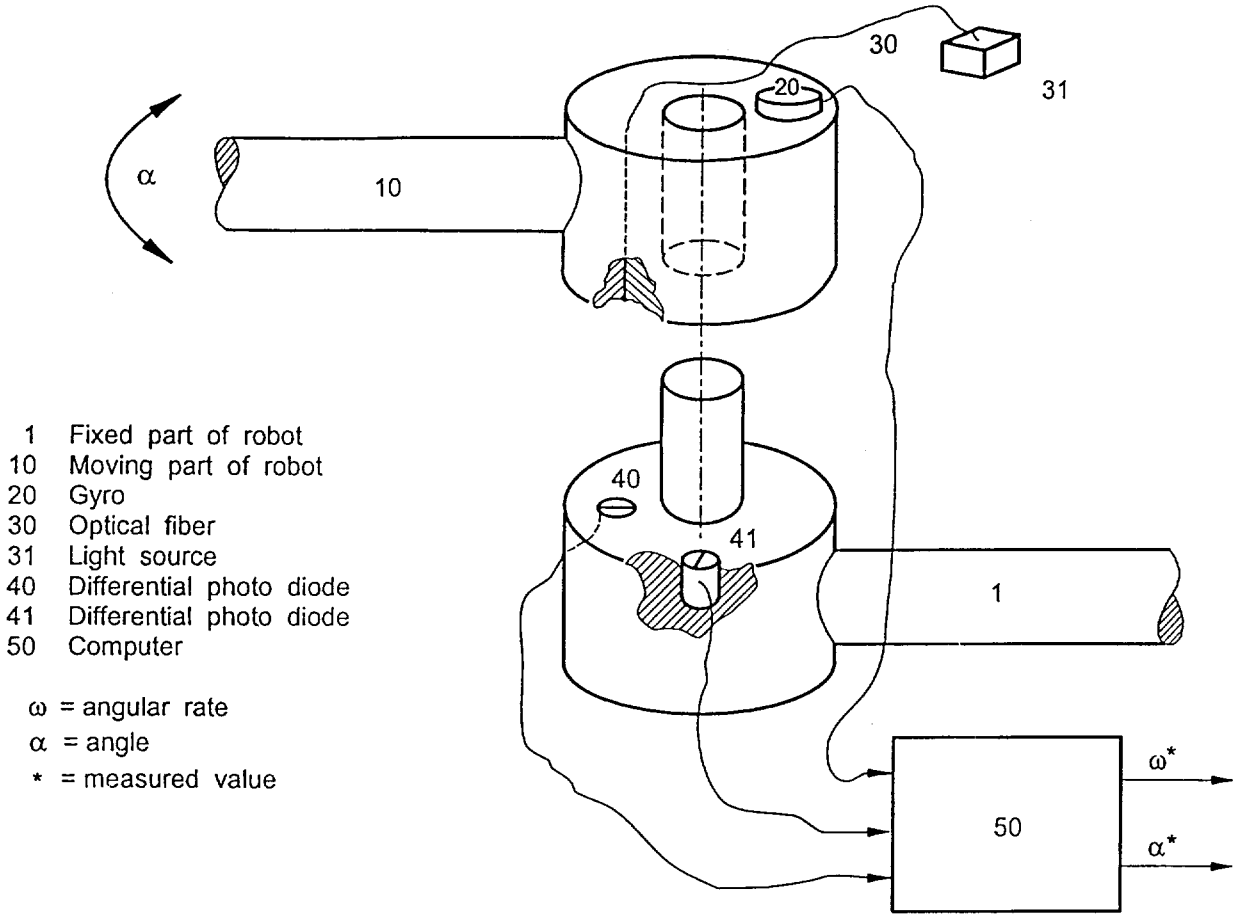


Figure 5.1 Setup for a Gyroscopic Angle Encoder in a Robot

In machine tools and robots which carry out motions in a repeating sequence the angular position references can be realized in very simple ways as pointed out in Fig. 5.1.

A light emitting diode (LED) is mounted on one of the two parts and one or two receiving diodes on the other. The relative angular position of the emitting and receiving diodes should be within the working range of both parts.

One receiving diode in a predefined angular position to zero allows the gyro initialization and the continuous drift calibration when the LED moves across it. A second receiving diode in a known angular position with respect to the first one allows in addition the continuous scalefactor error calibration.

The accurate positioning of the diodes with respect to a reference direction is either part of the manufacturing process or of a calibration process afterwards.

It is obvious that Kalman filtering has to be implemented within the computer for fusing the measurement data of the gyro and the receiving diodes.

In order to come up with a judgement on the achievable accuracy, we use again the quadrangle in Fig. 2.1 for trials. Kinematics are simulated by taking only one measurement in each position, as if the gyro would move across it. Measurements in the position m = 0 and 4 are taken for position

reference aiding and in the positions m = 1,2 and 3 for goniometric accuracy computations whose uncertainties are governed primarily by:

$$\sigma_\varepsilon^2 = \sigma_v^2 + r^2 t, \tag{5.1}$$

with σ_v = 1 arc min = uncertainty of the aiding information (s. Eq. 3.4) and r = 0.15 deg/$\sqrt{h}$ = gyro random walk coefficient (s. Eq. 2.9) and t = time with respect to the last position update in m = o or 4.

Fig. 5.2 shows estimation errors for one run which prove the relationship in (5.1). For the fairly long period of the manual motion of T ≈ 20s and the fairly high update uncertainty of 1 arc min the angle measurement uncertainty is limited to σ_ε = 0.02 deg, only.

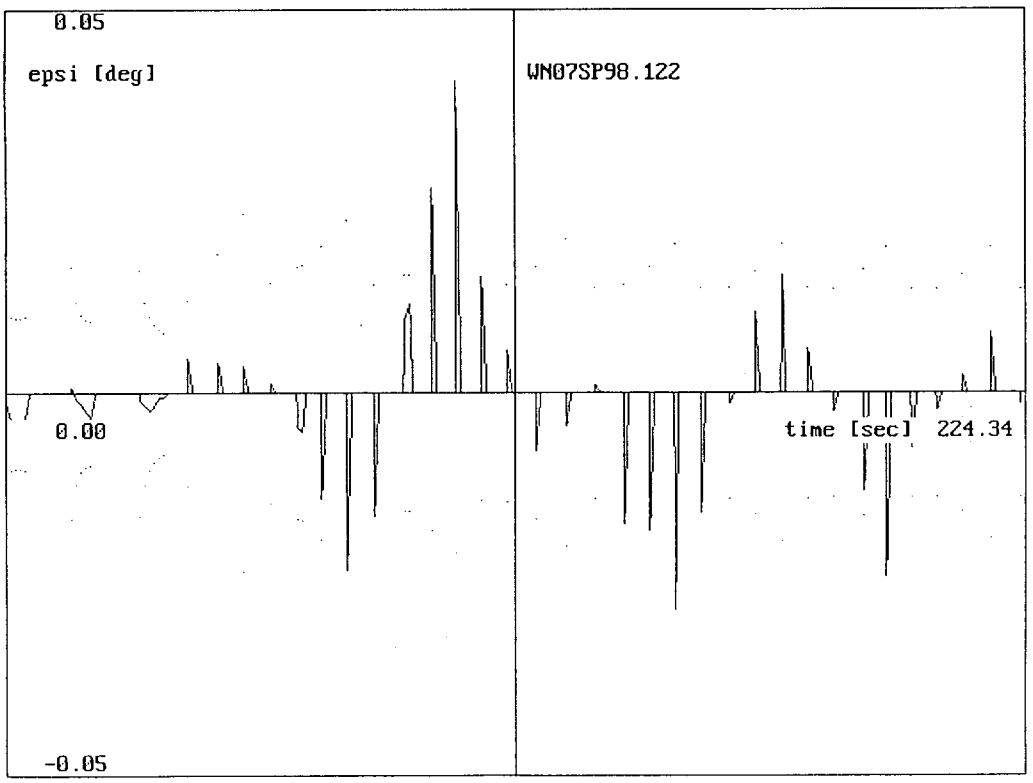


Figure 5.2 Test Results for Gyroscopic Goniometry During Repetitive Motions as in a Robot

6. Gyroscopic Goniometry under Kinematic Conditions Using Qualitative Angular Position References

6.1 Robots as Example.

If the position references m = 0 and 4 mentioned in Section 5 cannot be used for updating the robot during each cycle, but if it moves over recordable but unknown positions with high repeatability again and again - e.g. a gripping position - this information can be used for gyro aiding.

The quadrangle in Fig. 2.1 is again used for basic tests of this concept. At the beginning the gyro is precalibrated by rotating it into the positions m = 0,4, 0 and taking measurements for 5s. It is then rotated over the unknown positions m = 1, 2, 3, 2, 1 again and again, when only 1 measurement is taken for simulating kinematics.

The integers for m are registered in the computer, so that DIMT can be applied.

Fig. 6.1 shows the test results; in the upper half for conventional Kalman filtering, i.e. after the calibration no further aiding is applied and in the lower half DIMT is exploited. The results are convincing for the application of DIMT.

6.2 Track Angle Measurements as Example

Actual track angle measurements carried out for the St. Petersburg Metro are described in Chapter 14 of this book. A very simple sensor system concept with the complex DIMT data evaluation will be discussed here which can only be regarded as supplemental measuring concept to that described in Chapter 14.

According to the track angle components to be measured 1, 2 or 3 gyros are mounted onto a measurement carriage. This is also equipped with an odometer in order to locate the measured angles to the track location.

During a test campaign the measuring carriage moves in general from a starting point over the track to be measured and back again. So we obtain ideal information for applying DIMT to the recorded angle measurements, since at all track positions sensed with the odometer the track angles to be measured should remain stable. DIMT can then correlate all track positions which are identical on the way there and back with the corresponding gyro measurements. The results are of interest after the mission so that smoothing can be applied for data fusion and compensation of the long-term FOG errors. It should be remembered that the short term FOG measurements are in general of excellent quality.

Stationary FOG runs over 1 h are underlaid to the following update concepts in order to obtain a judgement on the achievable long term track direction measurement accuracy.

All update concepts rely on an equal position update time dtkal at the beginning and the end of the mission in the starting point. During the mission the measurement carriage moves across the track and returns back to the starting point when the time Tu has been reached. Tu is assumed to be 2/3 of the total mission time Te. Dtkal was chosen to be either 1 min or 2 sec. The latter case thus included no sensor error calibration but a position initialization only.

During some of the evaluated update concepts it is also assumed that alternatively or in addition to DIMT ZUPTs or ZUPT-As are taken when the measurement carriage stops 3 times for 1 min on its way to the point of return.

The FOG scalefactor error estimate is left out in this discussion; it is assumed that it is calibrated before the mission and that the angles α to be measured are so small that for the remaining uncertainty in κ the corresponding angle uncertainty $\varepsilon_\alpha = \kappa\alpha$ is negligible.

Summarily the measurement concepts to which the stationary FOG measurements were underlaid are as follows:

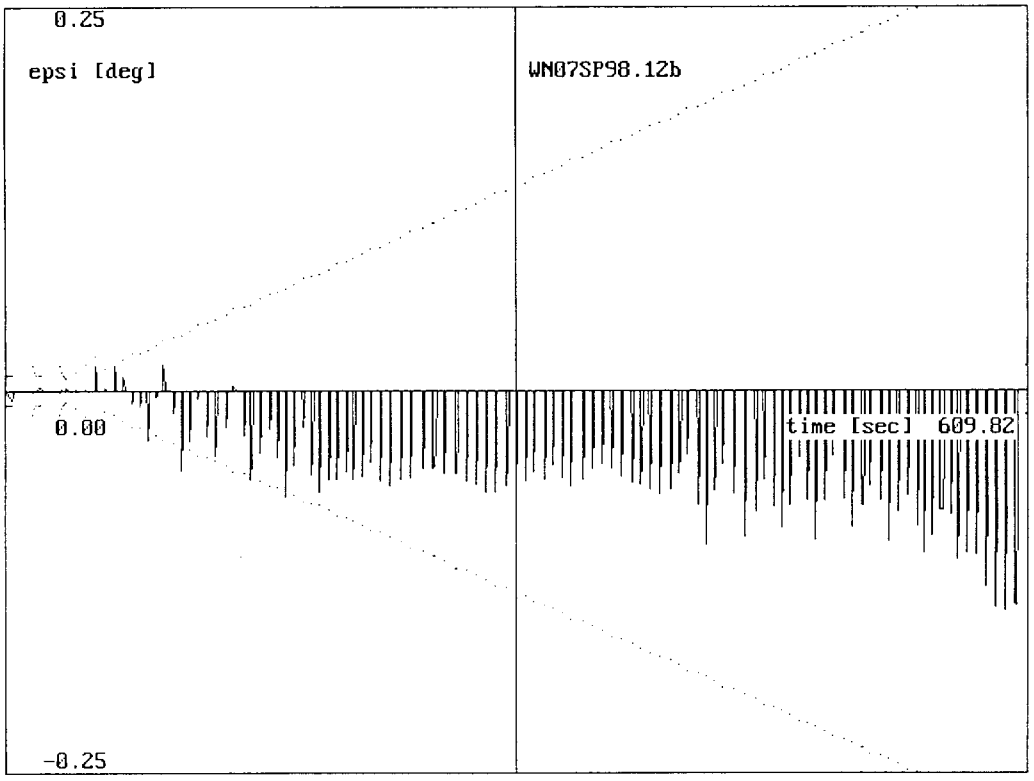
1. no DIMT & ZUPTs: only 1 min or 2 sec position reference update at the beginning and the end of the mission;

2. ZUPT: ZUPT updates for 1 min at Tu/4, Tu/2 and 3/4Tu in position m by taking the difference to the last position measurement;

3. ZUPT-A: same as ZUPT, but taking differences to all available position measurements in m out of the past;

4. DIMT: DIMT updates at all positions during the way back;

5. DIMT & ZUPT-As : updates 1, 3 and 4 together.



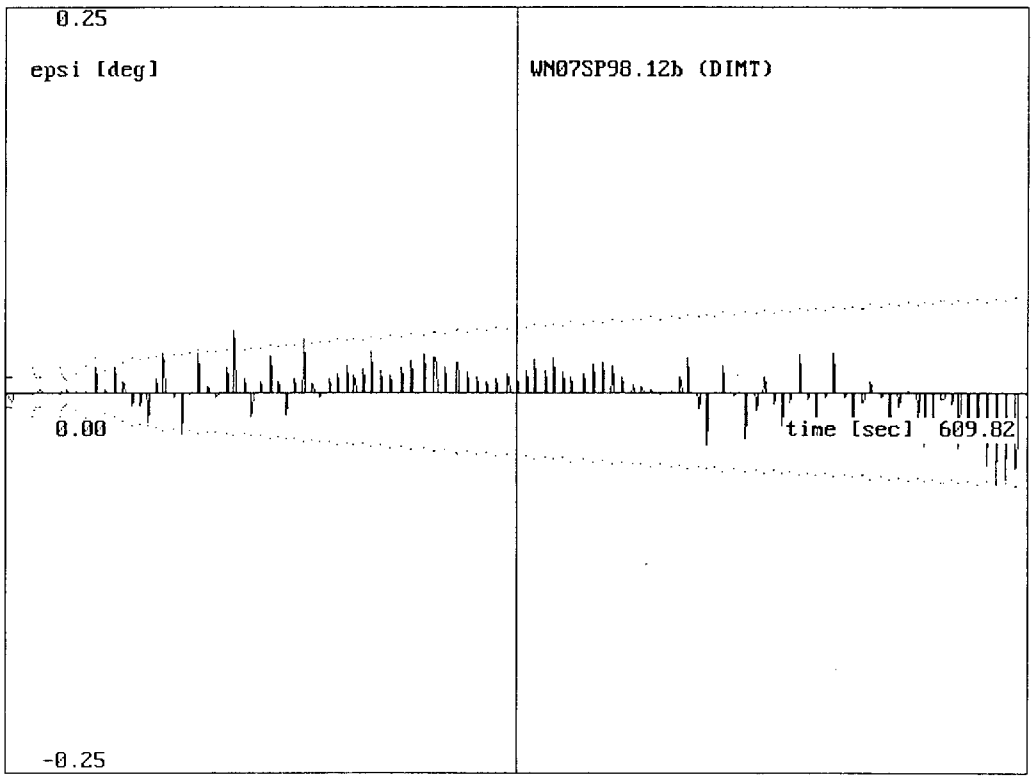


Figure 6.1 Test Results for Real-Time Gyroscopic Goniometry as in a Robot - Conventional Kalman Filtering compared with DIMT

The results are listed in Table 6.1 as RMS values of the angle residues RMS [$\varepsilon(t)$] over time for all the 10 runs and for the peaks RMS [ε_{max}] of all the residues for the whole mission. The mission time was chosen to be 1/2 and 1 h.

			dtkal = 1 min		dtkal = 2 sec	
no	mission time [s]	update	RMS [$\varepsilon(t)$] [deg]	RMS [ε_{max}] [deg]	RMS [$\varepsilon(t)$] [deg]	RMS [ε_{max}] [deg]
1	1800	no DIMT & ZUPTs	0.051	0.099	0.084	0.147
2	1800	3 ZUPTs	0.051	0.099	0.080	0.142
3	1800	3 ZUPT-As	0.050	0.096	0.075	0.133
4	1800	DIMT	0.045	0.084	0.059	0.098
5	1800	DIMT & 3 ZUPT-As	0.043	0.080	0.057	0.098
6	3600	No DIMT & ZUPTs	0.102	0.180	0.238	0.379
7	3600	3 ZUPTs	0.103	0.180	0.214	0.342
8		3 ZUPT-As	0.103	0.179	0.202	0.328
9	3600	DIMT	0.101	0.172	0.168	0.260
10	3600	DIMT & 3 ZUPT-As	0.101	0.173	0.162	0.250

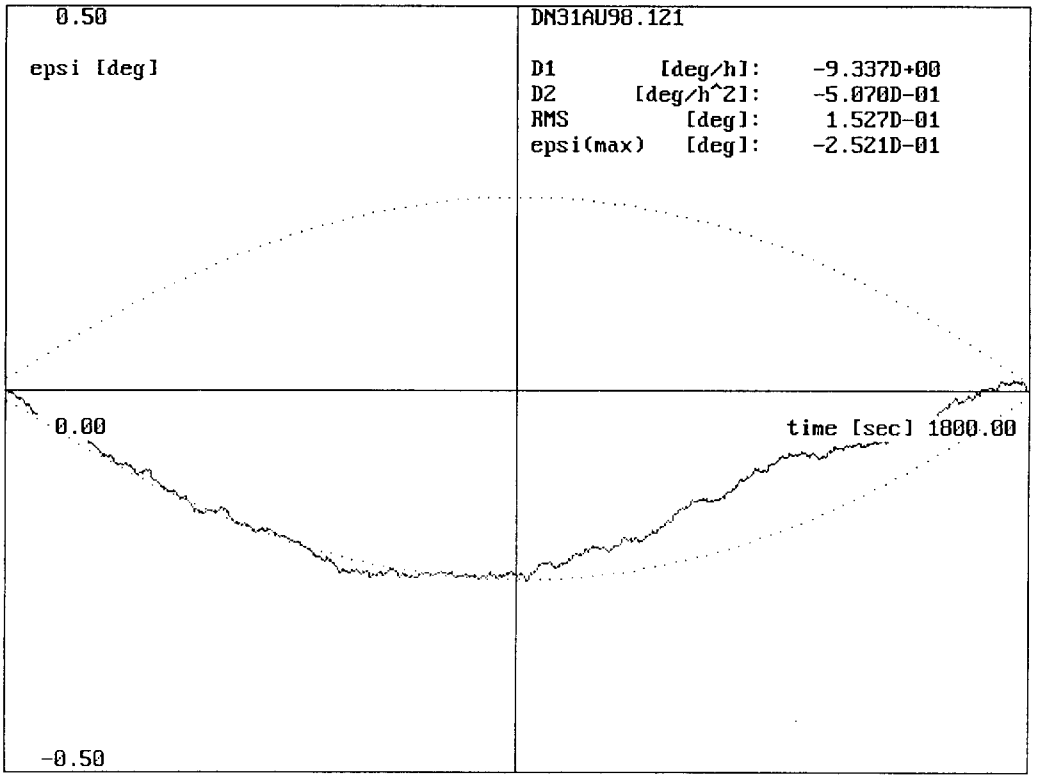
Table 6.1 Accuracy Results for Track Angle Measurement Investigations with Different Times for Sensor Calibration and Update Concepts Obtained with 10 FOG Measurements over 1/2 and 1h

Out of the 200 evaluations Figs. 6.2 a ÷ d show the angle residues for one FOG measurement and 4 update concepts.

For judging the results in Table 6.1 and Fig. 6.2 we use those of the regression analysis in Table 2.2 and Fig. 2.3 as reference. The latter were derived from a regression analysis over the whole mission time.

For dtkal = 1 min:

- It is surprising that the 1 min angular position updates before and after the mission including data smoothing after the mission result in angle deviations which are roughly only twice as big as those of the regression analysis;
- halving the mission time roughly halves the angle deviations;
- due to the relatively good sensor calibration ZUPT and DIMT bring an accuracy improvement of up to 10 % only.



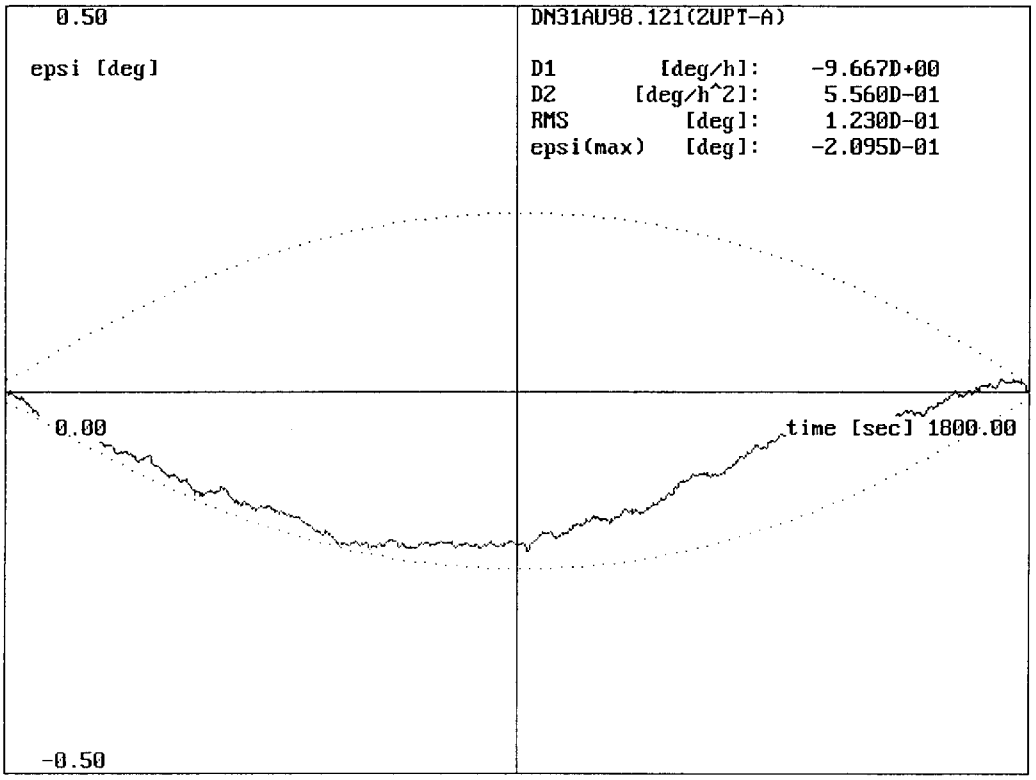
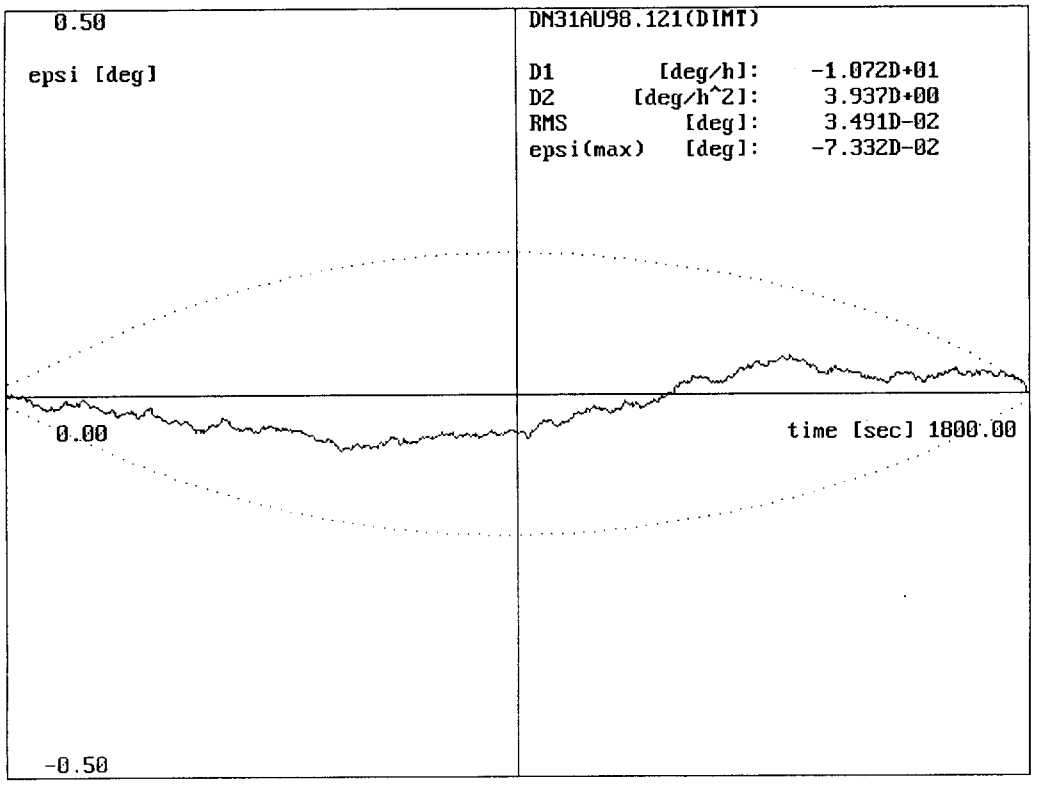


Fig. 6.2a, b Residues for Track Angle Measurement Investigations;
a) with 1 Angular Position Update at the Beginning and the End of the Mission
b) in Addition 3 × 1 min ZUPTs at 300s, 600s and 900s
(FOG Measurements s. Fig. 2.3 bottom)



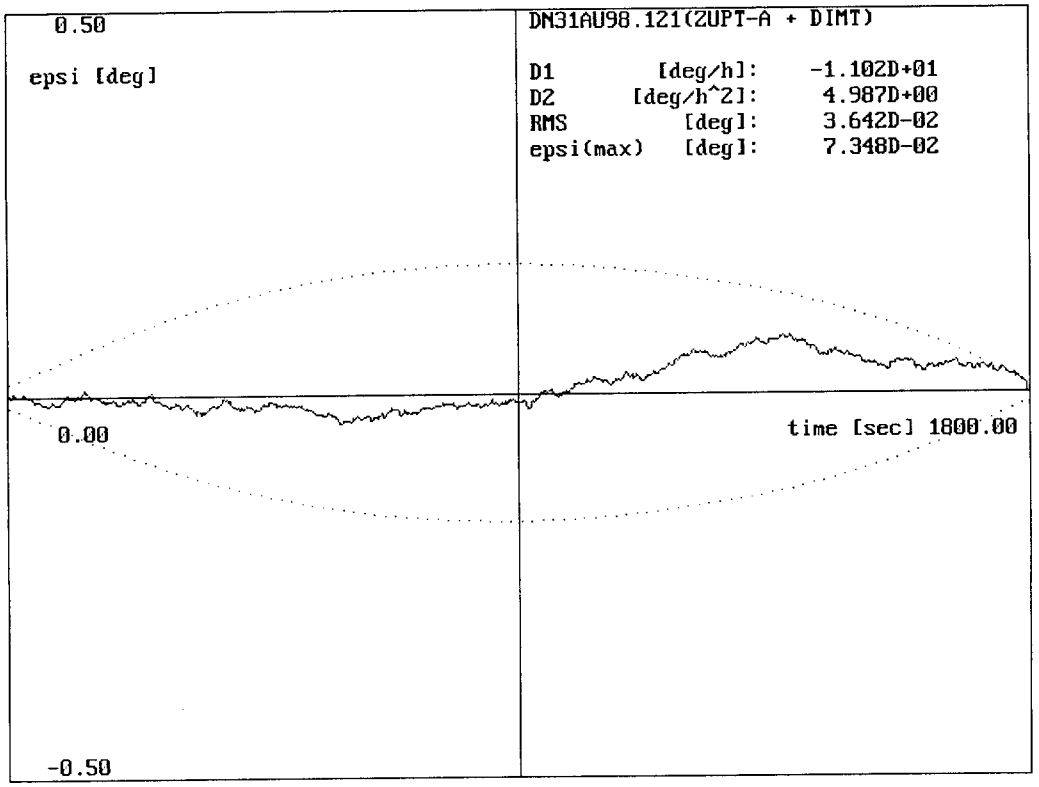


Fig. 6.2c, d Residues for Track Angle Measurement Investigations;
c) with 1 Angular Position Update at the Beginning and the End of the Mission plus DIMT During Way Back (t > 1200 s)
d) in Addition ZUPTs like in b
(FOG Measurements s. Fig. 2.3 bottom)

For dtkal = 2 s (1 position update only):

- Without ZUPT or DIMT the measurement uncertainty increases by 50% to 100%, approximately with respect to dtkal = 1 min;
- ZUPT brings an improvement by up to 10%; ZUPT-A by up to 15%;
- DIMT brings an improvement by up to 33%; when ZUPT-A is included the improvement is slightly higher.

The FOG measurements underlaid to these investigations were occasionally of limited quality. During the run shown in Fig. 2.3 bottom the drift changes at t = 800 s by approximately 1.7 deg/h probably due to thermal effects. This change cannot also be traced back to the results in Fig. 6.2 which are based upon the same gyro measurements. DIMT shows a remarkable improvement in this example.

6.3 Windtunnel Measurements as Example

Fig. 6.3 gives an insight into the German-Dutch Windtunnel (DNW) where an Airbus model is mounted to a so-called sting. It will be exposed to the airflow for measuring lift, drag and moments as a function of the angles of incidence, sideslip etc.

Figure 6.3 View into the German-Dutch Windtunnel with an Airbus Model

A typical windtunnel test result is the so-called angle of incidence polar, i.e. the dependence of the lift coefficient c_l (lift $L = c_L A \rho/2 V^2$, A = reference area of the wings, ρ = air density, V = air flow) as function of the angle of incidence α as shown in Fig. 6.4. The hysteresis at high values for α is due to flow separation.

At DNW the model attitude angles about the horizontal axes, i.e. the angles with respect to gravity - the angle of incidence, also called angle of attack and the roll angle - are measured with the "Model Attitude Measurement System (MAMS) consisting of 3 accelerometers. These inclinometers work accurately under quasi-static conditions. Under dynamic conditions, e.g. in the range of flow separation when the aircraft model/support structure is excited to oscillations, errors are introduced due to centrifugal horizontal accelerations. Since small perturbations are always present during windtunnel measurements, the inclinometer and aerodynamic measurements at DNW were lowpass filtered (3 Hz corner frequency) and averaged over 2s.

Another drawback of using inclinometers for windtunnel measurements lies in the fact that they cannot be used for goniometry about the vertical axis of the windtunnel, i.e. for conventional sideslip measurements. These limitations of using accelerometers for goniometry stimulates windtunnel users to look for supplemental means.

During a test campaign in the year 1995 the use of a FOG plus DIMT were investigated and compared with the MAMS measurements. The preparatory efforts and the results are documented in [Str 95].

When the true angle of attack is varied over time in the positive and negative directions as indicated by the dotted line in Fig. 6.5, the lift coefficient c_L and the measured angle of attack α^* will change as indicated in this figure. The stable relationship between c_L and the true angle α can be exploited for DIMT. It should be kept in mind that DIMT preconditions do not only offer themselves after the model has been moved across the whole α-range, but also when the model is kept at rest or oscillates slightly about a certain angular position (s. Section 6.1).

Gyroscopic goniometry always relies on an initialization. In the previous case a possible choice could be $\alpha = 0$ for $c_L = 0$ or the MAMS readings at $c_L = 0$ as indicated in Figs. 6.4 and 6.5. In these positions conventional Kalman filtering with angular position update can be used.

The results of the test campaign [Str 95] indicate that under the quasistatic measurement conditions including averaging and filtering the difference between both goniometric measurement and evaluation procedures - FOG plus DIMT and MAMS - remains below $\pm$ 0.02 deg. This changes drastically when the averaging and filtering process is interrupted.

The results in Fig. 6.6 were generated using a data acquisition method with one reading per second, no averaging and no filtering. This file contained the whole dynamics of wind tunnel measurements. The difference between FOG and MAMS in this figure has an amplitude of $\pm$ 1 deg (peak to peak) and more.

In order to separate the MAMS and FOG signals, a linear increasing angle of attack over time was assumed for the first upward segment of polar 80. The deviation of the two sensor signals with respect to this linear approximation is shown in Fig. 6.7. Major differences between the two sensors are obvious. The gyro measurement results have a performance of one order of magnitude better than MAMS due to the sensor's insensitivity to vibrations.

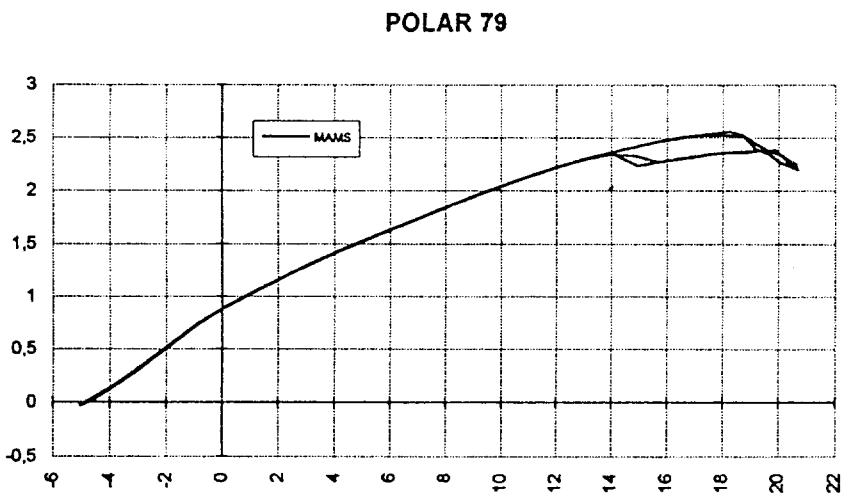


Figure 6.4 Lift Coefficient versus the Angle of Incidence as Windtunnel Test Result

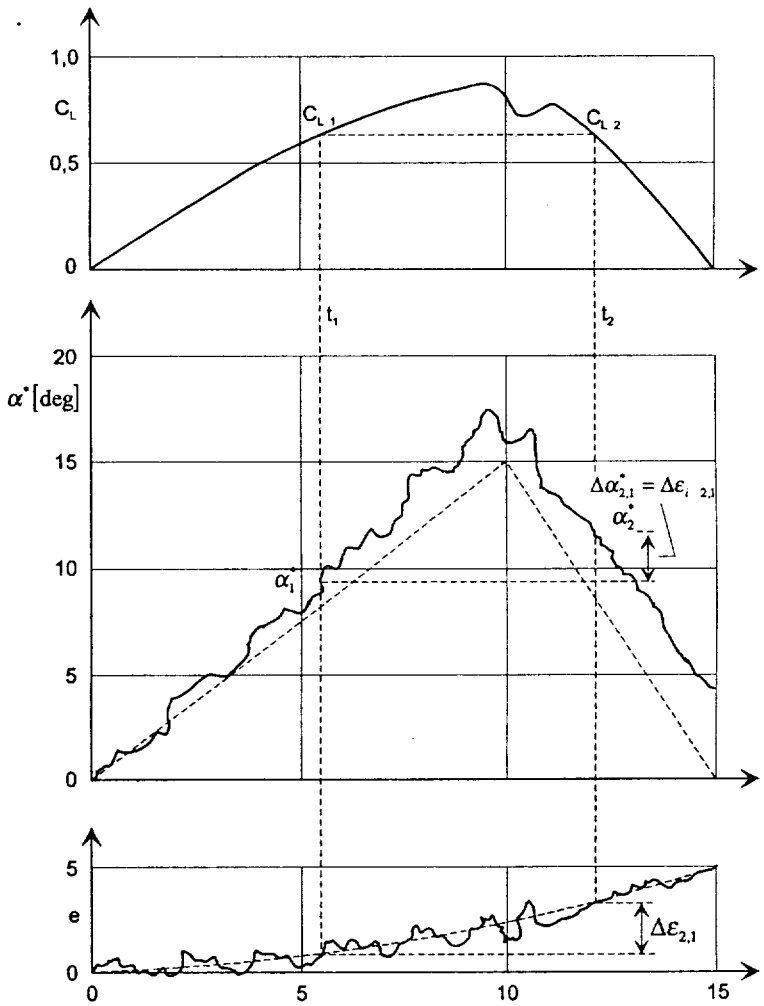


Figure 6.5 Sketch of Measured Angle of Incidence and Lift Coefficient over Time

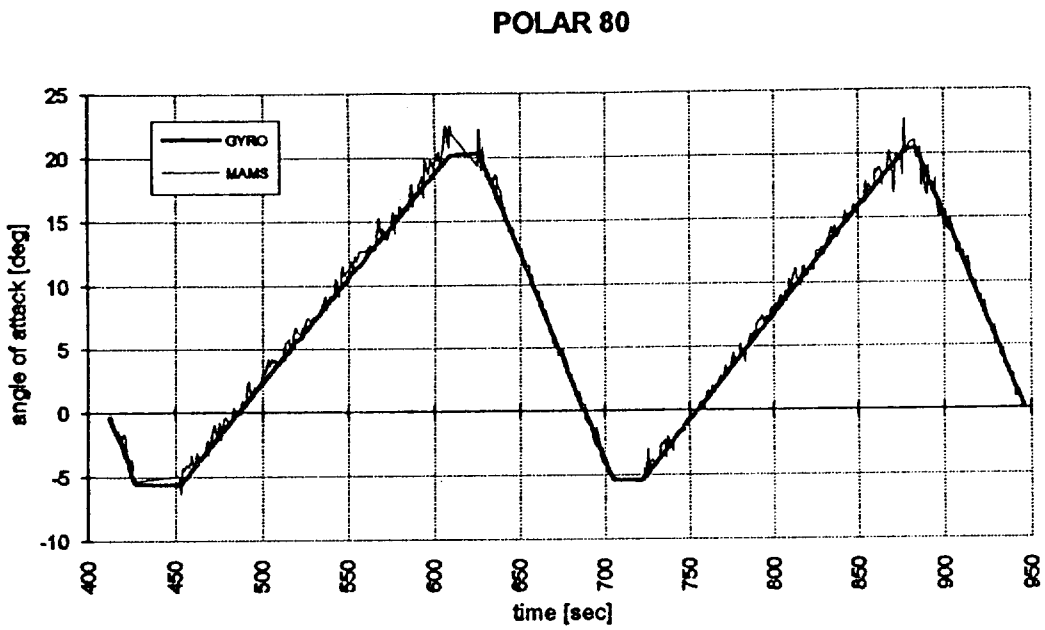


Figure 6.6 Angle of Incidence versus Time as measured by the FOG plus DIMT and by MAMS

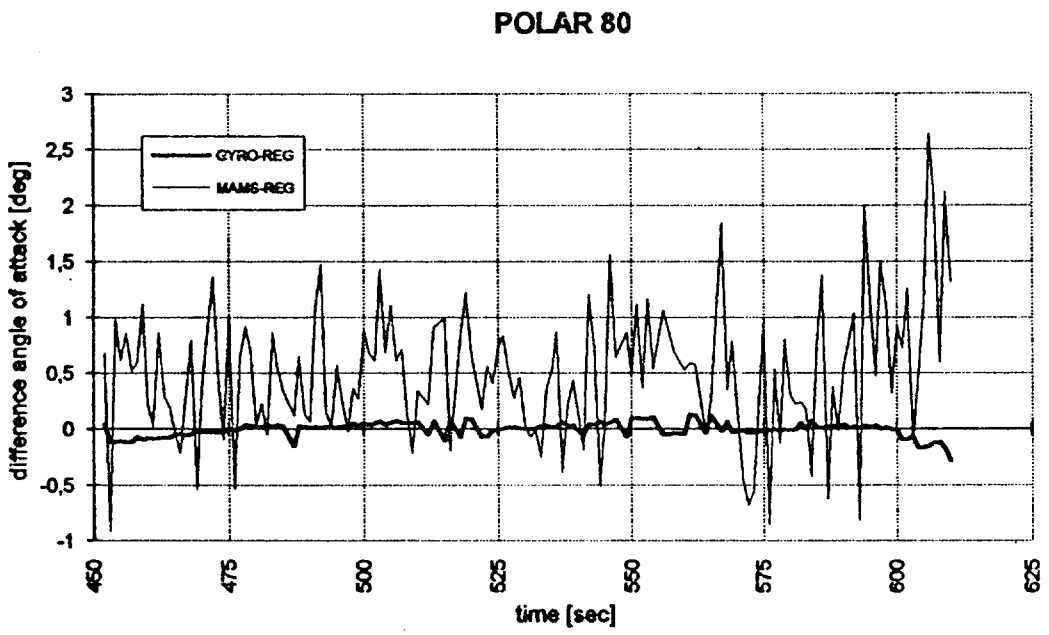


Figure 6.7 DIMT-Corrected FOG Data and MAMS Data compared to a Linear Curve over Time

7. Summary

Gyroscopic goniometry is discussed for different kinds of applications:

- measurement of the surface angles of a body under quasistatic conditions,
- measurement of the angles between hinged parts in motion as in machine tools and robots,
- measurement of the angles of a railway track and
- measurement of angular characteristics such as the lift of a windtunnel model as function of the angle of incidence.

Special attention is paid to different kinds of update concepts. The conventional methods are the angular position update and the zero velocity update (ZUPT). A new method is presented as Differential Inertial Measurement Technique (DIMT) which can be regarded as an extension to ZUPT, but without the constraint of the sensor or system being at rest. It exploits the difference of inertial measurements under identical measurement conditions taken at different points in time. Since DIMT relies on inertial measurement differences, the update information need not be based upon absolute external reference but only upon qualitative information that the measurement conditions at both points in time are identical. This identity can be based upon an integer for an angle with unknown direction but measured twice in time, upon an integer for identical gripping positions of a robot, upon the odometer readings on the way there and back of a measurement carriage during railway track incidence or direction measurements, or upon the measured lift of an aircraft model in a windtunnel. Examples thereof are investigated. All applications and update methods are exemplified with a fiberoptic gyro (FOG) of medium quality.

Abbrevations, Symbols

d	time dependent drift
D_1	drift bias
D_2	drift gradient
dtkal	calibration time prior and after the mission
DIMT	Differential Inertial Measurement Technique
$\underline{\underline{F}}$	system matrix
$\underline{\underline{H}}$	measurement matrix
m	subscript for characterizing same measurement condition, e.g. same position
$\underline{\underline{P}}$	covariance matrix
$\underline{\underline{Q}}$	system noise matrix
r	random walk coefficient
$\underline{\underline{R}}$	measurement covariance matrix
t	time
$\underline{v}$	measurement noise vector
w	system noise vector

$\underline{x}$ state vector

$\underline{y}$ measurement vector

ZUPT zero velocity update; the sensor is at rest in position m and its instantaneous velocity or its position differences are taken as measurement

ZUPT-A like ZUPT, but differences to all past position measurements in position m are taken

Greek Symbols

α angle to be measured

ε error angle

κ scalefactor error

$\underline{\underline{\phi}}$ transition matrix

Acknowledgement

The author wants to express his gratitude to his wife Ingrid for revising the English language, to Mrs. Renate Buehrig for typing the text, to Dipl.-Ing. Karlheinz Hurraß for reading through the text and to Mrs Kelly Edwards for her assistance in editing the text.

Literature

[Ge 74] Gelb, A. (Editor): "Applied Optimal Estimation". The M.I.T. Press, 1974.

[St 94] Stieler, B.: "Procedure for Measuring Angles and Trajectories by Means of Gyros and Inertial Systems".
US Patent No. 5, 331, 578, Date of Patent July 19, 1994

[Str 95] Stroezel, M.; Saaro, J.; Stieler B.; Eckert, D.:
"Wind Tunnel Model Attitude Measurement with a Fiberoptic Gyro."
Symposium Gyro Technology 1995, Stuttgart, Germany

[St 98] Stieler, B.: "Messanordnung zur Regelung von Robotern, Werkzeugmaschinen und dergleichen sowie damit durchgefuehrtes Messverfahren".
Deutsche Patentanmeldung, April 1999

14. APPLICATION OF A LASER GYRO IN TRACK MEASURING SYSTEMS

D.P. Loukianov, A.V. Mochalov
St.Petersburg Electrotechnical University
Professor Popov str. 5, 197376 St.Petersburg, Russia
M. Rechel.
DB Zentralbereich Basistechnologien
Pionier str. 10, 32423 Minden, Germany

14.1. DEVELOPMENT OF TRACK MEASURING SYSTEMS AND THE BASIC CHALLENGES

The need for the maintenance of safety and comfort of traffic calls for continual inspection and monitoring of the railway condition. The railway control involves detection and measuring of wheel track deviations from the standards established for the profile and curvature in plan.

14.1.1. Parameters to be measured and the measurement requirements

The main parameters to be checked are as follows: the gauge (checking with a pattern); the rail superele-vation (cross profile); curvature and the heading angle; short- and long-wave track and rail irregularities, etc. Besides the above-mentioned geometrical parameters it is also necessary to check the change in geometrical parameters under a load, rails for wear and defects, the quality of rail joints, switches, the condition of the overhead line (height and stagger of the wire with respect to the rails), etc.

Checking includes: regular visual walking inspection or travelling in a rolling-stock; control with the help of hand-operated measuring devices and tools and through the use of track measuring cars or carriages automatically recording the railway condition; control of the railway under a load by the use of specialized track measuring vehicles equipped with computers for registration and analysis of measurement results. The last method is the most complete and universal. It also permits various additional information to be recorded automatically, for example: information about measurement conditions, such as a track number, scale, date, time and a data file where the information is stored; synchronization marks to the track kilometers; events, like stations, switches, etc.

Inspection and surveying of the contact wire, its arrangement and wear (for the underground contact rail) are particularly complicated problems. The requirements for railway and underground track measurements are quite different in different countries and for different applications [1...3]. The examples of the measu-rement requirements and parameter specifications for the track recording car unit of Deusche Bahn (German Railways) are given in Tab. 14.1 and 14.2 [1].

Table 14.1. The main measurement requirements and their technical realizations

Requirements	Argument	Possible technical realizations
Precision in measuring geometrical parameters of the track	Need for accurate data about the track	Inertial measuring units
Measurement of long- wave rail irregularities	High speed traffic	Parameters - attitude, accelerations, location
High speed (250 km/h), the same accuracy at various speeds	Minimum occupation of tracks	Contactless optical sensors, typical rolling-stock
High precision, reliability, calibrations and self-test possibilities	Automatic data processing and self-tests	Reference frame with laser beams

Table 14.2. Some parameter measurement requirements and specifications

Parameter	Accuracy	Dimension
Gauge	± 0.5	mm
Subelevation of rails	± 0.8	mm
Leveling up to 25 m	± 0.8	mm
Alignment up to 25 m	± 1.0	mm
Leveling up to 50 m	± 1.5	mm
Alignment up to 50 m	± 1.5	mm
Long waves from 50 to 200 m	± 5.0	mm

Adequate measurement precision is basic to ensure that the differences between the results obtained in two successive measurement runs will be reliably recognized. As the measurement process is followed by automated on-line evaluation of the track condition, the data obtained must be highly reliable. These specifications are to

be observed up to the preset maximum speed of 250 km/h regardless of weather and for an outside temperature range from -10°C to + 60°C.

14.1.2 The current status of the research and employment of laser gyros

During the last years inertial sensors, devices and systems have found wider application in present-day automatic track measuring equipment. In the conditions of growing speeds and density of railway traffic, growing requirements to accuracy of measurements, the systems on laser gyro (LG) and accelerometers seem to be the most promising. Now the rail-ways and the underground are equipped with track measuring cars fitted with devices "of different generations" and sophistication, from elementary aircraft vertical gyros to modern marine inertial navigation systems (INS). One of the most widespread cars in Russia developed in the early sixties, is equipped with a marine gyro horizon which measures the car angular motion with respect to the horizon plane. Figure 14.1 shows the operation of the system measuring a track cross profile.

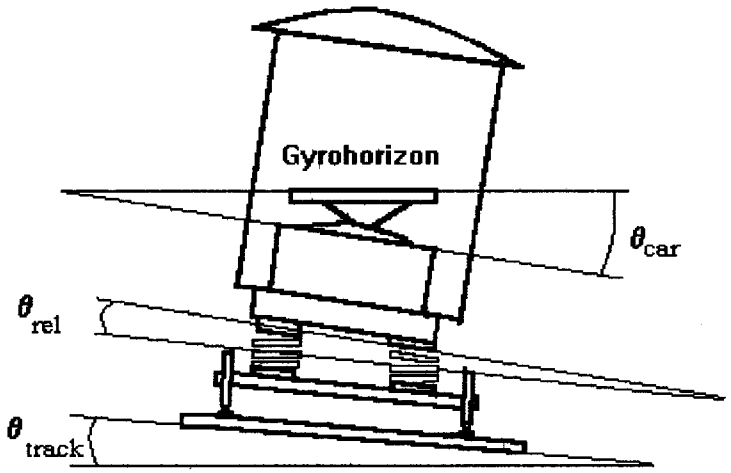


Fig. 14.1. A system measuring a track cross profile.

As the body of the measuring car is supported with springs, when moving it shakes and shows an angular misalignment θ_{rel} with respect to the track. The car body inclination angle with respect to the horizon θ_{car} is determined as

$$\theta_{car} = \theta_{track} + \theta_{rel},$$

where θ_{track} is the track inclination with respect to the horizon plane. Therefore, to measure an angle θ_{track} the angle θ_{rel} should be subtracted from the angle θ_{car} measured with the gyro horizon. In the car under consideration the subtraction is executed by the mechanical equalizer, the resultant θ_{track} changes being converted to the movement of the stylus with the use of a rather complex system of cables and blocks. The latter introduces additional errors, with the value usually exceeding the error of the inertial measuring unit.

In more perfect track recording cars [1...3] based on gyro horizons or on inertial navigation systems, inertial units are installed, as a rule, in the car. One way to increase the accuracy of measurements is to exclude the measurement of the car body motion with respect to the carriage. That can be achieved if the sensitive elements are installed directly on the axle-boxes which becomes possible with an LG and force-rebalanced accelerometers highly resistant to vibrations, shocks and climatic effects. In any case, the result of the measurements with inertial sensors will contain, besides the relative angular and linear motions, the components due to the Earth's rotation, cross-channel interferences, Coriolis accelerations, drifts of the sensors. To measure the relative motion, these components should be compensated for.

14.2. DESIGN PRINCIPLES OF TRACK MEASURING SYSTEMS ON LASER GYROS

14.2.1. The main kinematic equations

In the most general case the carriage has three LGs and three accelerometers mounted on it (Fig. 14.2).Let us introduce the following coordinate systems: Oxyz fixed to the carriage and sensors, with the axis Ox aligned with the transversal axis of the carriage, the axis Oy aligned with the longitudinal axis of the carriage and its motion direction, Oz perpendicular to the platform of the carriage; $O\xi\eta\zeta$ is the geographic frame, with the axis $O\xi$ directed to the East tangentially to the parallel, the axis $O\eta$ directed tangentially to the meridian to the North, and the axis $O\zeta$ directed upwards along the vertical.

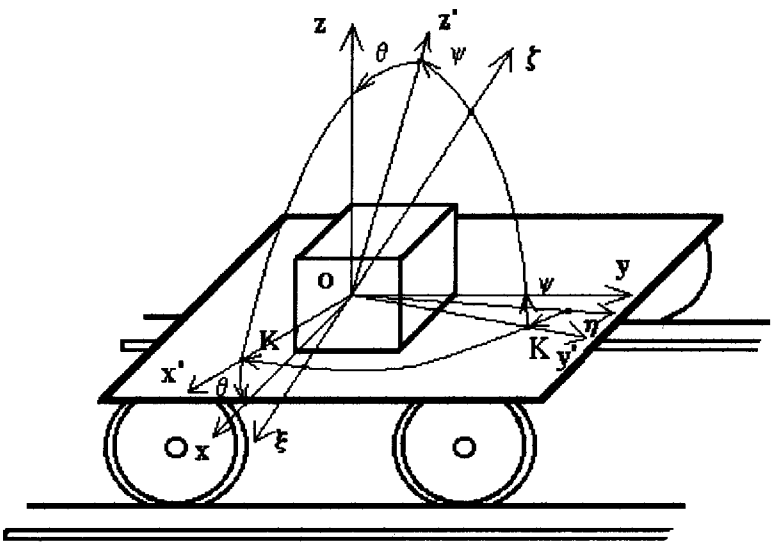


Fig. 14.2. Disposition of the sensors on the carriage

The absolute angular velocity Ω of the body-fixed frame Oxyz is determined by the expression

$$\Omega = \begin{vmatrix} \omega_x \\ \omega_y \\ \omega_z \end{vmatrix} = \omega_r + \omega_t = \begin{vmatrix} p \\ q \\ r \end{vmatrix} + B \begin{vmatrix} \omega_\xi \\ \omega_\eta \\ \omega_\varsigma \end{vmatrix}, \qquad (14.1)$$

where ω_r is the angular velocity of the carriage with respect to the geographic frame; ω_r is the reference-frame angular velocity of the geographic frame rotation due to the Earth's rotation and the carriage motion over the Earth's sphere; B is the matrix of transition from the geographic coordinates system $O\xi\eta\zeta$ to the body frame Oxyz, with

$$B = \begin{vmatrix} \cos\theta\cos K + \sin\theta\sin\psi\sin K; & -\cos\theta\sin K + \sin\theta\sin\psi\cos K; & -\sin\theta\cos\psi \\ \cos\psi\sin K; & \cos\psi\cos K; & \sin\psi \\ \sin\theta\cos K - \cos\theta\sin\psi\sin K; & -\sin\theta\sin K - \cos\theta\sin\psi\cos K; & \cos\theta\cos\psi \end{vmatrix}. \qquad (14.2)$$

Expression (14.2) includes K as the heading angle; θ - the roll angle (angle of the cross inclination); ψ - the pitch angle (angle of the longitudinal inclination) [4]. The components of the relative angular velocity are determined by the following equations [4] :

$$\begin{aligned} p &= \dot\psi\cos\theta + \dot K\sin\theta\cos\psi; \\ q &= \dot\theta - \dot K\sin\psi; \\ r &= \dot\psi\sin\theta - \dot K\cos\theta\cos\psi. \end{aligned} \qquad (14.3)$$

The components of the angular velocity of the geographic frame $O\xi\eta\zeta$ are similar to those given in [4]:

$$\omega_\xi = -\dot\psi = \frac{V_N}{R},$$

$$\omega_\eta = (\Omega_E + \dot\lambda)\cos\varphi = \Omega_E\cos\varphi + \frac{V_E}{R},$$

$$\omega_\varsigma = (\Omega_E + \dot\lambda)\sin\varphi = \Omega_E\sin\varphi + \frac{V_E}{R}\,\mathrm{tg}\varphi,$$

where φ, λ are the latitude and longitude of the carriage; Ω_E - the Earth's rotation angular velocity; V_N and V_E - the northern and eastern components of the carrier velocity. Taking into account the fact that in our case

$$\begin{vmatrix} V_\xi \\ V_\eta \\ V_\varsigma \end{vmatrix} = \begin{vmatrix} V_E \\ V_N \\ V_\varsigma \end{vmatrix} = B^{-1}\begin{vmatrix} 0 \\ V \\ 0 \end{vmatrix} = \begin{vmatrix} V\cos\psi\sin K \\ V\cos\psi\cos K \\ V\sin\psi \end{vmatrix}$$

and doing calculations in accordance with Equations (14.1) and (14.3), we derive the expressions for the input signals of LGs:

$$\omega_x = \dot{\psi}\cos\theta + \dot{K}\sin\theta\cos\psi + \Omega_E\cos\varphi(\sin\theta\cos\psi\cos K - \cos\theta\sin K) - \Omega_E\sin\varphi\sin\theta\cos\psi - \\ - \frac{V}{R}\cos\psi(\cos\theta + tg\varphi\cos\psi\sin K\sin\theta); \tag{14.4}$$

$$\omega_y = \dot{\theta} - \dot{K}\sin\psi + \Omega_E\cos\varphi\cos\psi\cos K + \Omega_E\sin\varphi\sin\psi + \frac{V}{2R}tg\varphi\sin K\sin 2\psi; \tag{14.5}$$

$$\omega_z = -\dot{K}\cos\theta\cos\psi + \Omega_E\sin\varphi\cos\theta\cos\psi - \Omega_E\cos\varphi(\sin\theta\sin K + \cos\theta\sin\psi\cos K) + \dot{\psi}\sin\theta + \frac{V}{R}\cos\psi\sin\theta + \\ + \frac{V}{R}tg\varphi\cos^2\psi\cos\theta\sin K. \tag{14.6}$$

Furthermore, it should be noted that the relation between the track inclination angle θ_h in the geographic coordinate system to the angle θ may be expressed by the equation

$$\theta = \theta_h\cos\psi$$

and the elevation of one rail over the other by

$$h = b\sin\theta,$$

where b is the track gauge.

As follows from Equations (14.4...14.6), in order for the angles K, θ, ψ to be determined, it is necessary that the system of three nonlinear differential equations should be integrated, thus, we will obtain a strapdown attitude orientation system. If we construct a simplified system with one or two measuring channels and LG correction by accelerometers, we may use these equations to estimate the errors due to incomplete compensation for the components to be excluded.

14.2.2. The classification of the systems

Integrating the system of the three nonlinear differential equations (14.4...14.6), we will obtain a strapdown attitude orientation system. If a measuring system of a maximum possible accuracy is required, we will have to use a strapdown inertial navigational system. That will permit us not only to measure the three-axis angular orientation and calculate the needed parameters but also to fix the train in reference to the mileage, as well as pinpoint irregularities in the tracks.

Another possible approach is to use these equations in constructing simplified systems with limited possibilities, containing one, two or three measuring channels with LG correction by accelerometer signals or other sources of navigation information. These systems may be simpler and cheaper than a SINS. A peculiar opportunity for the railway to use preliminary information about the railway condition and design parameters opens additional possibilities for improvement of such systems. The classification of track measuring systems is given in Fig. 14.3.

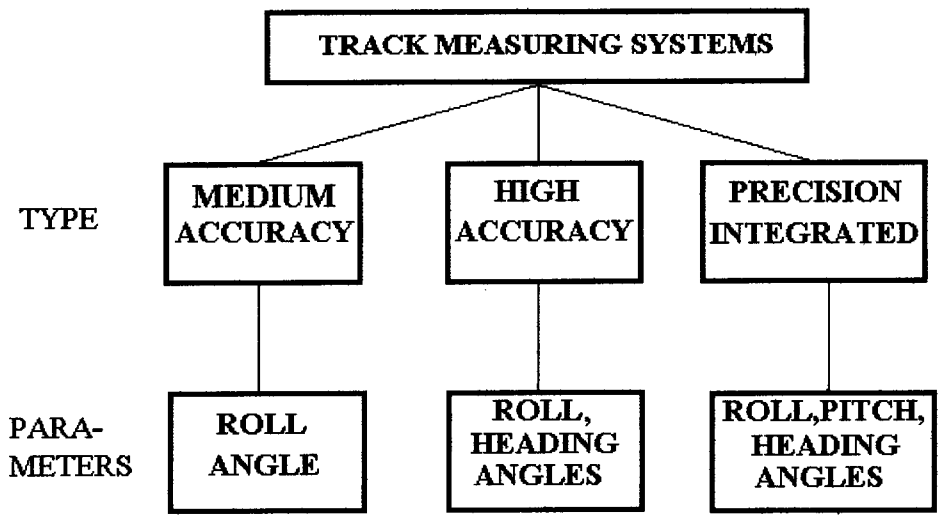


Fig. 14.3. A classification of track measuring systems

14.3. MEDIUM AND HIGH ACCURACY SYSTEMS

14.3.1. A medium accuracy system

The authors have developed a system for track cross inclination measurement and registration which is shown as a block diagram in Fig. 14.4. This system may be classified as a medium accuracy system (see Sect. 2.2).

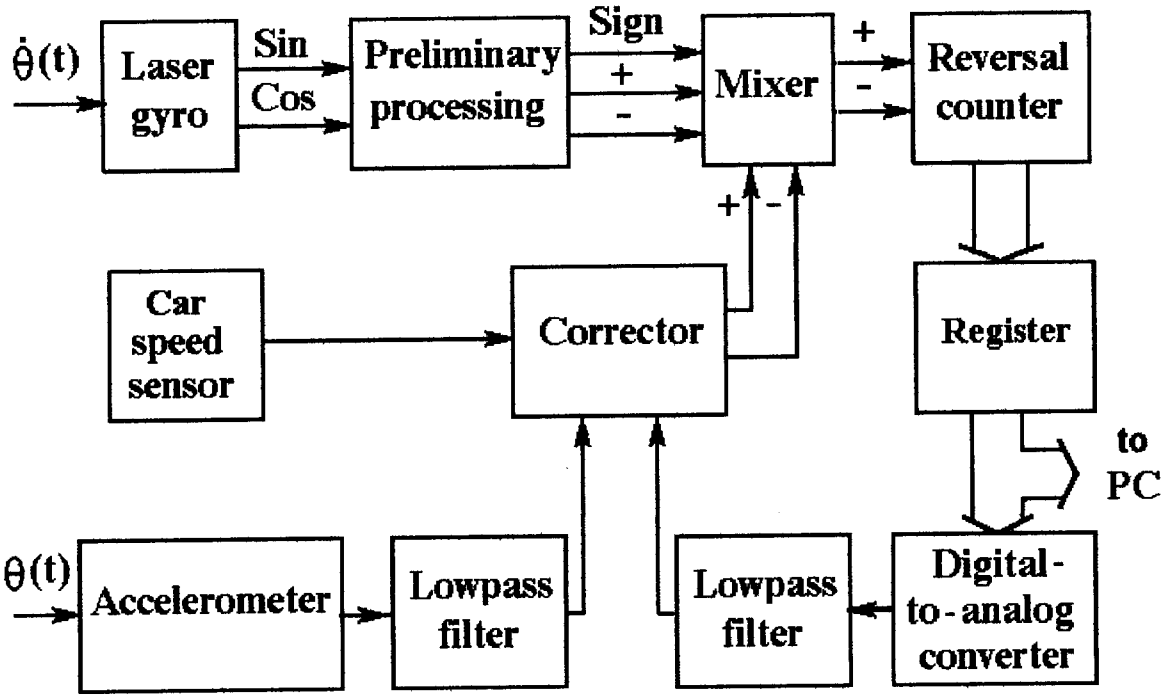


Fig. 14.4. A block diagram of a single-channel system

The laser gyro measures the absolute angular velocity of the carriage about the longitudinal axis of the car including the angular velocity with respect to the horizontal plane and the reference-frame angular velocity of the Earth's rotation. Moreover, the LG's output signal contains a drift. To eliminate the influence of the Earth's rotation, the LG drift and other components included in Equation (14.5), correction by the accelerometer is used. The accelerometer is mounted on the carriage in such a way that its measuring axis is aligned with the axis Ox (Fig. 14.2). The correction algorithm is based on the comparison of the average values of LG and accelerometer signals picked off the outputs of the lowpass filters. When the carriage changes the course and the accelerometer is disturbed by the centripetal acceleration the correction is disconnected automatically. For this purpose preliminary information about the way inclination in railway curves is used.

The system has the following technical specifications : the range of the measured inclination angles with respect to the horizontal plane ± 6 ° (±150 mm elevation of one rail over the other); accuracy of an inclination measurement with respect to the horizon ± 2.5′ (± 1 mm elevation); frequency of measurements of up to 154 Hz (every 10 cm of the track at a car velocity of 60 km / hour); bus interface 16, 24 bits; required power supply 100 W; size: the sensors unit 250 x 300 x 150 mm, the electronics unit 250 x 300 x 150 mm; weight, max. 20 kgs.

The data is read out (the indications of the reversal counter) and recorded in a PC or in a special storage unit on "flash cards" by the signals of the odometer. The prestarting procedures for the system (initial alignment) includes its start-up at a rather horizontal part of the track (elevation of one rail over the other no more than ± 20 mm) and 1 ...2 minutes' intervals until the transient for the correction unit has been completed. The preliminary analysis of the results may be made in a car, but the end-point analysis is usually made in the laboratory.

14.3.2. The experience of operation in the St.-Petersburg Metro

The laboratory and stand tests carried out on rotary tables have proved that the system meets the technical specifications and characteristics in various modes of operation. The single-channel system was put to experimental operation in a surveying car CSRI-2 in the St.Petersburg Metro (the underground). The inertial sensor unit was mounted on a special platform on the car's axle-box. Different railway segments were measured and the results compared with the readings of those of the standard system, consisting of a marine gyro horizon and a complex cable and block mechanism meant for compensation for the car angular fluctuations on

springs. Fig. 14.5 illustrates an example of track surveying carried out with an LG-based system (curve 1) and with a standard system CSRI-2 (curve 2).

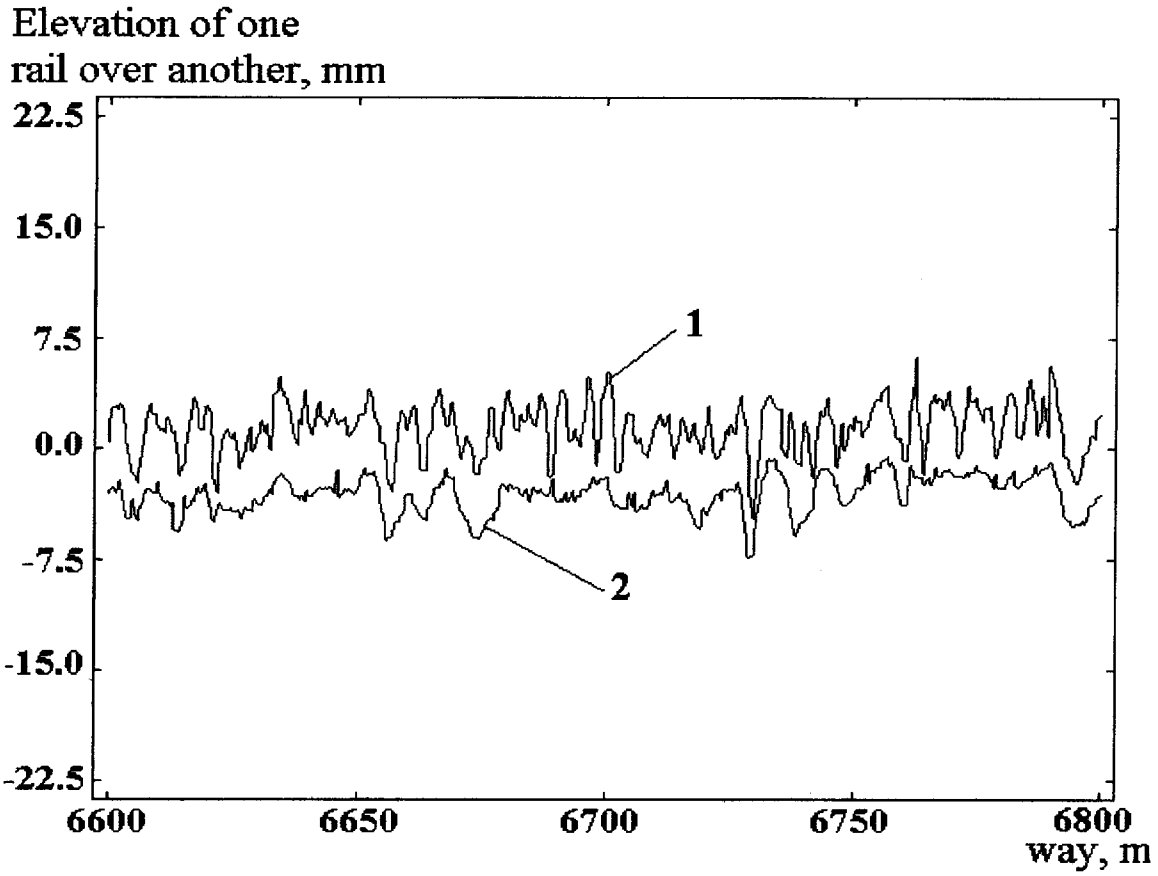


Fig. 14.5. Comparison of the curves obtained with an LG-based system (curve 1) and with a standard system CSRI-2 (curve 2)

The comparative analysis of these curves shows that the measurements obtained with the laser system give more complete information about the dynamic interaction of the railway with the loaded car. The spectral analysis shows that the output signal of the laser system contains information about the rail intersleeper deflections and rail butts. When the car was moving in a curve with a simultaneous slope, which is particularly characteristic of the under-ground, the influence of cross link was found to be rather strong. The signal of the heading angle changes is projected on the LG input axis (the second component in Equation (14.5)). To eliminate this influence an additional unit of correction is introduced. It determines an average angle of the railway cross inclination in the railway curve and maintains its value. The additional error in measurements in such correction does not exceed 1 mm. After going out of the curve, the system switches on the correction from the accelerometer again. The change of the railway over significant intervals of time were under investigation. The track recordings made two months later are given in Fig. 14.6.

Good reproducibility of the results testifies to high reliability of the measurements. The small horizontal shifts of the curves are the results of the errors in car's location measurements. All the correction procedures and reference of the results to the car location (the distance passed) can be considerably simplified by using a map of the profile and the results of the previous measurements stored in the computer.
The results of the experimental operation formed the basis for the recommendations aimed at refinement of the system. A modification of the system intended for application on high-speed railways has been developed, allowing the measurement frequency to be increased up to 2464 Hz and higher.

14.3.3. A high accuracy system

As a result of cooperation between St.Petersburg Electrotechnical University and the research center of the railway management of Germany in Minden, a modification of the system, complemented with second (heading) channel has been developed. Besides the LG and the accelerometer in the cross profile measuring channel (a cross LG and lateral accelerometer), it contains a heading LG and a longitudinal accelerometer. In this case in addition to the profile, the railway slope, the course and the heading angles are recorded to the following accuracies: inclination with respect to the horizon ±1′ (± 0,25 mm elevation of one rail over another); the railway slope with respect to the horizon ±10′; course ± 1°.

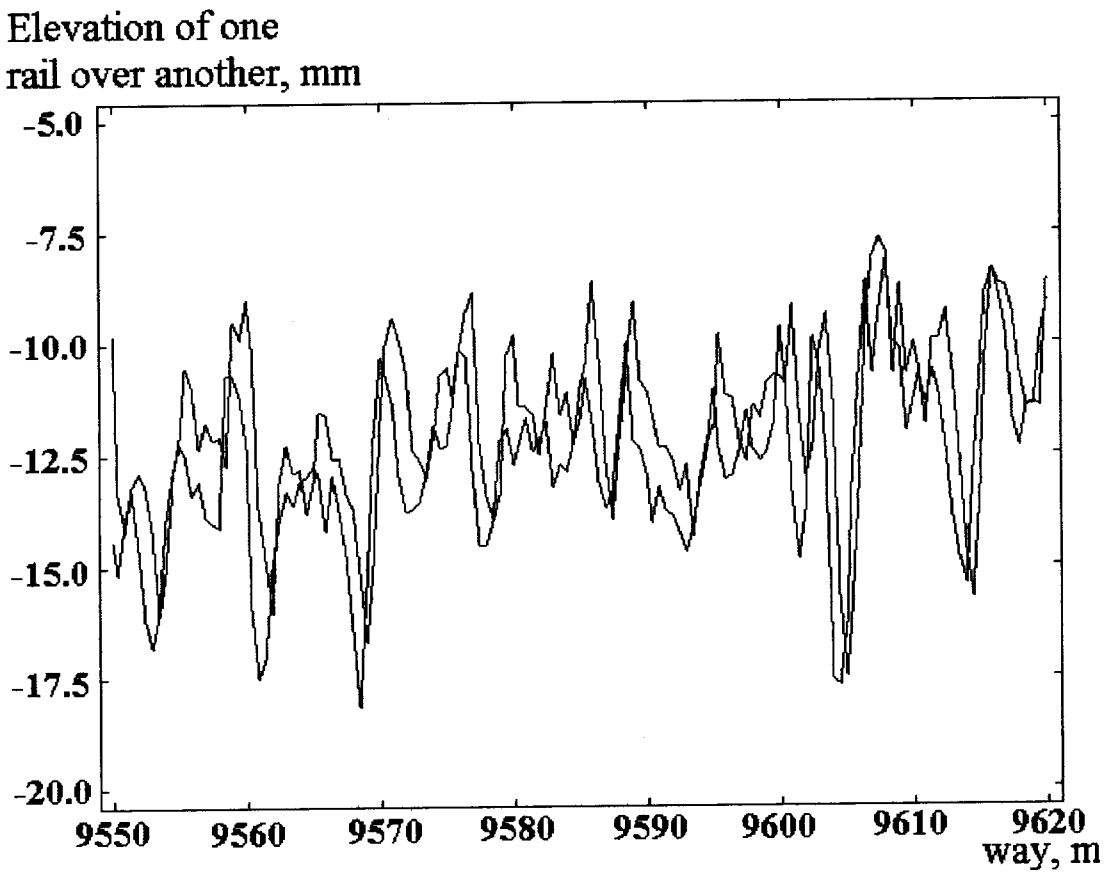


Fig. 14.6. Two realizations, made by LG system in two months

The accuracy of inclination measurement can be essentially increased in this case by the algorithm excluding the influence of the Earth's rotation and cross links according to Expressions (14.5) and (14.6) and subsequent correction of the results according to accelerometer readings.

Among the angles to be measured, the roll angle θ is the most important, as it gives the information about the elevation of one rail over the other (i.e. the cross profile of the railway). From Equation (14.5) it follows that besides $\dot{\theta}$ the output signal ω_y of the cross GL contains a number of the components, including the horizontal projection of the Earth's rotation velocity $\Omega_E \cos\varphi \cos K \cos\psi$. To exclude these components it is necessary to know the heading angle which can be derived by integrating the signal of the azimuth LG. The angle ψ of the railway slope with respect to the horizon (the pitch angle) is often also a parameter to be measured, however, in a number of problems (for example, for the underground tracks) it is necessary to know only an average value of this angle over rather long sections of the railway. In this case it is possible to use only two LGs (cross and azimuth) and two accelerometers and measure the rail-way slope with respect to the horizon with the longitudinal accelerometer. It simplifies the system and in the most cases yields a satisfactory result.

Estimations of the components in Expressions (14.5) and (14.6) are given in Table 14.3. As follows from Table 14.3, the influence of components 4, 2 and 8 is the most significant. That is why they must be calculated and excluded from the LG output signal, or else a strong correction should be introduced. Estimate, for example, the error introduced with the incomplete compensation for the horizontal component of the Earth's angular velocity (component 2) if the error in the course determination $\Delta K = 1°$. Assume that $K=K + \Delta K$. Then

$$\cos(K + \Delta K) = \cos K \cos\Delta K - \sin K \sin\Delta K \approx \cos K - \Delta K \sin K.$$

The error is maximal at $K= 90°$ and has a value of $\Omega\, \Delta K \cos\varphi \sin K= 0{,}13°$/hour. The maximum acceptable error of 0.5 mrad is accumulated in this condition over a period of $t=13.2$ min.
Component 4 may be the basic source of the error. If the track slope $\psi = 2°$ and the car changes the heading angle by $\Delta K=90°$, the accumulated error $\Delta\theta$ is more than 3°. The necessity to measure the slope angle ψ and compensate for component 4 is evident. A similar analysis may be carried out relative to the other components.

Table 14.3. The main components of the LG output signals and their estimations.

№	Component	Value of parameters	Value of component °/h
1	$\dot{\theta}$	Component to be measured	
2	$\dot{K}\sin\psi$	$\psi=2°$, $\dot{K}=1{,}25°/c$	100
3	$\Omega_E \cos\varphi\cos\psi\cos K$	$\varphi=60°$, $\psi=0°$, $K=0$	7.5
4	$\Omega_E \sin\psi\sin\varphi$	$\varphi=60°$, $\psi=2°$	0.455
5	$\frac{V}{2R_1}\mathrm{tg}\varphi\sin K\sin 2\psi$	$\varphi=60°$, $K=90°$, $\psi=2°$	$2{,}4\cdot10^{-3}$
6	$-\dot{K}\cos\theta\cos\psi$	Component to be measured	
7	$\Omega_E \sin\varphi\cos\theta\cos\psi$	$\varphi=60°$, $\theta=\psi=0$	13
8	$\Omega_E \cos\varphi\sin\theta\sin K$	$\varphi=60°$, $\theta=5°$, $K=90°$	0.65
9	$\Omega_E \cos\varphi\cos\theta\sin\psi\cos K$	$\varphi=60°$, $\theta=5°$, $\psi=2°$, $K=0$	0.26
10	$\dot{\psi}\sin\theta$	$\dot{\psi}=5°/h$, $\theta=1°$	0.09
11	$\frac{V}{R}\cos\psi\sin\theta$	V=250 km/h, $\psi=0$, $\theta=5°$	0.2
12	$\frac{V}{R}\mathrm{tg}\varphi\cos^2\psi\cos\theta\sin K$	V=250 km/h, $\varphi=60°$, $\psi=\theta=0$, $K=90°$	3.8

A full account of the above-mentioned errors may only be fulfilled after the third LG has been installed and the system of nonlinear differential equations (14.4...14.6) has been integrated.

Even in this case the two components in Equations (14.5, 14.6) $\dot{K}\sin\psi \approx \dot{K}\psi$ and $\dot{\psi}\sin\theta \approx \dot{\psi}\theta$ may be the source of an additional methodical error at conic vibrations [5]. The mechanism of this error can be explained in the following way. In the case of harmonic single-frequency angular vibrations with the relative phase shift αo

$$\psi = \psi_0 \sin\omega t;\ \theta = \theta_0 \sin(\omega t + \alpha_0).$$

We rewrite Equation (14.6) in the form of

$$\omega_z = \psi\,\theta + \Delta,$$

where Δ denotes all the other components.

Combining the two foregoing equations, we derive :

$$\omega_z = \omega\,\psi_o\theta_o\cos\omega t\,\sin(\omega t + \alpha_o) + \Delta =$$
$$= \frac{\omega\,\psi_o\theta_o}{2}[\sin(2\omega t + \alpha_o) + \sin\alpha_o] + \Delta = \Delta f + \Delta.$$

So, besides the basic component Δ, there appears an additional error Δf with the constant component

$$\Delta f_{=} = \frac{\omega\psi_0\theta_0}{2}\sin\alpha_0$$

and a variable component at a double frequency of vibrations. If the variable component is averaged after the integration, the constant component accumulates. This error reaches its maximum value at $\alpha o = \pi/2$ and grows as the frequency ω and the vibration amplitudes ψo and θo grow . For example, at ω = 440 rad/s (70 Herz), $\psi o = \theta o = \ =1{,}75\cdot10^{-2}$ rad (1°); $\Delta f = 6{,}6\cdot10^{-4}\ ^o/h$. It is easy to see that this error is negligible and should only be included if the measurements are extremely accurate.

Let us make a similar analysis of the output signals of the horizontal accelerometers. The accelerations of the car, the components of the gravitational acceleration g, the centripetal and Coriolis accelerations will be projected onto the measuring axes of the accelerometers during the car motion. The projections of the gravitational acceleration onto the car axes are defined by the following expression:

$$\begin{vmatrix} g_x \\ g_y \\ g_z \end{vmatrix} = B^{-1} \begin{vmatrix} 0 \\ 0 \\ g \end{vmatrix} = \begin{vmatrix} -g \sin\theta \cos\psi \\ g \sin\psi \\ g \cos\theta \cos\psi \end{vmatrix}$$

When the train changes the course the centripetal acceleration has the value of $\dot{K}V$ and is directed along the measuring axis of the lateral accelerometer (OX). The other components, alternating-sign, are of the zero average value and may be filtered off. The input signals of the accelerometers may be presented by the following expressions (including Coriolis accelerations [4]):

$$\left.\begin{aligned} W_x &= -g \sin\theta \cos\psi + \dot{K}V - 2V\Omega_E \cos K \sin\varphi - \frac{V^2}{2R_1} \sin 2K \, tg\varphi, \\ W_y &= g \sin\psi + \dot{V} + 2V\Omega_E \sin K \sin\varphi + \frac{V^2 \sin^2 K}{R_1} tg\varphi. \end{aligned}\right|$$

The quantitative evaluation shows that at V=250 km/h the components of Coriolis accelerations are no more than 10^{-3} g and may be neglected. The components $\dot{K}V$ and $\dot{V}$ have to be calculated and compensated for. In a two-channel system the railway slope ψ with respect to the horizon is calculated according to the readings of the longitudinal accelerometer whose output signal may be expressed by the approximate equation

$$W_y \approx g \sin\psi + \dot{V},$$

where: g - gravitational acceleration; V - the car speed measured by the sensor. Initial alignment in azimuth and correction in azimuth during continuous operation (more than 3 ...4 hours) become a serious problem if a two-channel system is used.

Initial azimuth alignment (gyrocompassing) of LG-based systems may be carried out by one of the methods described in [6]. On completion of the horizon alignment, this procedure is reduced to determination of the system angular orientation with respect to the projection of the Earth's rotation vector to the horizon plane. If the LG measuring axis is in the horizon plane and makes an angle α with the meridian plane, the frequency of its output signal will be

$$f = k\Omega_E \cos\varphi \cos\alpha,$$

where k is the LG scale factor. As the angular velocity of the Earth's rotation is very small, the LG output frequency has to be integrated over a time interval T which will ensure high accuracy in measurements. Then the number of periods of the output signal will be

$$N = Tk\Omega_E \cos\varphi \cos\alpha,$$

from where

$$\alpha = \arccos \frac{N}{Tk\Omega_E \cos\varphi}. \qquad (14.7)$$

The latitude φ in Expression (14.7) may be an unknown value. Besides this, the accuracy of the measurements in this case depends on the value of the angle measured. In the range of small angles the cosine function varies slowly with the changes of the argument, therefore the sensitivity of the method decreases.

LG bias exerts a noticeable influence on the accuracy of azimuth determination. To compensate for this influence, the LG has to be oriented in four orthogonal positions. A scheme of rotation for this purpose is shown in Fig. 14.7.

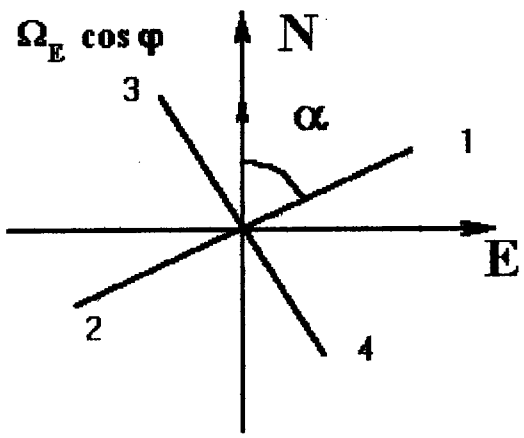


Fig. 14.7. A scheme of LG rotation during gyrocompassing

After the rotation the projection of the Earth's angular velocity to the measuring axis changes the sign, and the bias remains the same, which permits it to be excluded from the result of the measurements. As a result we obtain

$$N_1 = (k\Omega_E \cos\varphi \cos\alpha + \Delta)T,$$
$$N_2 = -(k\Omega_E \cos\varphi \cos\alpha + \Delta)T,$$

where Δ is the LG bias.

$$N_1 - N_2 = 2Tk\Omega_E \cos\varphi \cos\alpha,$$
$$N_3 = [k\Omega_E \cos\varphi \cos(\alpha + 90°) + \Delta]T = (k\Omega_E \cos\varphi \sin\alpha + \Delta)T,$$
$$N_4 = (-k\Omega_E \cos\varphi \sin\alpha + \Delta)T,$$
$$N_3 - N_4 = 2Tk\Omega_E \cos\varphi \cos\alpha,$$
$$\frac{N_3 - N_4}{N_1 - N_2} = \mathrm{tg}\alpha,$$
$$\alpha = \mathrm{arctg}\frac{N_3 - N_4}{N_1 - N_2}. \qquad (14.8)$$

It is possible to show that the accuracy of the azimuth measurements according to the algorithm (14.8) is identical over the whole range.

The experimental research on the operation of the system in this regime, was made. In each angular position (see Fig. 14.7) the number of periods of the LG output signal accumulated over the time interval T, was measured. These intervals were assumed to be equal to 10 or 100 seconds. Several series of measurements, each of 12-15 cycles, were made. The results show that the azimuth may be determined with an error of 2 ...3° in a period of 10 seconds and 0.2 ..0.7° in 100 seconds.

The initial alignment of the azimuth LG may also be carried out by the use of the information from a satellite navigation system (for example, the GPS) while the track measuring car is moving along a rectilinear railway segment. Assume that there is a rectilinear railway segment between points A and B, then the azimuth angle may be deter-mined according to the equation

$$\alpha = \mathrm{arctg}\frac{\Delta y_{BA}}{\Delta x_{BA}} + \pi k,$$

where: Δx_{BA} is an increment in the northern coordinate and Δy_{BA} is an increment in the eastern coordinate; k is an even number, if the sign of $\Delta x_{BA} = +1$, and k is an odd number, if the sign of $\Delta y_{BA} = -1$. Let us estimate the errors of the initial alignment by the use of the GPS. The errors ΔA and ΔB in determining the coordinates of the points A and B yield

$$\Delta\alpha = \mathrm{arctg}\frac{\Delta A + \Delta B}{L},$$

where L is the distance between the points A and B (if $\Delta A = \Delta B = 0.6$ m and L = 1000 m, $\Delta\alpha = 1.2$ mrad = 4.13′). The error resulting from the displacement Δ of the car carriages within the track gauge in the opposite directions is de-fined by the expression

$$\Delta\alpha = 2\mathrm{arctg}\frac{\Delta}{L_b},$$

where L_b is the base of the car (with $\Delta = 10$ mm, $L_b = 12600$ mm, $\Delta\alpha = 1.6$ mrad =5.5′). The total alignment error will also contain the error caused by nonlinearity of the railway segment AB. A significant gain in the course determination accuracy can be achieved by integration of the measurement system and the satellite navigation system with the use of a Kalman filter.

14.4. Integrated navigational systems

A precision track recording system can only be designed with the use of an inertial navigational system. It will guarantee the most complete and accurate information for determination of necessary parameters. But even very good INS have navigation errors of about 1 nautical mile within 24 hours or ± 1.2 m per minute which will result in rather big errors in de-termination of the measuring car location. So in tunnels longer than 3 km we will risk to have errors of more than ± 0.5 m. The best way to solve this problem is integration of a measuring system with other sources of navigation information.

In our case the integrated system will permit the solution of two problems: a) to estimate and reduce the influence of the INS sensor errors and to improve INS accuracy in angular orientation measurement; б) to improve INS accuracy in location determination.

14.4.1. INS/Satellite Navigation System Integration

The INS exhibits low noise but tends to drift over time and, on the contrary, the GPS errors are noisy but exhibit no long-term drifts. Integration of the information from both sensors yields a navigation system which operates as a drift-free INS [7] .

In order to obtain error equations, two coordinate systems are introduced: $O\xi\eta\zeta$ is the geographic frame ($O\xi$ is directed tangentially to the meridian to the North, $O\eta$ - to the East tangentially to the parallel and the axis $O\zeta$ - down-wards along the vertical; Oxyz is fixed to the gyro stabilized platform of the local-level INS or the "calculated platform" for the case of a strapdown INS (SINS). Mutual angular orientation of these two coordinate systems is shown in Fig. 14.8.

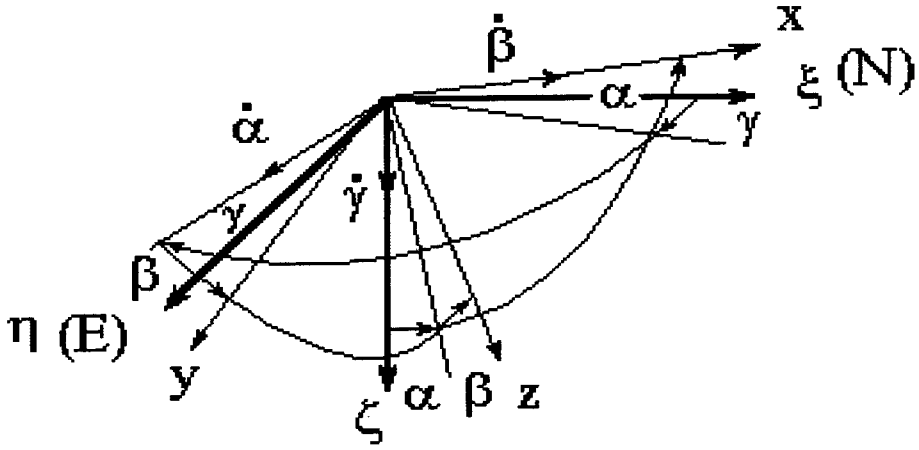


Fig. 14.8. Mutual angular orientation of the geographic and stabilized platform coordinate systems.

Making some allowable assumptions (the Earth does not rotate and has a spherical shape), we can write simplified equations for the local-level INS errors:

$$\dot{\Delta x} = \Delta V_x - \gamma V_E; \qquad \dot{\alpha} = -\frac{V_x}{R} + \delta_{yc} + \delta_{yr};$$

$$\dot{\Delta y} = \Delta V_y + \gamma V_N; \qquad \dot{\beta} = \frac{V_y}{R} + \delta_{xc} + \delta_{xr};$$

$$\dot{\Delta z} = \Delta V_z; \qquad \dot{\gamma} = \delta_{zc} + \delta_{zr};$$

$$\dot{\Delta V_x} = g\alpha + a_{xc} + a_{xr}; \qquad \dot{a}_{ic} = 0;$$

$$\dot{\Delta V_y} = -g\beta + a_{yc} + a_{yr}; \qquad \dot{a}_{ir} = -\mu_{ai} a_{ir} + \sigma_{ai}\sqrt{2\mu_{ai}}\,w;$$

$$\dot{\Delta V_z} = -g\alpha - g\beta + a_{zc} + a_{zr}$$

$$\dot{\delta}_{ic} = 0; \qquad (14.9)$$

$$\dot{\delta}_{ir} = -\mu_{\delta i}\delta_{ir} + \sigma_{\delta i}\sqrt{2\mu_{\delta i}}\,w$$

where: Δx , Δy, Δz - the error components in determination of the coordinates ; ΔV_x , ΔV_y , ΔV_z - the error components in determination of the velocity ; α,β,γ - the angular errors in determination of the vertical and the course; a_{ic}, a_{ir} $(i=x,y,z)$ - the constant and the random components of the accelerometer; δ_{ic}, δ_{ir} - the constant and the random components of the gyro drifts. The random gyro and accelerometer errors are given by the correlation function $R(\tau) = = \sigma_i^2 e^{-\mu|\tau|}$, where σ_i - RMS, and μ_i - the factor of damping.

This mathematical model can easily be transformed into a model of a SINS by projecting the gyro and accelerometer errors from the platform coordinate system into the geographic frame.

The Kalman filtering algorithm in an ordinary form can be used to simulate the integrated system . Equation (14.9) must be presented in a vector-matrix form and added by the equation for vector Z measurement. After transition from the differential equations to the difference equations for the discrete system it is possible to write:

$$x_{k+1} = \Phi_{k+1,k} x_k + \Gamma_k u_k \,, \qquad (14.10)$$

$$z_k = H_k x_k + v_k \,,$$

where: x_k - the state vector of the system of [21x1] dimension at a k-th cycle; the discrete step response matrix Φ can be derived from the expression $\Phi \approx \exp(AT) \approx E+AT$, where A is a state, T - a discrete time interval, E - a unit matrix; Γk - a matrix of input; zk и Hk - a vector and matrix of measurement, vk - a vector of measurement errors.

14.4.2. Integration of an INS with other devices of the measuring car

Integration of the system with other devices of the measuring car and sources of navigation information offers new possibilities for the system improvement . Let us consider a measuring car which is equipped with an inertial navigation system (INS), a GPS receiver, a sensor for detecting the navigation marks (NM) position, and an odometer-speedometer (OS) operating from wheels. NM are marks on the track whose locations are well known (milestones, InduSI, a gap of switches, a gap of rail crossings). We want to get accurate information about the car position by integration of these information sources using a Kalman filter. A block diagram of an integrated navigation system is shown in Fig. 14.9.

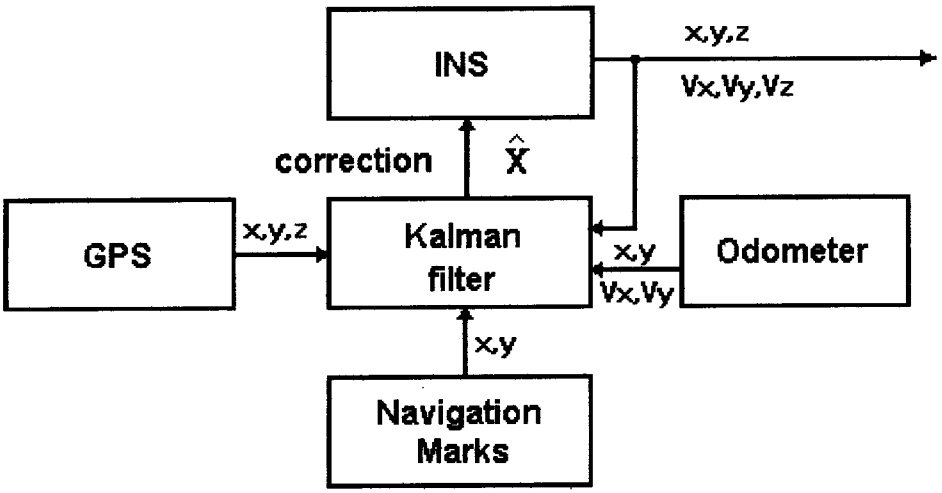


Fig. 14.9. A block diagram of an integrated navigation system.

Operation of the system was simulated in different modes. The examples of the simulation are shown in Figures 14.10 and 14.11. At an instant of 15 min the car enters the tunnel (shadow) and the GPS signals are lost. In

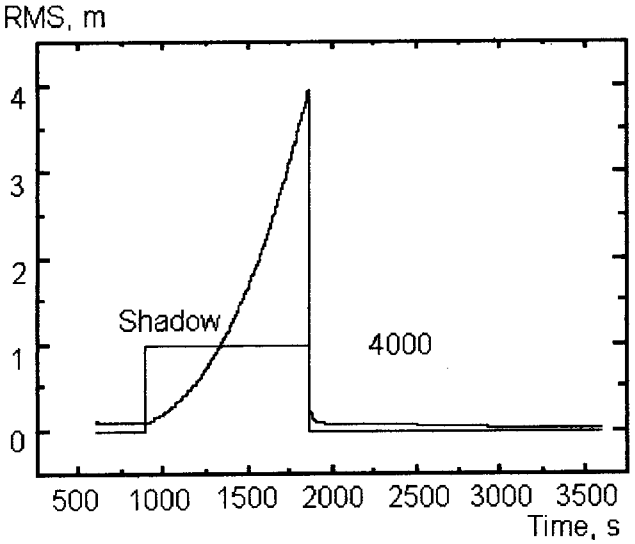


Fig 14.10.

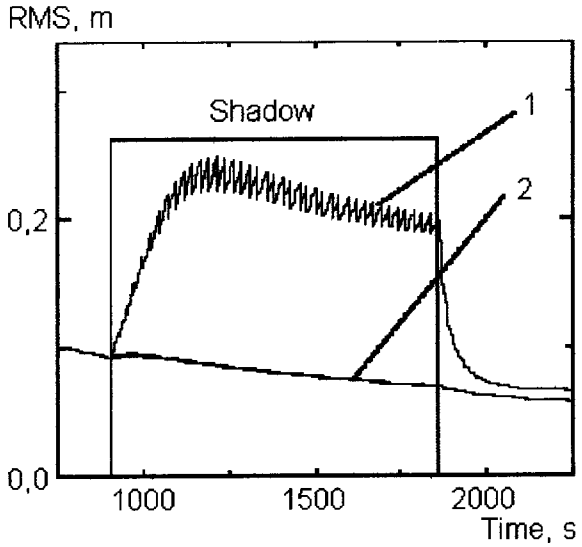


Fig. 14.11.

Fig.14.10 there is not any correction in the tunnel and the coordinate estimation error is growing. In Fig.14.10 the correction with navigation marks is carried out in the tunnel. Curve 1 corresponds to case 1 at 1 km mark, and curve 2 - to case 1 at 100 m mark.

The parameters of the devices and systems in these simulations were as follows: INS updating frequency - 10 Hz; position error growth rate - σ = 1 m/min; GPS updating frequency - 2 Hz; position error - σ= 1 m; NM position determination error - σ= 0.5 m. The car speed - 150 km/h. The parameters of the INS state vector: gyro drifts - 0.01°/h; initial attitude alignment errors - 1.5'; errors in determination of the initial speed - 1 m/s; initial coordinate errors - 1.5 m; accelerometer drifts - 0.0005g.

Figures 14.12 and 14.13 present the simulation results of the filter operating in the other two modes: Fig. 14.12 - GPS + NM - the whole time except that of the run in the tunnel (shadow); in the shadow - only NM, NM every 5 km, NM and GPS errors - 0.5 m . All the coordinates (x,y,z) have practically the same graphs. Fig. 14.13 presents the same mode, except the error of the GPS which is 2.5 m. Curve 1 - without the GPS and NM, curve 2 - only NM. Updating frequency - 100 Hz, GPS updating frequency - 2 Hz, car speed - 100 km/h, gyro drifts - 0.001 °/h, errors in determination of the initial speed - 0.5 m/s, initial coordinate errors - 1 m; accelerometer drifts - 0.00005g.

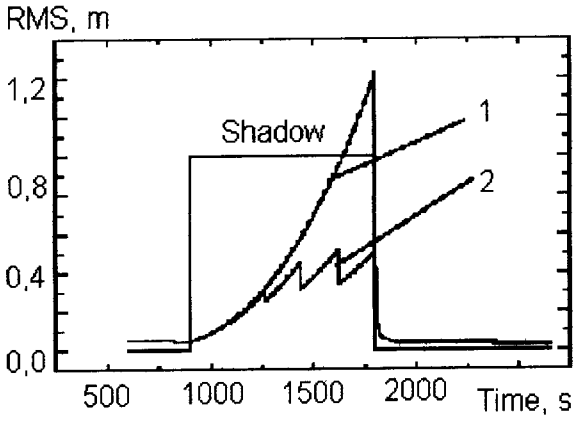


Fig. 14.12.

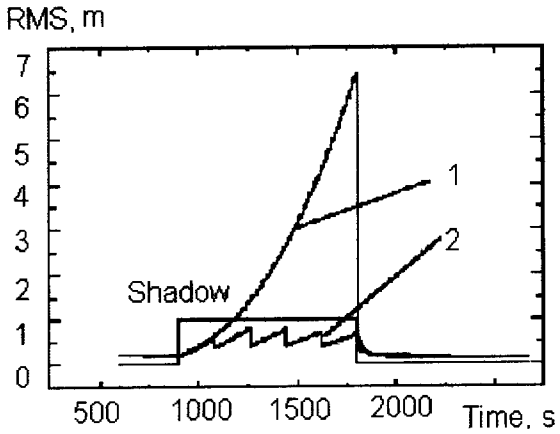


Fig.14.13.

So, to design a precision track measuring system, we need a highly accurate INS and a reference to an error less than ± 0.5 m. Therefore, in the differential mode the GPS should be in an error of ± 0.5 m. The simulation also shows that NMs should be placed at a distance of 5 km from each other in the regions where no GPS is available. If it is too expensive, it would be desirable to have a more accurate INS. An odometer can improve the accuracy of the system only when the position information is not available, taking into account that the system observability with the position information is more complete than with the information about the velocity, and the odometer information has a position determination error accumulated in time.
The prospects for the further development of the integrated system depend on the right choice of accurate and adequate methods and types of the units used and algorithms implemented in the navigation system.

14.5. The prospects for the track measurement system development

An essential increase of the system accuracy and extension of its functional possibilities can be ensured by the use of preliminary information about the railway conditions and geometry. High-frequency output of information permits measurement of not only long deviations, but also intersleepers deflections, quality of joints, etc., which is particularly important for monitoring of high-speed railways.

A rather low cost, small dimensions and simplicity of service will allow each high-speed train to be equipped with such a system. The possibility to record, store and transfer the measurement results, as well as the opportunity to analyse the current changes in the condition of the railway, will make it possible to choose an optimum high-speed mode and make traffic safer. Some other problems concerning inspection of mutual location of the track and the overhead line (for the railway) or the contact rail (for the underground) have to be solved. Special sensors measure the location of the overhead line with respect to the car, and it is necessary to measure and exclude the car angular motion with respect to the track. Simultaneous elaborate processing of all data provides accurate and complete information about the condition of the track and the power supply equipment. The processing has to be carried out in real time by special computing means on the basis of high-speed signal processors.

The area of the system application may be extended for: recording of a dynamic condition of various railways (tram-lines, tunnel and special-purpose railways, etc.) under a load; monitoring of asphalt and concrete roads, pipelines, airport runways, ground surfaces, etc.
Of special interest is the possibility to use the measuring system in construction and maintenance of the railway.

References:

1. Hans-Jörg Höhberger, Martin Rechel und Heinz-Herbert Zück. Die Oberbau-Meßwagen-Einheit - Neuer Standard für die Gleismeßtechnik. Gleismeßtechnik, 41, H.6-Juni, 1992, S. 405-411. (In German).
2. Alekseyev A.E., Zenzinov B.N. Measuring Station Ways Structure by the Track Recording Cars. - Izvestia of LIIZhT, №6, 1986, p. 41-44. (In Russian).
3. Luis R.B. High-speed Track Recording Car. - Zheleznyie dorogi mira, №12, 1977, p.66-68. (In Russian).
4. Inertial navigation systems of marine objects / D.P. Loukianov, A.V. Mochalov, A.A. Odintsov, I.B. Vaisgant.- L: Sudostroenie, 1989, 184 p. (In Russian).
5. Mochalov A.V., Buravlev A.S. Peculiarity of laser gyro operation on a rocking base. - Izvestia of LETI. Issue 286, 1981, pp 56 -62 . (In Russian).
6. Laser measuring systems /D.P. Loukianov, A.V. Mochalov, J.V. Filatov, and al. M: Radio i Sviaz, 1981, 456 p. (In Russian).
7. Phillips R.E., Schmidt G.T., " GPS / INS integration", System Implications and Innovative Applications of Satellite Navigation, AGARD-LS-207, Neuilly-sur-Seine, France, July 1996.

15. A SYSTEM FOR MEASURING DEFORMATIONS OF LARGE-SIZED OBJECTS

A.V. Mochalov
St.Petersburg Electrotechnical University
Professor Popov str. 5, St.Petersburg, Russia

15.1. A METHOD FOR MEASURING STATIC AND DYNAMIC DEFORMATIONS

Present-day mobile objects (vehicles) are equipped with a large number of different devices and systems whose mutual angular position must be known to a high degree of accuracy. Owing to static and dynamic deformations of the object, the angular position of peripheral apparatus (radar antennas, launching gears, optical systems, etc.) may differ essen-tially from that of central devices. The problems of evaluation and proper allowance for angular misalignments become particularly urgent for large-sized vehicles (tankers, big ships, groundborne vehicles, spacecrafts and orbital bases, etc.)

The main reasons of the vehicle's static angular deformations (that can be as high as 1-1.5°) are redistribution of freight and fuel on it and non-uniform heating of different parts of the vehicle under the sun. Dynamic (elastic) angular defor-mations are caused by motion disturbances, sea waves, helm's operation, etc. The experience shows that the values of such deformations for ships are equal to units of angular minutes in the plane of the deck (yawing angle), dozens of angular minutes in the longitudinal central plane (pitching angle) and fractions of an angular minute for the torsion angle (rolling angle). These values can essentially grow if the apparatus are installed on superstructures and masts. The prob-lem of deformation measurement has to be solved for the case of initial alignment of inertial navigation systems (INS) on moving objects, for example, initial alignment of an airborne INS aboard an aircraft-carring ship or alignment of a winged missile INS aboard a heavy missile-carrier aircraft. Besides, continuous monitoring of deformations of a moving vehicle is necessary for prediction and prevention of possible destructions and occurrence of emergencies. It is particularly important for vehicles transporting petroleum and other dangerous freights or for those with nuclear power installations, when a crash could make the ecological situation worse.

Among various methods for measuring angular deformation (a rather complete review of these methods can be found in [1]), optical, photographic and photogrammetric methods are the most widespread. But often they are not universal or accurate enough. A new promising method was offered in [2] and is being quickly developed. In accordance with it, two three-axis laser gyro units (LGUs) are installed: the first one - near the basic device, and the second - near the peripheral device. Comparing their indications, it is possible to determine the value of their current angular misalignment. If the deformation angles are assumed to be small enough, they linearly depend on the LGU output signals, making it possible to use a Kalman filter for processing measurement results.

15.1.1. Equations of the method

Three-axis LGUs (LGU1 and LGU2 in Fig.15.1) are fixed at two or more points of the ship where controllable devices are located. The LGU1's axes oxyz are aligned with the coordinate system of the ship, and the axis oy is aligned with the longitudinal axis of the ship, oz is perpendicular to the deck and directed upwards, ox supplements the two previous axes to form the right-handed orthogonal coordinate system.

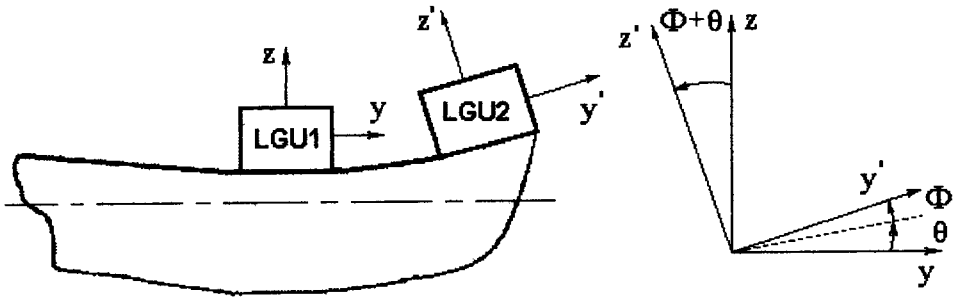


Fig. 15.1

Assume that the angular misalignment of the two LGUs has a static component $\vec{\Phi}$ and a dynamic one $\vec{\theta}$. The total turn vector $\vec{\varphi} = \vec{\Phi} + \vec{\theta}$ can be presented in a matrix form

$$\varphi = \left| \varphi_x \ , \varphi_y \ , \varphi_z \right|^T .$$

The LGU1 measures the absolute angular velocity of the ship $\vec{\Omega}$ in the projections on the axes oxyz, while the LGU2

measures the angular velocity $\vec{\Omega}'$ in the projections on the axes $ox'y'z'$, with

$$\vec{\Omega}' = \vec{\Omega} + \dot{\vec{\varphi}} , \qquad (15.1)$$

where $\dot{\vec{\varphi}}$ is the angular velocity of the LGU1 and LGU2 relative motion caused by the elastic deformations ($\dot{\vec{\varphi}} = \dot{\vec{\theta}}$, $\dot{\vec{\Phi}} = 0$). The coordinate system OXYZ is converted to that of $ox'y'z'$ in accordance with the expression

$$[x', y', z'] = B[x, y, z], \qquad (15.2)$$

where B is the direction cosine matrix.

Let us introduce the following designation: α as a bend angle in the plane of the deck; β - a bend angle in the longitudinal plane of the ship; γ - a torsion angle around the longitudinal axis. In view of the small value of the deformation angles

$$B \approx \begin{vmatrix} 1 & \alpha & -\gamma \\ -\alpha & 1 & \beta \\ \gamma & -\beta & 1 \end{vmatrix} = \begin{vmatrix} 1 & \varphi_z & -\varphi_y \\ -\varphi_z & 1 & \varphi_x \\ \varphi_y & -\varphi_x & 1 \end{vmatrix}.$$

After projecting the vector equation (15.1) on the axes $ox'y'z'$, we shall obtain

$$\Omega' = B\Omega + \dot{\varphi}.$$

The difference between the angular velocities measured by the two LGUs will be

$$\Delta\Omega = \Omega - \Omega' = \Omega - B\Omega - \dot{\varphi} = (E - B)\Omega - \dot{\varphi}$$

or in an expanded form

$$\begin{vmatrix} \Delta\Omega_x \\ \Delta\Omega_y \\ \Delta\Omega_z \end{vmatrix} = \begin{vmatrix} \Omega_x - \Omega'_x \\ \Omega_y - \Omega'_y \\ \Omega_z - \Omega'_z \end{vmatrix} = \left\{ \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} - \begin{vmatrix} 1 & \varphi_z & -\varphi_y \\ -\varphi_z & 1 & \varphi_x \\ \varphi_y & -\varphi_x & 1 \end{vmatrix} \right\} \begin{vmatrix} \Omega_x \\ \Omega_y \\ \Omega_z \end{vmatrix} - \begin{vmatrix} \varphi'_x \\ \varphi'_y \\ \varphi'_z \end{vmatrix} =$$

$$= \begin{vmatrix} 0 & -\varphi_z & \varphi_y \\ \varphi_z & 0 & -\varphi_x \\ -\varphi_y & \varphi_x & 0 \end{vmatrix} \begin{vmatrix} \Omega_x \\ \Omega_y \\ \Omega_z \end{vmatrix} - \begin{vmatrix} \varphi'_x \\ \varphi'_y \\ \varphi'_z \end{vmatrix} = \hat{\varphi}\Omega - \varphi',$$

Where $\hat{\varphi}$ is a skew-symmetric matrix made of projections of the vector $\vec{\varphi}$ on the axes ox'y'z'. But it is known that $\hat{\varphi}\Omega = -\hat{\Omega}\varphi$, therefore, with due account of $\dot{\varphi} = \dot{\theta}$, we shall derive

$$\Delta\Omega = \hat{\Omega}\varphi - \dot{\varphi} = \hat{\Omega}\varphi - \dot{\theta}, \qquad (15.3)$$

where $\hat{\Omega}$ is a skew-symmetric matrix

$$\hat{\Omega} = \begin{vmatrix} 0 & \Omega_z & -\Omega_y \\ -\Omega_z & 0 & \Omega_x \\ \Omega_y & -\Omega_x & 0 \end{vmatrix}.$$

It is easy to see that the system of the differential equations (15.3) has no analytical decision. This testifies to the fact that the measuring system constructed in such a way is not quite observable (the observability analysis will be discussed in Section 15.3). At the same time the deformation parameters $\Phi, \vartheta, \dot{\vartheta}$ can be estimated by the results of repeated successive measurements and processing with a Kalman filter.

15.1.2. Use of a Kalman filter.

In order for a Kalman filter (KF) to be used, the measurement equation (15.3) should be presented in a matrix form as

$$z = Hx + v, \qquad (15.4)$$

where z is a measurement vector; H - a measurement matrix; x - a state vector; v - a matrix-column formed from measure-ment errors. The state vector is of the [1 x 9] size

$$x = \left| \Phi_x, \Phi_y, \Phi_z, \theta_x, \theta_y, \theta_z, \dot{\theta}_x, \dot{\theta}_y, \dot{\theta}_z \right|^T,$$

with the measurement matrix in this case taking the form of

$$H = \begin{vmatrix} 0 & \Omega_z & -\Omega_y & 0 & \Omega_z & -\Omega_y & -1 & 0 & 0 \\ -\Omega_z & 0 & \Omega_x & -\Omega_z & 0 & \Omega_x & 0 & -1 & 0 \\ \Omega_y & -\Omega_x & 0 & \Omega_y & -\Omega_x & 0 & 0 & 0 & -1 \end{vmatrix}.$$

If necessary, the state vector may be augmented, for example, by introducing gyro drifts or other parameters. The dynamic deformation components may be represented as random processes approximated by the correlation function

$$K_{\theta_i}(\tau) = D_i e^{-\mu_i |\tau|} \left(\cos\lambda_i \tau + \frac{\mu_i}{\lambda_i} \sin\lambda_i |\tau|\right),$$

where μ_i is the irregularity coefficient; λ_i - the prevailing variation frequency of the deformation angles. So, the shaping filter equation of the random processes with such a correlation function will be:

$$\ddot{\theta}_i + 2\mu_i \dot{\theta}_i + b_i^2 \theta_i = 2 b_i \sqrt{D_i \mu_i}\, w(t),$$

where $b_i^2 = \mu_i^2 + \lambda_i^2$, w(t) is the white noise. Thus, a set of state equations may be written in the following way:

$$\begin{aligned} &\dot{x}_1 = 0; \dot{x}_2 = 0; \dot{x}_3 = 0; \\ &\dot{x}_4 = x_7; \dot{x}_5 = x_8; \dot{x}_6 = x_9; \\ &\dot{x}_7 = -2\mu_x x_7 - b_x^2 x_4 + 2 b_x \sqrt{D_x \mu_x}\, w; \\ &\dot{x}_8 = -2\mu_y x_8 - b_y^2 x_5 + 2 b_y \sqrt{D_y \mu_y}\, w; \\ &\dot{x}_9 = -2\mu_z x_9 - b_z^2 x_6 + 2 b_z \sqrt{D_z \mu_z}\, w; \end{aligned} \qquad (15.5)$$

These equations may be also represented in a matrix form as

$$\dot{x} = Ax + Bw, \qquad (15.6)$$

by the use of the state matrix

$$A = \begin{vmatrix} 0_3 & & 0_3 & & & 0_3 & \\ 0_3 & & 0_3 & & & E_3 & \\ & -b_x^2 & 0 & 0 & -2\mu_x & 0 & 0 \\ 0_3 & 0 & -b_y^2 & 0 & 0 & -\mu_y & 0 \\ & 0 & 0 & -b_z^2 & 0 & 0 & -\mu_z \end{vmatrix},$$

and the noise matrix

$$B = \begin{vmatrix} & 0_3 & \\ & 0_3 & \\ 2b_x\sqrt{D_x\mu_x} & 0 & 0 \\ 0 & 2b_y\sqrt{D_y\mu_y} & 0 \\ 0 & 0 & 2b_z\sqrt{D_z\mu_z} \end{vmatrix},$$

where 0_3 and E_3 are the null and identity [3 x 3] matrices. After Equations (15.4, 15.6) have been represented in the form of difference equations, a standard Kalman filtering procedure for optimal estimation of the system state vector can be implemented [2].
The effectiveness of Kalman filtering in various modes of an object motion was checked by simulation of the measuring system. An example of such simulation results for the case of a marine vehicle is given below. The parameters of the vehicle's angular velocity and angular deformation are shown in Table 15.1 [3].

Table 15.1 Initial data for simulation

Angular velocity parameters	Amplitude, deg	Period, s	Static deformation, min	Dynamic deformation, min
Rolling	14	7	$\sigma_\gamma = 12$	$\tilde{\sigma}_\gamma = 0,5...4$
Pitching	3	9	$\sigma_\beta = 20$	$\tilde{\sigma}_\beta = 3...7$
Yawing	4	30	$\sigma_\alpha = 5$	$\tilde{\sigma}_\alpha = 1...2$

The results of the computations are shown in Fig. 15.2. a, b. As follows from it, the dynamic deformation components may be estimated to an error of 10 ...20″ in 1... 2 minutes. Static misalignments relative to transversal and vertical axes are estimated in about 3...5 minutes. As for the estimation of static misalignment relative to the longitudinal axis under given conditions, the process takes more than 10 minutes. In order to accelerate it, it would be desirable to conduct a special manoeuvre changing the course [3].

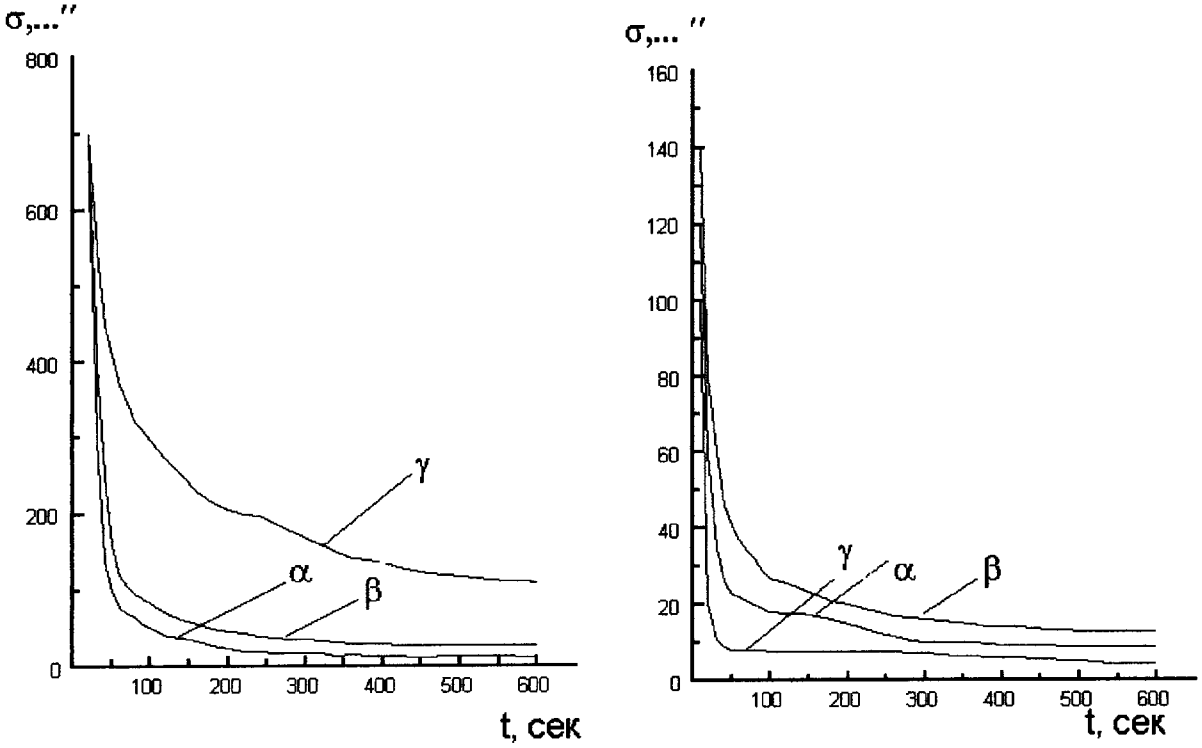


Fig. 15.2 a,b

Observability of the process and estimation of the accuracy were found to have strong dependence on mutual correlation between changes in angular motion and deformation angles of the vehicle. If the correlation is strong, estimation errors grow. In this connection, the problem of the system observability, and related to it, the accuracy estimation problem, as well as the ways of improving their solution, require closer consideration.

15.2. Observability and errors of the system

15.2.1. Static deformation

It is known that the errors of measuring systems contain instrumental, methodical and calculation errors. This section considers only methodical errors due to partial observability of the system. Following the reference [2], we shall analyze the cases of static and dynamic deformations separately.

For the case of static deformation the matrix equation (15.3) takes the form

$$\Delta\Omega = \hat{\Omega}\varphi . \qquad (15.7)$$

Thus, to determine the deformation components φ_x , φ_y , φ_z, we have three linear algebraic equations with an infinite set of decisions or having no decisions in view of $\det \hat{\Omega} = 0$. From the linear systems theoretical standpoint, System (15.7) is not quite observable, as the rank r=2 of the matrix $\hat{\Omega}$ is less than the number of undetermined values n=3: r <n. With Equation (15.7) presented in a vector form

$$\Delta\vec{\Omega} = \vec{\Omega} \times \vec{\varphi}, \qquad (15.8)$$

it is easy to show that in Equation (15.8) only the deformation component, perpendicular to the basic vector $\vec{\Omega}$, can be completely estimated. Really, multiplying the vector $\vec{\Omega}$ by the both parts of Equation (15.8), we shall derive

$$\vec{\Omega} \times \Delta\vec{\Omega} = \vec{\Omega} \times (\vec{\varphi} \times \vec{\Omega}) = \vec{\varphi}(\vec{\Omega} \cdot \vec{\Omega}) - \vec{\Omega}(\vec{\Omega} \cdot \vec{\varphi}) . \qquad (15.9)$$

From Equation (15.9) we can rderive the known formula for decomposing the vector $\vec{\varphi}$ in the direction of the vector $\vec{\Omega}$ and the one perpendicular to it, i.e.

$$\vec{\varphi} = \frac{\vec{\Omega} \times (\vec{\varphi} \times \vec{\Omega})}{|\vec{\Omega}|^2} + \frac{\vec{\Omega}(\vec{\Omega} \cdot \vec{\varphi})}{|\vec{\Omega}|^2} = \vec{\varphi}_{N\Omega} + \vec{\varphi}_{\Omega} . \qquad (15.10)$$

From Equation (15.10) it follows that using Formula (15.8), it is possible to determine only normal to the basic vector $\vec{\Omega}$ deformation component $\vec{\varphi}_{N\Omega}$. In this case the deformation component $\vec{\varphi}_{\Omega}$ is a methodical error.

It is known [4] that in order to determine mutual angular orientation of two triads without a methodical error, it is necessary either to measure the same vector at two different instants (if this vector changes its angular orientation), or to measure two vectors of different physical nature simultaneously. We shall consider both of these cases. Let the vector $\vec{\Omega}$ has the values $\vec{\Omega}_1$ and $\vec{\Omega}_2$ in two successive measurements, and in accordance with Equation (15.10)

$$\vec{\varphi}_1 = \vec{\varphi}_2 = \vec{\varphi}_{N\Omega_1} + \frac{\vec{\Omega}_1(\vec{\Omega}_1 \cdot \vec{\varphi})}{|\Omega_1|^2} = \vec{\varphi}_{N\Omega_2} + \frac{\vec{\Omega}_2(\vec{\Omega}_2 \cdot \vec{\varphi})}{|\Omega_2|^2} ,$$

whence it follows that

$$\frac{\vec{\Omega}_1(\vec{\Omega}_1 \cdot \vec{\varphi})}{|\Omega_1|^2} - \frac{\vec{\Omega}_2(\vec{\Omega}_2 \cdot \vec{\varphi})}{|\Omega_2|^2} = \vec{\varphi}_{N\Omega_2} - \vec{\varphi}_{N\Omega_1} . \qquad (15.11)$$

Multiplying both parts of Equation (15.11) by the vector $\vec{\Omega}_2$, we shall obtain

$$\frac{\vec{\Omega}_1 \times \vec{\Omega}_2(\vec{\Omega}_1 \cdot \vec{\varphi})}{|\Omega_1|^2} = \vec{\varphi}_{N\Omega_2} \times \vec{\Omega}_2 - \vec{\varphi}_{N\Omega_1} \times \vec{\Omega}_2 . \qquad (15.12)$$

Doing scalar multiplication of both parts of Equation (15.12) by the vector $\vec{\Omega}_3 = \vec{\Omega}_1 \times \vec{\Omega}_2$, we come to equality

$$\frac{(\vec{\Omega}_1 \cdot \vec{\varphi})|\vec{\Omega}_3|^2}{|\Omega_1|^2} = \left[\frac{(\vec{\Omega}_2 \times d\Omega_2) \times \Omega_2}{|\Omega_2|^2} - \frac{(\vec{\Omega}_1 \times d\vec{\Omega}_1) \times \vec{\Omega}_1}{|\Omega_1|^2}\right] \cdot (\vec{\Omega}_1 \times \vec{\Omega}_2) \quad ,(15.13)$$

from which it is possible to determine the unknown component $\vec{\varphi}_{\Omega_1}$. In view of

$$(\vec{\Omega}_2 \times d\vec{\Omega}_2) \times \vec{\Omega}_2 = -\vec{\Omega}_2\left(\vec{\Omega}_2 \cdot d\vec{\Omega}_2\right) + d\vec{\Omega}_2\left(\vec{\Omega}_2 \cdot \vec{\Omega}_2\right) = d\vec{\Omega}_2\left|\vec{\Omega}_2\right|^2 ;$$

$$(\vec{\Omega}_1 \times d\vec{\Omega}_1) \times \vec{\Omega}_2 = -\vec{\Omega}_1\left(\vec{\Omega}_2 \cdot d\vec{\Omega}_1\right) + d\vec{\Omega}_1\left(\vec{\Omega}_2 \cdot \vec{\Omega}_1\right) ;$$

$$\left[\vec{\Omega}_1\left(\vec{\Omega}_2 \cdot d\vec{\Omega}_1\right)\right] \cdot \left(\vec{\Omega}_1 \times \vec{\Omega}_2\right) = 0 ,$$

from Expressions (15.10) and (15.13) we shall finally derive

$$\vec{\varphi} = \frac{\vec{\Omega}_1 \times d\vec{\Omega}_1}{\left|\vec{\Omega}_1\right|^2} - \frac{\left(\vec{\Omega}_1 \times \vec{\Omega}_2\right) \cdot \left[d\vec{\Omega}_2\left|\vec{\Omega}_1\right|^2 - d\vec{\Omega}_1\left(\vec{\Omega}_2 \cdot \vec{\Omega}_1\right)\right]}{\left|\vec{\Omega}_1\right|^2\left|\vec{\Omega}_3\right|^2}\vec{\Omega}_1 . \quad (15.14)$$

In Equation (15.14) the first term is the same as the observable component in the first measurement, and the second term excludes the methodical error by the use of the data obtained in the second measurement. Thus, as a result of two measurements of the absolute angular velocity it is possible to estimate the static deformation without a methodical error. The system becomes quite observable under the condition of

$$\vec{\Omega}_3 = \vec{\Omega}_1 \times \vec{\Omega}_2 \neq 0.$$

15.2.2. Dynamic and composite deformation

In this case Expression (15.3) is to be used. It may be represented in a vector form as

$$d\vec{\Omega} = \vec{\varphi} \times \vec{\Omega} - \dot{\vec{\varphi}} , \qquad (15.15)$$

where $\dot{\vec{\varphi}}$ is an additional angular velocity of the peripheral device with respect to the central one due to elastic deformations. Similarly to the previous variant, using Expression (15.10), decompose $\vec{\varphi}$ into the directions, one of which is perpendicular to the measured vector $\vec{\Omega}$, and the other is aligned with it. In view of Equation (15.15) the decomposition (15.10) will take the following form

$$\vec{\varphi} = \frac{\vec{\Omega} \times d\vec{\Omega}}{\left|\vec{\Omega}\right|^2} + \frac{\vec{\Omega} \times \dot{\vec{\varphi}}}{\left|\vec{\Omega}\right|^2} + \frac{\vec{\Omega}\left(\vec{\varphi} \cdot \vec{\Omega}\right)}{\left|\vec{\Omega}\right|^2} = \vec{\varphi}_{N\Omega} + \vec{\varphi}_{N\dot{\varphi}} + \vec{\varphi}_{\Omega} . \quad (15.16)$$

According to Expression (15.16) only the first member can be determined from the measurement results, the two others introduce a methodical error. Thus, even the component of deformation perpendicular to the basic vector is estimated with an error that is as much higher as $\left|\vec{\Omega} \times \dot{\vec{\varphi}}\right|$ is.

In this case even the repeated measurements of successive states of the vector $\vec{\Omega}$ do not permit the deformation to be determined without methodical errors. They can only be reduced with a Kalman filter, as shown in Section 15.2. Thus, the system is not quite observable, the filter convergence and stability depend on the rate of the change in orientation of the vector $\vec{\Omega}$, the correlation of the vectors $\vec{\varphi}$ and $\vec{\Omega}$ and other parameters of the system.

Many ways of increasing the system observability can be offered. Let us consider one of them, namely, invoking the second vector to the measurements, for example, the one formed by the indications of the output data of three-axis accelerometer units adjacent to the LGUs. The difference of the accelerations, measured at two points of the ship, may be presented as

where $\vec{\varepsilon}=\dot{\vec{\Omega}}$ is the angular acceleration of the ship; $\vec{R}$ is the radius vector of the point, where the peripheral device is installed, with respect to the centre of rotation; $\vec{\varepsilon}\times\vec{R}$ is the linear acceleration due to the angular acceleration of the ship, and $\vec{\Omega}\times\left(\vec{\Omega}\times\vec{R}\right)$ is the centripetal acceleration. The contribution of the components $\dot{\vec{\varphi}}$ and $\ddot{\vec{\varphi}}$ into the second and third terms, induced by elastic deformation, are neglected in view of the fact that they give the second-order terms. Designating

$$\vec{c}=\vec{\varphi}\times\vec{A}=d\vec{A}+\vec{\varepsilon}\times\vec{R}+\vec{\Omega}\times\left(\vec{\Omega}\times\vec{R}\right), \qquad (15.18)$$

similar to Equation (15.10), we shall write

$$\vec{\varphi}=\frac{\vec{A}\times\vec{c}}{\left|\vec{A}\right|^2}+\frac{\vec{A}\left(\vec{A}\cdot\vec{\varphi}\right)}{\left|\vec{A}\right|^2}, \qquad (15.19)$$

whence it follows that the acceleration measurement with two three-axis units permits the deformation to be estimated with a methodical error, the value of which is determined by the second term. Solving Equations (15.16) and (15.19) simultaneously, by the same methods that were used for Equation (15.14), we shall be able to estimate the deformation from the results of simultaneous measurements of the vectors $\vec{\Omega}$ and $\vec{A}$:

$$\vec{\varphi}=\frac{\vec{\Omega}\times\vec{B}}{\left|\vec{\Omega}\right|^2}+\frac{\vec{\Omega}}{\left|\vec{\Omega}\times\vec{A}\right|^2}\left[\frac{\vec{A}\times\left(\vec{\Omega}\times\vec{B}\right)}{\left|\vec{\Omega}\right|^2}-\frac{\vec{A}\times\left(\vec{A}\times\vec{c}\right)}{\left|\vec{A}\right|^2}\right]\cdot\left(\vec{\Omega}\times\vec{A}\right). \qquad (15.20)$$

where $\vec{B}=d\vec{\Omega}+\dot{\vec{\varphi}}$.

Similarly, decomposing the vector $\vec{\varphi}$ into the direction of the vector $\vec{A}$ and the one perpendicular to it, it is possible to obtain

$$\vec{\varphi}=\frac{\vec{A}\times\vec{c}}{\left|\vec{A}\right|^2}+\frac{\vec{A}}{\left|\vec{A}\times\vec{\Omega}\right|^2}\left[\frac{\vec{\Omega}\times\left(\vec{\Omega}\times\vec{B}\right)}{\left|\vec{\Omega}\right|^2}-\frac{\vec{\Omega}\times\left(\vec{A}\times\vec{c}\right)}{\left|\vec{A}\right|^2}\right]\cdot\left(\vec{\Omega}\times\vec{A}\right). \qquad (15.21)$$

In the right parts of Equations (15.20) and (15.21) among the terms that are known from the measurements, there is an unknown one, $\vec{B}$, containing the angular velocity of the deformation $\dot{\vec{\varphi}}$, which induces a methodical error. Let us separate the terms, containing $\dot{\vec{\varphi}}$, by transforming Equation (15.21) as follows

$$\begin{aligned}\vec{\varphi}=&\frac{\vec{A}\times\vec{c}}{\left|\vec{A}\right|^2}+\frac{\vec{A}}{\left|\vec{A}\times\vec{\Omega}\right|^2}\left[\left(\vec{\Omega}\cdot\vec{A}\right)\left(\vec{\Omega}\cdot\vec{\varphi}_{NA}\right)+\left(\vec{\Omega}\times d\vec{\Omega}\right)\cdot\vec{A}\right]+\\&+\frac{\vec{A}}{\left|\vec{A}\times\vec{\Omega}\right|^2}\left[\left(\vec{A}\times\vec{\Omega}\right)\cdot\dot{\vec{\varphi}}\right].\end{aligned} \qquad (15.22)$$

The analysis of Equation (15.22) leads to the following conclusion. Simultaneous measuring of the angular velocity $\vec{\Omega}$ and the acceleration $\vec{A}$ makes the system completely observable and permits the deformation to be estimated with a methodical error due to the angular velocity of the deformation $\dot{\vec{\varphi}}$. This error is aligned with the vector $\vec{A}$, its value depending on $\left|\dot{\vec{\varphi}}\right|$ and mutual orientation of the vectors $\dot{\vec{\varphi}}$, $\vec{\Omega}$ and $\vec{A}$. The larger is the

angle between the vectors $\vec{A}$ and $\vec{\Omega}$ (within $\pm\pi/2$) and the smaller is the angle between the vector $\dot{\vec{\varphi}}$ and plane containing the vectors $\vec{A}$ and $\vec{\Omega}$, the less is the error.

15.3. A practical implementation of the method

15.3.1. Three-axis magnetometer alignment

One of the examples of the practical implementation of the method under consideration is a system for determining angular orientation of the three-axis magnetometer [5]. The magnetometer is intended for measuring the components of the Earth's magnetic field on board a moving metal ship whose own magnetic field can be practically of any level. If the angular motion of the ship is known to a sufficient accuracy, there is a possibility to separate the Earth's stationary field through the induced and the ship's own fields. In this case, by making use of the LGU rigidly connected with the magnetometer, it is possible to solve two problems simultaneously :

1. To determine the angular orientation of the magnetometer installed at different points on the ship, with respect to the central device, for example, attitude-and-heading reference device (AHRD), by determining static deformation;
2. To measure the angular motion of the magnetometer more precisely and in more details, taking into account elastic deformations of the ship.

In order to handle the first problem, determination of mutual angular orientation of the AHRD and the magnetometer, the method considered in Sections 3.1 and 3.2, should be slightly modified. The output signals from the AHRD and the LGU, located on the same base as the magnetometer, can be used instead of the readings from two LGUs. In this case it is necessary to determine the absolute angular velocity of the ship Ω (in Equation 15.3), using the AHRD's output readings. This angular velocity consists of the relative Ω_{rel} with respect to the Earth's velocity and the reference-frame velocity Ω_{rf} due to the Earth's rotation and the motion of the ship over the Earth's spherical surface. Transition from the geographic frame $o\xi\eta\zeta$ to the body frame $ox_b y_b z_b$ is carried out with the use of the matrix C according to Equation

$$[x_b, y_b, z_b] = C[\xi\eta\zeta],$$

where: $c = c_\theta c_\psi c_k$; k is the heading; ψ is the pitch angle; θ is the roll angle. The absolute angular velocity of the ship Ω will be

$$\Omega = \Omega_{rel} + C\Omega_{rf}, \qquad (15.23)$$

where:

$$\Omega_{rel} = C_\theta C_\psi \begin{vmatrix} 0 \\ o \\ -k \end{vmatrix} + C_\theta \begin{vmatrix} \dot{\psi} \\ 0 \\ 0 \end{vmatrix} + \begin{vmatrix} 0 \\ \dot{\theta} \\ 0 \end{vmatrix}, \qquad \Omega_{rf} = \begin{vmatrix} u_\xi \\ u_\eta \\ u_\zeta \end{vmatrix} = \begin{vmatrix} -\dot{\varphi} \\ (\Omega_E + \dot{\lambda})\cos\varphi \\ (\Omega_E + \dot{\lambda})\sin\varphi \end{vmatrix}.$$

φ, λ are the latitude and the longitude of the point. Further, the AHRD and LGU mutual angular orientation may be determined according to the procedure, described in Section 15.2, with the use of a Kalman filter.

As the AHRD does not accumulate errors in measuring roll and pitch angles, there is an additional possibility to estimate the LG drifts and improve the system accuracy. It takes a few minutes to de-determine the static components of the angular misalignment between the AHRD and LGU . If the LG drift model is assumed to be linear, the state vector of the system can be augmented to

$$x = \left| \varphi_x, \varphi_y, \varphi_z, \varepsilon_x, \varepsilon_y, \varepsilon_z, \right|^T,$$

where $\varepsilon_x, \varepsilon_y, \varepsilon_z$ are the drifts of the LG. The second problem, determination of the magnetometer's angular motion with respect to the geographic frame using the LGU, is solved in two stages. First, the LGU output data is corrected to eliminate drifts, and then substituting matrix A=BC for matrix C in Expression (15.23) we derive.

$$\Omega_{rel} = \Omega_{LGU} - A\Omega_{rf}$$

15.3.2. The results of the tests in the Baltic sea

The technique described above was tested under field conditions on a surveying vessel GS-403 in the Baltic Sea in 1993-1994. The measuring unit, consisting of a magnetometer and an LGU, was fixed at various points of the

ship, and the mutual angular orientation of the AHRD and LGU (with a magnetometer) and the angular motion of the magnetometer with respect to the geographic frame were measured for each of them. Fig. 15.4 shows the AHRD and magnetometer angular misalignment β with respect to the axis x of the ship as a function of time (upper curve) and the root-mean-square error of its estimation (lower curve).

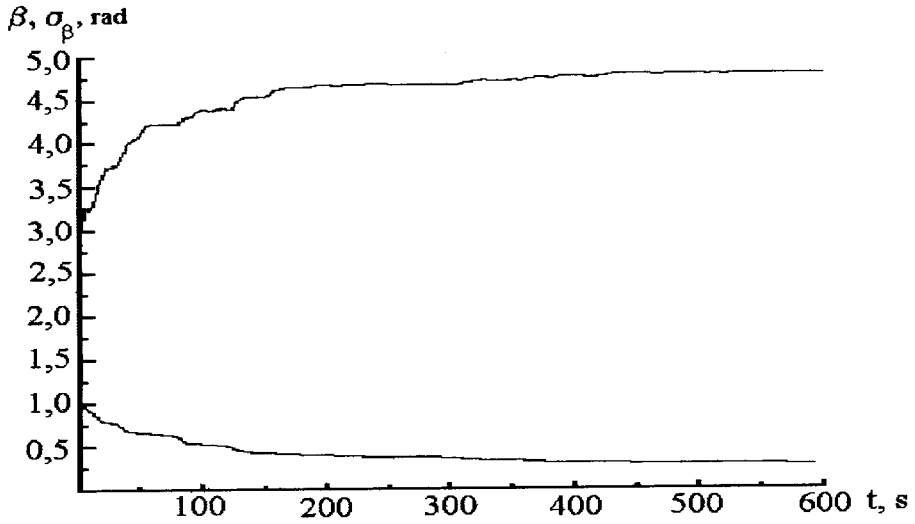


Fig. 15.4. AHRD and magnetometer angular misalignment β with respect to the axis x of the ship (upper curve) and the root-mean-square error of its estimation (lower curve)

The results of one of 10-minutes' measurements of AHRD and magnetometer angular misalignment relative to all the three axes are presented in Table 15.1.

Table 15.1 Estimations of angular misalignments and their errors.

Angular misalignment estimation, ...°			Estimation errors , ... °		
α	β	γ	σ_α	σ_β	σ_γ
-0.29	1.72	3.48	0.53	0.64	0.57

In Table 15.1 σ_i $(i=\alpha,\beta,\gamma)$ are the steady-state estimation errors at the end of the Kalman filter operational cycle. The repeated measurements of the same angular misalignment, taken at different time of the day and in different meteo-rological conditions, show good reproducibility of the result. The root-mean-square errors of twelve angular misalign-ment measurements of α ,β ,γ were 0,086°; 0.13°; 0,19°, correspondingly. Far greater measurement errors in com-parison with the results of the system simulation (Fig.15.3) may be due to simplification of the system model (neglect of the dynamic component of the deformation) and lower accuracy of the AHRD as compared to that of the LGU in the mode of angular velocity measuring (angular velocity was calculated as angle increments over the readout cycle).

As it was already mentioned at the beginning of this section, the LGU permits not only estimation of angular orientation of the magnetometer with respect to the AHRD, but also makes it possible to record more exactly dynamics of the ship's angular motion. The tests of the system at sea have shown that the accuracy in separation of the Earth's static magnetic field and the magnetic fields of the ship has increased 10 times through the use the LGU as compared to the use of the AHRD.

References

1. Okon I.M., Rivkin S.S., et al. Determination of Ship's Pitching and Rolling Parameters with Account of Deformations. Central Scientific Research Institute "Rhumb", 1989.
2. Alan M. Schnider. Kalman Filter Formulations for Transfer Alignment of Strapdown Inertial Units. Navigation, V.30, № 1, Spring 1983, p.72-89.
3. Okon I.M., Visegant I.B., Lukjanov D.P., Mochalov A.V., Brykov V.G. Three-axis Laser Gyroscopes Used in Measurement Tasks of Relative Angular Position of Shipborne Systems. Proc. of the Second Soviet-Chinese Symp., Leningrad, 9 - 15 October, 1991, p. 63-66.
4. Lipton A. Initial Alignment of Inertial Systems on Moving Objects. Moscow, "Nauka", 1971.
5. Blednov V.A., Korobejnikov A.G., Mochalov A.V., Kolchin D.L., Guzevich S.N. Use of a Laser Gyro and Accelerometer Unit for Determination of the Components of the Geomagnetic Field. Proc. of the Second Russian Conference "Present Condition and Problems of Navigation and Oceanography", St. Petersburg, 1995, p.116.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's References	3. Further Reference	4. Security Classification of Document
	RTO-AG-339 AC/323(SCI)TP/9	ISBN 92-837-1014-2	UNCLASSIFIED/ UNLIMITED

5. Originator

Research and Technology Organization
North Atlantic Treaty Organization
BP 25, 7 rue Ancelle, F-92201 Neuilly-sur-Seine Cedex, France

6. Title

Optical Gyros and their Application

7. Presented at/sponsored by

the Systems Concepts and Integration Panel (SCI) of RTO.

8. Author(s)/Editor(s)

D. Loukianov
R. Rodloff
H. Sorg
B. Stieler

9. Date

May 1999

10. Author's/Editor's Address

Multiple

11. Pages

340

12. Distribution Statement

There are no restrictions on the distribution of this document. Information about the availability of this and other RTO unclassified publications is given on the back cover.

13. Keywords/Descriptors

Gyroscopes
Laser gyroscopes
Ring lasers
Zeeman effect
Fiber optics
Accuracy
Utilization
Deformation
Depolarization
Drift (instrumentation)
Dynamic range
Flight control
Inertial navigation
Integrated optics
Manufacturing
Goniometers
Performance
Random walk
Sagnac effect
Signal processing
Measurement
Metrology
Surveying instruments
Railroad tracks

14. Abstract

This AGARDograph is written by leading experts from France, Germany, Russia and the United States. It consists of 15 chapters that can be divided into four main parts. The first one gives a detailed description of the laser and fiber optical gyro theory, the main sources of errors and the methods used to reduce their influence. The analysis of many types of modern laser gyros allowing various requirements to accuracy, dimension, weight and cost is presented in the second part. The reader can also find here some areas of practical applications of laser gyros. The third part is devoted to fiber optical gyros that have been developed recently and which are very successful technologically and commercially. The last part of the book considers special applications of optical gyros for laser dynamic goniometry and metrology, fundamental and applied research (such as nuclear physics, aerodynamics, etc.), railway track surveying and estimation of deformation of objects.

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION
BP 25 • 7 RUE ANCELLE
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

DIFFUSION DES PUBLICATIONS RTO NON CLASSIFIEES

L'Organisation pour la recherche et la technologie de l'OTAN (RTO), détient un stock limité de certaines de ses publications récentes, ainsi que de celles de l'ancien AGARD (Groupe consultatif pour la recherche et les réalisations aérospatiales de l'OTAN). Celles-ci pourront éventuellement être obtenues sous forme de copie papier. Pour de plus amples renseignements concernant l'achat de ces ouvrages, adressez-vous par lettre ou par télécopie à l'adresse indiquée ci-dessus. Veuillez ne pas téléphoner.

Des exemplaires supplémentaires peuvent parfois être obtenus auprès des centres nationaux de distribution indiqués ci-dessous. Si vous souhaitez recevoir toutes les publications de la RTO, ou simplement celles qui concernent certains Panels, vous pouvez demander d'être inclus sur la liste d'envoi de l'un de ces centres.

Les publications de la RTO et de l'AGARD sont en vente auprès des agences de vente indiquées ci-dessous, sous forme de photocopie ou de microfiche. Certains originaux peuvent également être obtenus auprès de CASI.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Fachinformationszentrum Karlsruhe
D-76344 Eggenstein-Leopoldshafen 2

BELGIQUE
Coordinateur RTO - VSL/RTO
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evère, B-1140 Bruxelles

CANADA
Directeur - Recherche et développement - Communications et gestion de l'information - DRDCGI 3
Ministère de la Défense nationale
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Defence Research Establishment
Ryvangs Allé 1, P.O. Box 2715
DK-2100 Copenhagen Ø

ESPAGNE
INTA (RTO/AGARD Publications)
Carretera de Torrejón a Ajalvir, Pk.4
28850 Torrejón de Ardoz - Madrid

ETATS-UNIS
NASA Center for AeroSpace Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320

FRANCE
O.N.E.R.A. (Direction)
29, Avenue de la Division Leclerc
92322 Châtillon Cedex

GRECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ISLANDE
Director of Aviation
c/o Flugrad
Reykjavik

ITALIE
Aeronautica Militare
Ufficio Stralcio RTO/AGARD
Aeroporto Pratica di Mare
00040 Pomezia (Roma)

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PAYS-BAS
NDRCC
DGM/DWOO
P.O. Box 20701
2500 ES Den Haag

PORTUGAL
Estado Maior da Força Aérea
SDFA - Centro de Documentação
Alfragide
P-2720 Amadora

ROYAUME-UNI
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

TURQUIE
Millî Savunma Başkanliği (MSB)
ARGE Dairesi Başkanliği (MSB)
06650 Bakanliklar - Ankara

AGENCES DE VENTE

NASA Center for AeroSpace Information (CASI)
Parkway Center
7121 Standard Drive
Hanover, MD 21076-1320
Etats-Unis

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
Royaume-Uni

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council
Document Delivery,
Montreal Road, Building M-55
Ottawa K1A 0S2
Canada

Les demandes de documents RTO ou AGARD doivent comporter la dénomination "RTO" ou "AGARD" selon le cas, suivie du numéro de série (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Des références bibliographiques complètes ainsi que des résumés des publications RTO et AGARD figurent dans les journaux suivants:

Scientific and Technical Aerospace Reports (STAR)
STAR peut être consulté en ligne au localisateur de ressources uniformes (URL) suivant:
http://www.sti.nasa.gov/Pubs/star/Star.html
STAR est édité par CASI dans le cadre du programme NASA d'information scientifique et technique (STI)
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
Etats-Unis

Government Reports Announcements & Index (GRA&I)
publié par le National Technical Information Service
Springfield
Virginia 2216
Etats-Unis
(accessible également en mode interactif dans la base de données bibliographiques en ligne du NTIS, et sur CD-ROM)

Imprimé par le Groupe Communication Canada Inc.
(membre de la Corporation St-Joseph)
45, boul. Sacré-Cœur, Hull (Québec), Canada K1A 0S7

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION
BP 25 • 7 RUE ANCELLE
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Telefax 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

DISTRIBUTION OFUNCLASSIFIED RTO PUBLICATIONS

NATO's Research and Technology Organization (RTO) holds limited quantities of some of its recent publications and those of the former AGARD (Advisory Group for Aerospace Research & Development of NATO), and these may be available for purchase in hard copy form. For more information, write or send a telefax to the address given above. **Please do not telephone**.

Further copies are sometimes available from the National Distribution Centres listed below. If you wish to receive all RTO publications, or just those relating to one or more specific RTO Panels, they may be willing to include you (or your organisation) in their distribution.

RTO and AGARD publications may be purchased from the Sales Agencies listed below, in photocopy or microfiche form. Original copies of some publications may be available from CASI.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Coordinateur RTO - VSL/RTO
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evère, B-1140 Bruxelles

CANADA
Director Research & Development
Communications & Information
Management - DRDCIM 3
Dept of National Defence
Ottawa, Ontario K1A 0K2

DENMARK
Danish Defence Research Establishment
Ryvangs Allé 1, P.O. Box 2715
DK-2100 Copenhagen Ø

FRANCE
O.N.E.R.A. (Direction)
29 Avenue de la Division Leclerc
92322 Châtillon Cedex

GERMANY
Fachinformationszentrum Karlsruhe
D-76344 Eggenstein-Leopoldshafen 2

GREECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ICELAND
Director of Aviation
c/o Flugrad
Reykjavik

ITALY
Aeronautica Militare
Ufficio Stralcio RTO/AGARD
Aeroporto Pratica di Mare
00040 Pomezia (Roma)

LUXEMBOURG
See Belgium

NETHERLANDS
NDRCC
DGM/DWOO
P.O. Box 20701
2500 ES Den Haag

NORWAY
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PORTUGAL
Estado Maior da Força Aérea
SDFA - Centro de Documentação
Alfragide
P-2720 Amadora

SPAIN
INTA (RTO/AGARD Publications)
Carretera de Torrejón a Ajalvir, Pk.4
28850 Torrejón de Ardoz - Madrid

TURKEY
Millî Savunma Başkanliği (MSB)
ARGE Dairesi Başkanliği (MSB)
06650 Bakanliklar - Ankara

UNITED KINGDOM
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

UNITED STATES
NASA Center for AeroSpace Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320

SALES AGENCIES

NASA Center for AeroSpace Information (CASI)
Parkway Center
7121 Standard Drive
Hanover, MD 21076-1320
United States

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
United Kingdom

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council
Document Delivery,
Montreal Road, Building M-55
Ottawa K1A 0S2
Canada

Requests for RTO or AGARD documents should include the word 'RTO' or 'AGARD', as appropriate, followed by the serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Full bibliographical references and abstracts of RTO and AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
STAR is available on-line at the following uniform resource locator:
http://www.sti.nasa.gov/Pubs/star/Star.html
STAR is published by CASI for the NASA Scientific and Technical Information (STI) Program
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
United States

Government Reports Announcements & Index (GRA&I)
published by the National Technical Information Service
Springfield
Virginia 22161
United States
(also available online in the NTIS Bibliographic Database or on CD-ROM)

Printed by Canada Communication Group Inc.
(A St. Joseph Corporation Company)
45 Sacré-Cœur Blvd., Hull (Québec), Canada K1A 0S7

ISBN 92-837-1014-2